AERONAUTICS AND ASTRONAUTICS

Edited by **Max Mulder**

INTECHWEB.ORG

Aeronautics and Astronautics
Edited by Max Mulder

Published by InTech
Janeza Trdine 9, 51000 Rijeka, Croatia

Copyright © 2011 InTech
All chapters are Open Access articles distributed under the Creative Commons Non Commercial Share Alike Attribution 3.0 license, which permits to copy, distribute, transmit, and adapt the work in any medium, so long as the original work is properly cited. After this work has been published by InTech, authors have the right to republish it, in whole or part, in any publication of which they are the author, and to make other personal use of the work. Any republication, referencing or personal use of the work must explicitly identify the original source.

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published articles. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

Publishing Process Manager Mia Devic
Technical Editor Teodora Smiljanic
Cover Designer Jan Hyrat
Image Copyright hunta, 2010. Used under license from Shutterstock.com

First published August, 2011
Printed in Croatia

A free online edition of this book is available at www.intechopen.com
Additional hard copies can be obtained from orders@intechweb.org

Aeronautics and Astronautics, Edited by Max Mulder
p. cm.
ISBN 978-953-307-473-3

INTECH OPEN ACCESS PUBLISHER

INTECH open

free online editions of InTech
Books and Journals can be found at
www.intechopen.com

Contents

Preface IX

Part 1 Aerodynamics 1

Chapter 1 **Visualization of Complex Flow Structures by Matched Refractive-Index PIV Method 3**
Kazuhisa Yuki

Chapter 2 **Plasma Flow Control 21**
Ying-hong Li, Yun Wu, Hui-min Song, Hua Liang and Min Jia

Chapter 3 **Nonequilibrium Plasma Aerodynamics 55**
Andrey Starikovskiy and Nickolay Aleksandrov

Chapter 4 **Numerical Investigation of Plasma Flows Inside Segmented Constrictor Type Arc-Heater 97**
Kyu-Hong Kim

Chapter 5 **Physico - Chemical Modelling in Nonequilibrium Hypersonic Flow Around Blunt Bodies 125**
Ghislain Tchuen and Yves Burtschell

Chapter 6 **A Frequency-Domain Linearized Euler Model for Noise Radiation 159**
Andrea Iob, Roberto Della Ratta Rinaldi and Renzo Arina

Chapter 7 **High-Order Numerical Methods for BiGlobal Flow Instability Analysis and Control 185**
Javier de Vicente, Daniel Rodríguez, Leo González and Vassilis Theofilis

Part 2 Flight Performance, Propulsion, and Design 235

Chapter 8 **Rotorcraft Design for Maximized Performance at Minimized Vibratory Loads 237**
Marilena D. Pavel

Chapter 9 **Concurrent Subspace Optimization for Aircraft System Design 257**
Ke-shi Zhang

Chapter 10 **The Assessment Method for Multi-Azimuth and Multi-Frequency Dynamic Integrated Stealth Performance of Aircraft 279**
Ying Li, Jun Huang, Nanyu Chen and Yang Zhang

Chapter 11 **Aircraft Gas-Turbine Engine's Control Based on the Fuel Injection Control 305**
Alexandru-Nicolae Tudosie

Chapter 12 **Plasma-Assisted Ignition and Combustion 331**
Andrey Starikovskiy and Nickolay Aleksandrov

Chapter 13 **O_2/CH_4 Kinetic Mechanisms for Aerospace Applications at Low Pressure and Temperature, Validity Ranges and Comparison 369**
Angelo Minotti

Part 3 Materials and Structures 401

Chapter 14 **Creep Behaviors and Influence Factors of FGH95 Nickel-Base Superalloy 403**
Tian Sugui and Xie Jun

Chapter 15 **Multi-Dimensional Calibration of Impact Models 441**
Lucas G. Horta, Mercedes C. Reaves, Martin S. Annett and Karen E. Jackson

Part 4 Avionics, Control and Operations 459

Chapter 16 **An Agile Cost Estimating Methodology for Aerospace Procurement Operations: Genetic Causal Cost CENTRE-ing 461**
R. Curran, P. Watson and S. Cowan

Chapter 17 **Developing Risk Models for Aviation Inspection and Maintenance Tasks 487**
Lee T. Ostrom and Cheryl A. Wilhelmsen

Chapter 18 **Novel Digital Magnetometer for Atmospheric and Space Studies (DIMAGORAS) 499**
George Dekoulis

Chapter 19 **Aeronautical Data Networks 515**
Mustafa Cenk Erturk, Wilfrido Moreno, Jamal Haque and Huseyin Arslan

Chapter 20 **Air Traffic Control Decision Support for Integrated Community Noise Management 533**
Sander J. Hebly and Hendrikus G. Visser

Chapter 21 **A Conceptual Framework and a Review of Conflict Sensing, Detection, Awareness and Escape Maneuvering Methods for UAVs 549**
B. M. Albaker and N. A. Rahim

Chapter 22 **Collision Probabilities, Aircraft Separation and Airways Safety 571**
Luís Campos and Joaquim Marques

Chapter 23 **Development of an Aircraft Routing System for an Air Taxi Operator 589**
F.M. van der Zwan, K. Wils and S.S.A. Ghijs

Preface

In the next decades, aerospace will be facing significant challenges. Important themes are the reduction of environmental impact and the increase in efficiency of aerospace operations. The most important and daunting challenge, however, is to design aircrafts and spacecrafts that, ultimately, do not depend on fossil fuels. The dependency of aerospace operations on fossil fuels, which is currently almost 99%, needs to be dramatically reduced. The availability of fossil fuels will be significantly decreased by 2035, increasing their cost. Given the fact that aerospace innovations typically take 20 to 30 years to move from a laboratory setting to a real-life, safe implementation, we have to hurry in developing radically novel aircrafts and spacecrafts, and the enabling technology.

This book captures many of the ongoing investigations to develop more sustainable and environmentally-friendly aerospace vehicles. It contains twenty-three chapters organized in four main sections: aerodynamics, flight performance and propulsion, materials and structures, and avionics and operations. Throughout these sections, the research presented is often geared towards radical innovations that may well be the basis for the new era of aerospace operations.

The section on aerodynamics covers subjects ranging from the visualization of complex flow using particle image velocimetry techniques, to the reduction of aerodynamic drag through plasma flow actuation techniques. In the propulsion and performance section, the chapters range from helicopter performance improvements through design, the improvement of gas turbine techniques and exhaust measurements, to radically new forms of propulsion. The section on structures and materials, novel metallic alloys with increased performance regarding creep and fracture, and the development of better models for calculating the impact of crashes. Many of these innovations are mandatory to design the next generation of aerospace vehicles that allow for a sustainable air transportation.

In the near future, changes in how to operate aerospace vehicles more effectively and efficiently may be implemented. Novel air traffic management concepts, aircraft routing schemes and methods, new avionics sensors and aeronautical data networks facilitating a system-wide information management are on the drawing table and likely to be implemented in the next decade, even before the ‘new’ aerospace vehicles become real.

The book clearly illustrates that the next generation aerospace vehicles and their operation require a multi-disciplinary approach, ranging from pure aerodynamics to operations research. After one hundred years of developments and the maturation of the aerospace domain from a pioneering activity into an established, indispensable field of study which enables our daily life activities, we now face an incredible challenge indeed.

I hope you like the book. I would like to thank all authors for their efforts and assistance in completing the book. Special thanks to the InTech team for their help, great editing job, and for making this book possible in the first place.

Max Mulder
Department Control & Operations
Faculty of Aerospace Engineering
Technical University Delft
The Netherlands

Part 1

Aerodynamics

1

Visualization of Complex Flow Structures by Matched Refractive-Index PIV Method

Kazuhisa Yuki
Tokyo University of Science, Yamaguchi
Japan

1. Introduction

With the development of computers and their surrounding equipments, the simulation of complicated flow structures around aircrafts will further become easier and cheaper by applying computational fluid dynamics. However, in order to judge whether the flow field obtained is reasonable or not, turbulent models and/or numerical schemes should be selected based on the comparison with experimental results. On the other hand, three-dimensional measurement of unsteady flow structures especially around obstacles with complicated geometry is still difficult due to some problems. For instance, where a three-dimensional flow structures around obstacles is visualized by a PIV technique, it is extremely difficult to grasp the whole flow-structure including the flow behind the obstacles even if transparent materials are used, because the difference of the refractive index between the working fluid and the transparent material causes distortion in the image. Therefore, in this chapter, I introduce a special visualization technique to match the refractive index of the working fluid with that of the transparent material that is called "matched refractive-index PIV measurement" and show some complicated flow fields visualized by this technique.

2. PIV visualization utilizing a matched refractive-index method

2.1 Refractive-index adjustment of NaI solution

Where the whole three-dimensional flow structure around obstacles is visualized by a PIV technique, it is necessary to match the refractive index of the working fluid with that of the obstacle material. This research employs a sodium iodide solution (NaI solution), which is easy to handle and chemically stable, as the working fluid. This solution is deliberately chosen in order to be able to adjust the refractive index of the working fluid to that of the acrylic obstacle with the index of 1.49. Normally the refractive index of this solution is not so sensitive to temperature change, so that the refractive index of the NaI solution is adjusted by changing its concentration. Figure 1 shows a light path difference caused by refraction, where a YAG laser used in the PIV measurement is irradiated to an acrylic cylinder of 30mm in diameter fixed at a center in a 10cm square acrylic box filled with the NaI solution at 30 degrees Celsius. The light path difference, δ, is measured at a location of 660mm from the back of the cylinder. The difference decreases with the increase in the NaI concentration and reaches zero at 61.6wt%. That means that the refractive index of the NaI solution completely corresponds with that of the acrylic cylinder at this concentration. In actual visualization

experiments, a refractive index at this concentration under visible light, which is 1.485, was always checked by using a portable refractometer before each experiment, because the change in the refractive index might be caused by deposition of NaI crystals onto the pipe wall and/or volatilization from the solution.

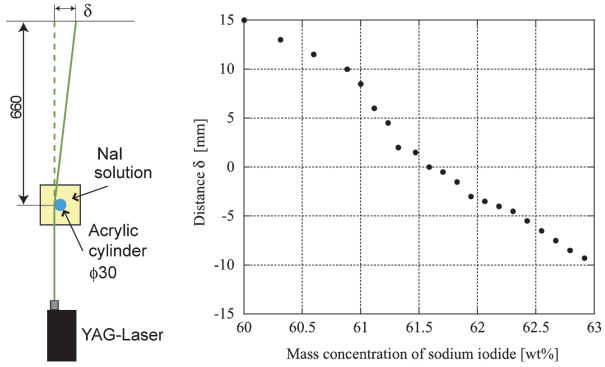


Fig. 1. Matched refractive-index experiment using NaI

2.2 PIV measurement with fluorescence particles

The PIV utilized in this experiment is a double-pulse YAG laser system manufactured by Japan Laser Corporation. The laser output is 25mJ@532nm and the maximum oscillatory frequency is 30Hz. In the PIV measurement, a time series of tracer particles' images in a sheet laser is taken with a high speed camera, and, then, a two-dimensional flow structure is quantitatively visualized from the movement of the tracer particles. The time interval of the double pulse and the tracer concentration are adjusted depending on the flow conditions. To process the obtained particle images, a cross-correlation scheme is adopted to get spatially dense velocity information. Furthermore, melamine fluorescence resin particles with 1~20 μm diameter are utilized as the tracer particles. The specific gravity of NaI solution at the above mentioned concentration is relatively close to that of this tracer, so that buoyancy influence can be ignored. When this fluorescence particle is irradiated with the YAG laser, it causes excitation in the fluorescent agent which emits light of 580nm wavelength. By taking only this newly emitted light into a CCD camera with an attached filter lens, it makes it possible to obtain a clearer particle image than the usual tracer image, because the diffused reflection light of the laser observed on the pipe wall surface and on the acrylic sphere surface can be completely removed simultaneously.

3. Experimental apparatus and details of test section

Figure 2 shows a diagram of the apparatus for the visualization experiment under isothermal conditions. The apparatus consists of the following components: a circulating pump, a flow rate measuring section, a flow-straightened section, a test section, a bag filter, and a mixing tank. All piping materials and components have been made of polyvinyl chloride or acrylic materials etc. which co-exist in a stable state with the NaI solution. The magnetic pump circulates the working fluid inside the loop, and its maximum flow rate under obstacle-unpacked conditions is approximately 200l/min. The flow rate of the working fluid is adjusted

by valves: two valves located between the magnetic pump and the flow rate measuring section and a valve of a bypass line which directly returns to the mixing tank from the magnetic pump. A turbine flowmeter or an ultrasonic Doppler velocimeter is utilized to measure the flow rate. The mixing tank has the following functions: injection of tracer particles, de-aeration of bubbles existing in the fluid, and heat exchange to control the fluid temperature. The section upstream of the test section has a flow-straightener with a honeycomb structure consisting of stainless steel pipes, which straightens and counteracts a swirling flow formed in the bend upstream. The bug filter is a polypropylene-made cartridge with strong corrosion resistance which separates the tracer particles from the NaI solution.

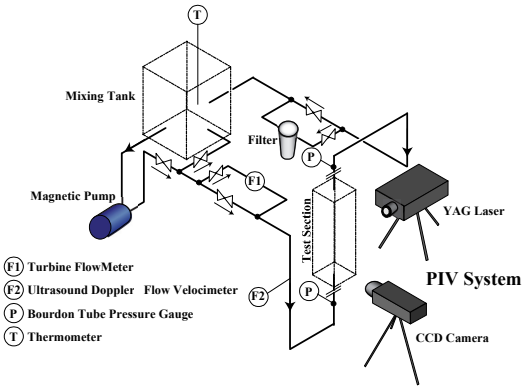


Fig. 2. Experimental apparatus for visualization

Figure 3 shows a detailed view of the visualization test section. Here, we focus here on the flow field in a Sphere-Packed Pipe (hereafter, SPP) that is utilized as a heat exchanger and/or a cooling device in various fields (e.g. Yuki et al. 2007, 2008). The test section is an acrylic vertical riser-pipe with D=56mm as inner diameter and 670mm in length. The visualizing area is located at 8.2D (=460mm) downstream from an inlet of the test section where a fully-developed flow is anticipated. In addition, there is a rectangular jacket surrounding the test section in order to reduce image distortion resulting from the geometry of the circular pipe. The NaI solution at the same temperature as the working fluid is also filled into the jacket. In order to visualize the flow field in the lateral cross section of the circular pipe, an acrylic observation window is attached to the upper part of the test section. Figure 2 also shows the packing structure of the acrylic spheres. The sphere size prepared for this research is D/2.0 (27.6mm) in diameter, and 68 spheres can be packed in the test section with a porosity of 0.548. An acrylic baffle plate set between the flanges, which exist at the inlet and outlet of the test section, fixes the acrylic spheres.

The temperature of the NaI solution is 30 degrees Celsius, and the visualization of the flow field is conducted at three Reynolds numbers (Re_d=Ud/ν) of 800, 2000 and 4900, based on the sphere diameter, d, and mean inlet-velocity U. The mean inlet-velocity, which is equivalent to superficial velocity in the SPP, is 0.0376, 0.0940 and 0.230 m/s, respectively. Fand et al. (1990, 1993a, 1993b) have classified the SPP pipe flow with D/d>1.4 into: turbulent regime (Re_d>120), Forchheimer regime (5<Re_d<80), and Darcy regime (Re_d<2.3) from the data of friction factor. In that sense, the flow conditions in this research could be in the turbulent regime. This fact was also confirmed by the authors' experiments (Yuki et al., 2005).

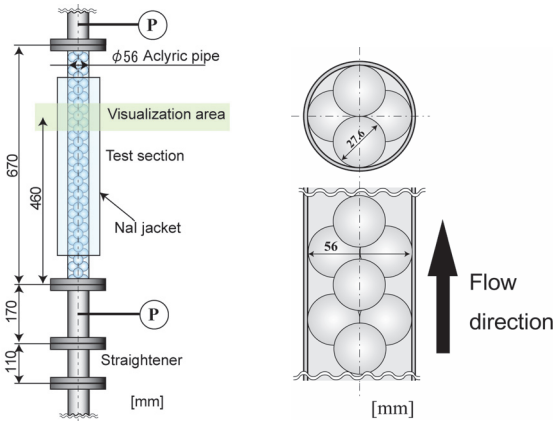


Fig. 3. Test section and packing structure

4. Flow-structures in sphere-packed pipe

4.1 Flow structures in a central longitudinal-section

First, a central longitudinal section S_1 shown in Figure 4 is visualized. Figure 5 shows a time-series of the flow fields at the lowest Re number of Re_d=800. The matched refractive-index method makes it possible for the flow field in the central area of the pipe, which is usually impossible to see, to be successfully and vividly visualized. The whole flow field fluctuates intensely and is extremely unsteady. In order to discuss the flow field more qualitatively, the time-averaged flow field is shown in Figure 6(a). This clearly provides that a high-velocity flow spouting from the upstream of a central area A (spouting flow, hereafter) is pushed back toward the center area again from around the middle of the center area and the wall. Furthermore, judging from the instantaneous flow fields together, a flow toward the pipe wall, which is considered as a part of a wake, forms a strong impinging flow to the pipe wall. After this impingement, a backward-flow in the upstream direction forms a circular vortex in the area B between the sphere and the wall. It can be also confirmed that there exists a small circular vortex behind the impinging flow. An interesting feature is the formation of a low-velocity area in the middle of the spouting flow and the impinging flow, which seems to capture a coexisting area of different kinds of several flows in the SPP. In addition, there is stagnation in the area C closely located to the contact point between the sphere and the wall. If the SPP is utilized

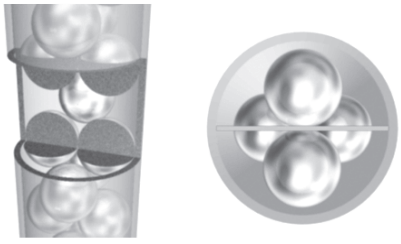


Fig. 4. Visualizing section: S_1

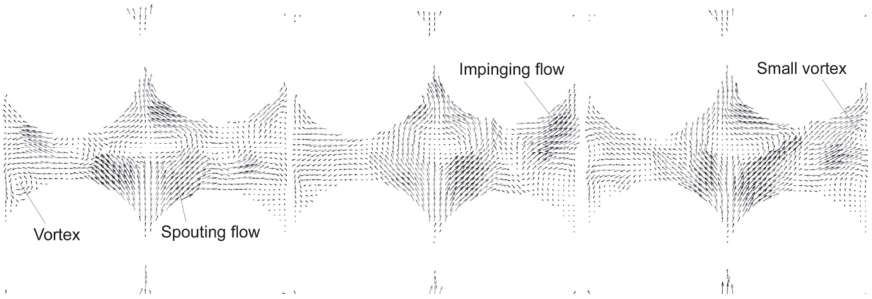


Fig. 5. A time-series of flow fields in the section S_1 (Re_d=800, Time interval=0.07sec.)

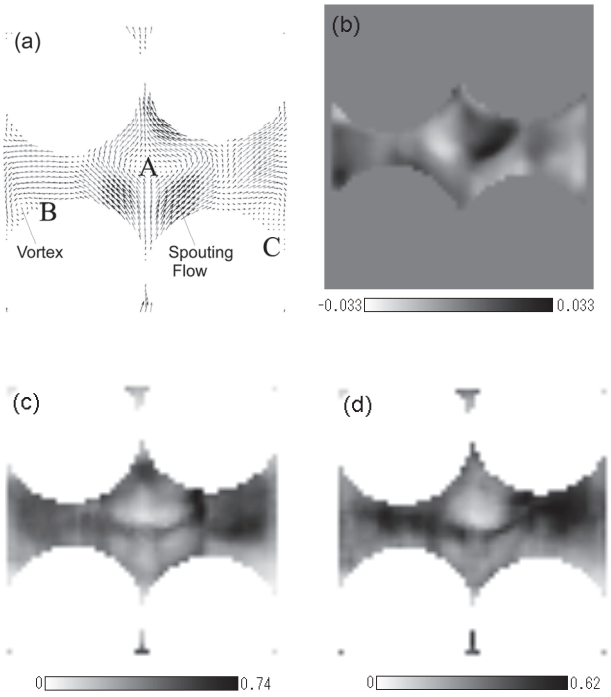


Fig. 6. Time-averaged flow field (a), vorticity (b), and intensity of velocity fluctuation (c): U_p & (d): V_p

as a heat transfer promoter, it is impossible to remove enough heat in this area under the heating conditions, so that the wall-temperature could rise sharply here, especially for higher Pr number fluid. However, the area downstream of this stagnation seems to have a high heat transfer performance due to the effect of the impinging flow. Figure 6(b) shows vorticity distribution for the time-averaged flow field. The regions around A and near the wall with high vorticity hold a high heat transport performance. Hence, the spouting flow

around the central area and its returning effect efficiently transport enthalpy that was transported from the heating wall by the circular vortex toward the center of the pipe, which could work more effectively if there existed active heat conduction from the wall to the sphere. Figures 6(c) and (d) respectively represent intensities of velocity fluctuation in the radial and streamwise directions, U_p and V_p, which are calculated by RMS and normalized by each inlet velocity U. A strong velocity-fluctuation area is formed upstream of the sphere existing downstream. Their maximum values over 0.6 indicate that the intensity of this variation is quite high. These fluctuations could significantly contribute to the heat transport from the wall, because both the areas with high velocity fluctuation spread toward the pipe wall.

Figures 7 present two kinds of profiles at the horizontal center line of the visualizing area: the one is the averaged velocity profile about u and v in the radial and streamwise directions and the other is the RMS profiles of velocity fluctuation U_p and V_p. The radial velocity near the wall indicates the strength of the impinging flow and is higher than the inlet velocity. The velocity profiles at Re_d=4900 differ from the other cases at Re_d=800 and 2000, and indicates a lower value, especially near the wall. This is conceivably because the impinging flow area moves more downstream at Re_d=4900. As to the streamwise velocity profile, the spouting flow around the center area is accelerated as the Re_d number increases. Both figures show that the velocity peak position near x/R=±0.25 shifts toward the center of the pipe with increasing Re_d number, which suggests the prevention of the spread of the spouting flow and also suggests the development of something strong flow pushing the spouting flow back. Additionally, there are two areas with high velocity fluctuation near x/R=±0.5 for each Re_d number, which indicates that the above-mentioned two characteristic flow-structures themselves have strong variation and are closely linking with each other

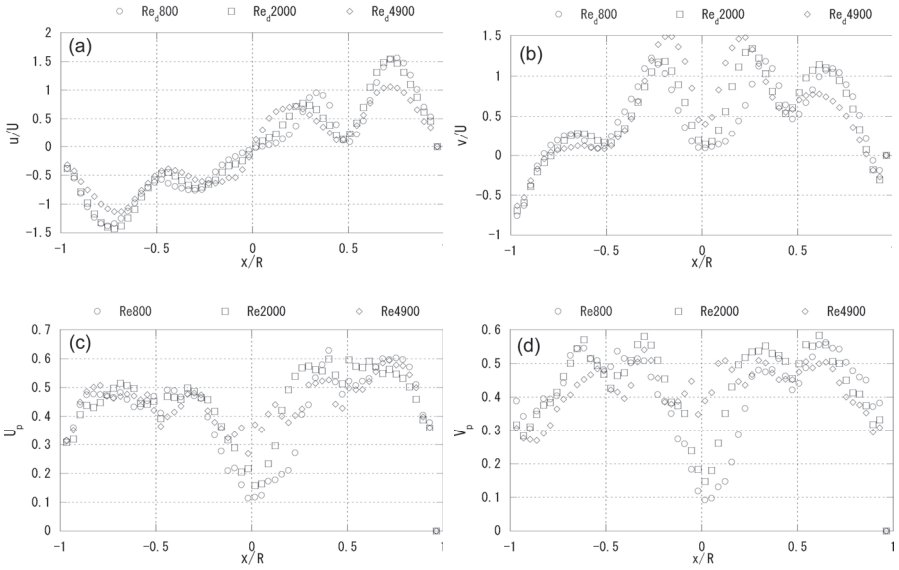


Fig. 7. Mean velocity and velocity-fluctuation intensity profiles in the radial and streamwise directions

because of there not being much difference in the variation in both the flow structures. In all data, there exists low flow velocity and a low velocity fluctuation area at the location of x/R=±0.5. This location corresponds to the coexistence area for the two flow structures that seems to exist around the right and left sides of the sphere existing in front and back of the A area. The flow structure in this coexistence area will be clarified later by comparison with the flow structures on the other cross sections.

4.2 Flow structure of bypass flow due to wall effect

Since the present SPP has four wide gap-channels in the streamwise direction due to the existence of the pipe wall, the flow characteristics of a high-velocity channeling flow, that is to say a bypass flow formed in this area, strongly affect not only the flow structures of the spouting flow and the vortices behind the sphere but also heat transfer characteristics. To capture this flow, a longitudinal section S_2 shown in Figure 8 is visualized. Figures 9(a) through (d) are the distributions of time-averaged flow field, vorticity, and the RMS of velocity fluctuations, respectively, in the right half area of the S_2 at the Re_d of 800. The bypass flow flowing in parallel with the pipe wall is observed in the gap area between the sphere and the pipe wall. The wake area behind the sphere has high vorticity, because the flow direction shifts to the back from the front of the paper depending on the packing structure, as well as the influence of the wake. In addition, the RMS of velocity fluctuation, U_p and V_p, are more overwhelmingly intense than those in the case of the above mentioned longitudinal section S_1, which indicates strong unsteadiness of the flow field in this section, especially beside and behind the sphere. A time-series of the flow fields allow us to confirm that a circular vortex or a flow with low velocity is intermittently formed in this area (note: this is caused by strong variation at the bend of the bypass flow described later on). To quantitatively evaluate the influence of the Re_d number on the flow structure of the bypass flow, Figures 10(a) through (d) show the flow velocity profiles and the RMS of velocity fluctuation in the radial and streamwise directions at the horizontal center line in the middle of S_2. Minus value of the radial flow velocity means that the sphere shape contributes to a certain amount of momentum transfer toward the center of the pipe. Additionally, Figure 10 (b) shows that the bypass flow has quite a high velocity, whose maximum value is around 6 times that of the mean velocity. The Re_d number doesn't seem to affect the value much. Judging from this, it is evident that the bypass flow could be the something strong flow for the spouting flow being pushed back and for producing some vortices including the impinging flow. Though the RMS of velocity fluctuation in both the radial and streamwise directions goes up as the Re_d number increases, the data at Re_d=800 shows a peculiar profile. It is considered that the flow field at Re_d=800 still has the characteristics in lower Re_d number regime where there exists a different kind of vortex shedding behind the sphere.

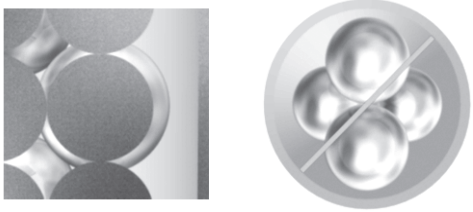


Fig. 8. Visualizing section: S_2

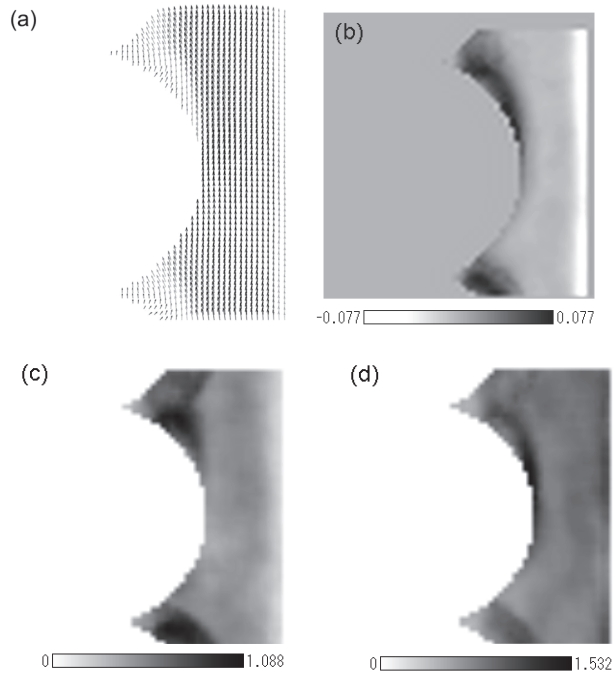


Fig. 9. Time-averaged flow field, vorticity, and RMS of velocity fluctuation of bypass flow

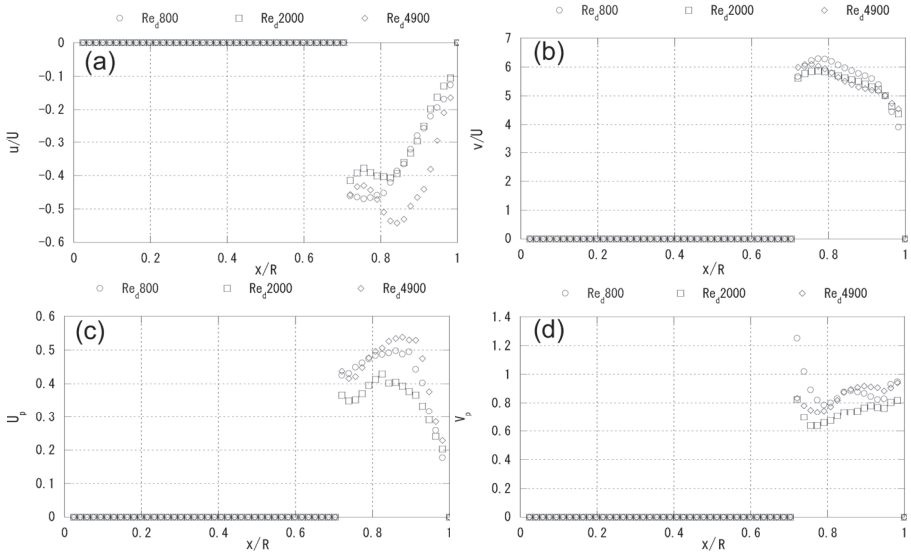


Fig. 10. Mean velocity and RMS profiles in the radial and streamwise directions

To evaluate the three-dimensional structure of the bypass flow, the flow field in the longitudinal section S_3 perpendicularly to S_2 is focused on (see Figure 11). Figures 12(a) through (d) show the same distributions as those of Fig.11. The bypass flow is flowing with a meandering motion through the spheres in a circumferential direction. Furthermore, a circular vortex is formed between the upstream and downstream spheres. As this circular vortex is formed near the downstream sphere, the flow that is branched by the meandering bypass flow colliding with the downstream sphere significantly plays a role in forming the vortex. In addition, the existence of a high-vorticity area A being released from the sphere represents the boundary of the bypass-flow route. Regarding the velocity fluctuation U_p in the horizontal direction of S_3, the area with high velocity fluctuation is formed on the bypass-flow route, which means that the bypass flow itself unsteadily changes in the circumferential direction. Furthermore, an area B with high velocity fluctuation is formed forward of the downstream sphere. As already described, a part of the bypass flow collides with the downstream sphere due to its meandering motion. In such areas, the flow becomes unstable and increases the intensity of velocity fluctuation by the influence of the sphere geometry. Regarding the intensity of velocity fluctuation V_p in the vertical direction of S_3, an area C with high velocity fluctuation exists beside the bypass-flow route between the upstream and downstream spheres. The flow loses stability in this large gap area by being released from the influence of the upstream sphere-surface, which could also lead to the formation of the impinging flow mentioned before though this flow can't be confirmed in this section S_3. Figure 13 is a time-series of the flow fields at Re_d=4900. Looked at through the time-averaged flow field characteristics, the flow structures don't show such significant change as is the case at Re_d=800. However, the time-series of the flow fields show that two apparent circular vortices are formed between the downstream and upstream spheres. In other words, the wake structure in the large gap area between the spheres is characterized by the co-existence of the circular vortex formed by the colliding effect of the bypass flow with the downstream sphere and by the separation vortex shedding from the upstream sphere. In particular, the small vortex shown in Fig. 5 corresponds to a part of this downstream vortex. The existence of these two vortices has been also confirmed at Re_d=800. In the SPP flow, it is quite interesting that several vortices exist between the upstream and downstream spheres. In addition, the bypass-flow seems to meander in larger area with increasing Re_d number as shown in Figures 14(c) and (d), which show the RMS of the velocity fluctuation in the horizontal and vertical directions at the horizontal center line of S_3. In Fig. 14(a), the horizontal velocity indicates a minus value because of the separation vortex near the wall. Observing from the streamwise velocity profiles in Fig. 14(b), the maximum flow velocity of the bypass flow is almost 5 through 6 times higher than the mean velocity, regardless of the Re_d number. This fact accords with the result shown in Figure 10(b). Moreover, the central axis of the bypass flow also seems not to be influenced by the Re_d number, because the peak location of the streamwise velocity doesn't shift for the change in the Re_d number.

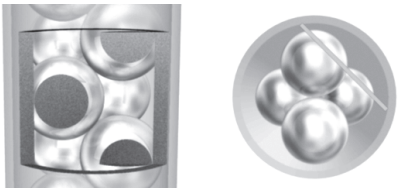


Fig. 11. Visualizing section: S_3

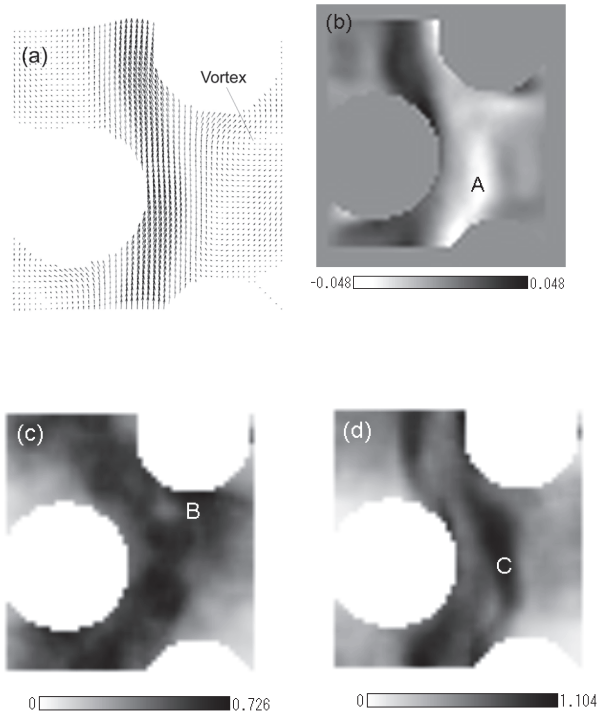


Fig. 12. Time-averaged flow field, vorticity, and RMS of velocity fluctuation of bypass flow

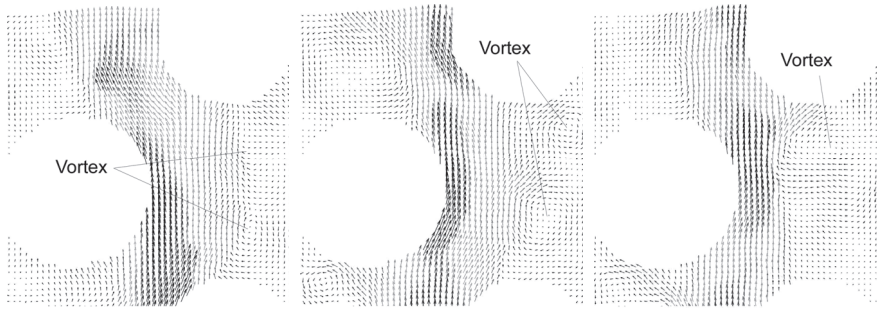


Fig. 13. Time sequence of flow fields of bypass flow (Re_d=4900, Time interval=0.07sec.)

4.3 Flow structure behind the sphere

To identify further the wake structures formed between the spheres, the longitudinal section S_4 shown in Figure 15 is visualized. Figure 15 also show instantaneous flow fields at Re_d=4900, and Figures 16(a) through (c) show the distributions of vorticity and intensities of velocity fluctuation in the horizontal and vertical directions of S_4, respectively. The high-

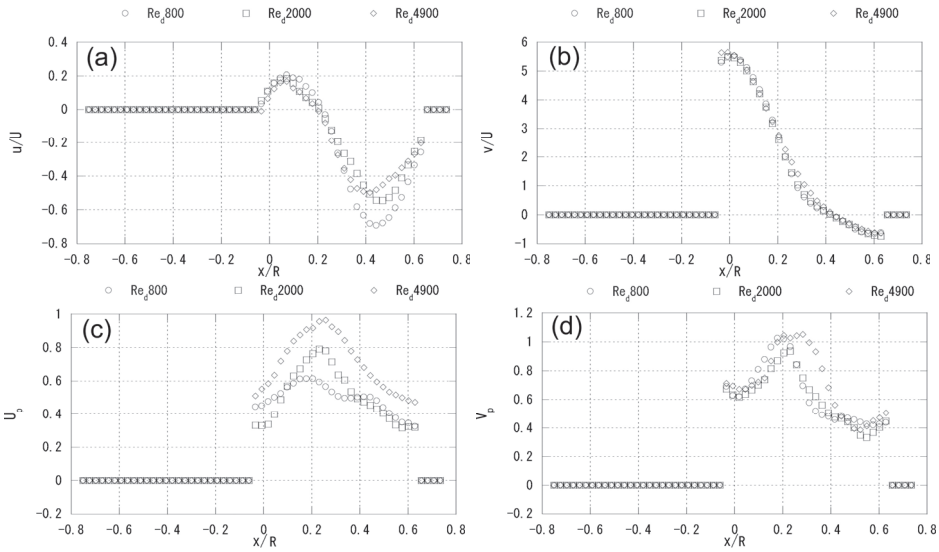


Fig. 14. Mean velocity and velocity-fluctuation intensity profiles in horizontal and vertical directions of S_3

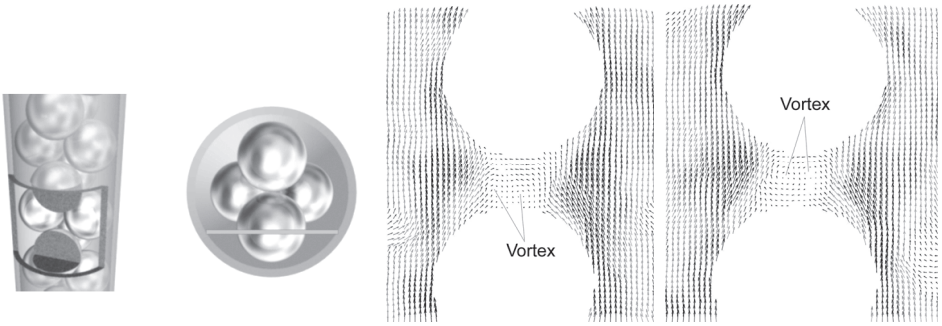


Fig. 15. Visualizing section: S_4 and instantaneous flow fields behind sphere (Re_d=4900, Time interval=0.07sec.)

velocity flow existing in the gap between the sphere and the pipe wall is the bypass flow itself. Observation through the instantaneous flow fields confirms the generation and disappearance of a circular vortex such as Karman-like twin vortices behind the sphere that is strongly affected by the inflowing of a part of the bypass flow. According to this inflowing, an area with high vorticity exists behind the sphere as shown in Fig. 16(b). The shape and behavior of these vortices are another aspect of the separation vortex mentioned above. However, as it is difficult to visualize both the two kinds of vortices simultaneously, the downstream circular vortex, which should be also twin vortices, seems to be somewhat flat in the circumferential direction and be formed near the pipe wall as shown in Fig. 5. In addition, the gap area between the spheres has high velocity fluctuation, this also indicates strong time variation of these two pairs of twin vortices. These vortices don't affect the

bypass flow structure and its behavior significantly, because the area with high velocity fluctuation by the wake does not spread out to the bypass flow area. To sum up the matter, there are two factors that form the structure of the wake between the spheres and their unsteady behavior (see Fig. 16(d) that is illustrations of flow structures in S_3 and S_4): 1, Twin separation vortices shedding from the sphere in the process that a part of each high-velocity bypass flow passing beside the sphere flows into the large gap area located behind the sphere (this process also generates the strong impinging flow as mentioned later on); 2, Twin circular vortices generated by the colliding and branching effects of two bypass flows in front of the sphere in downstream.

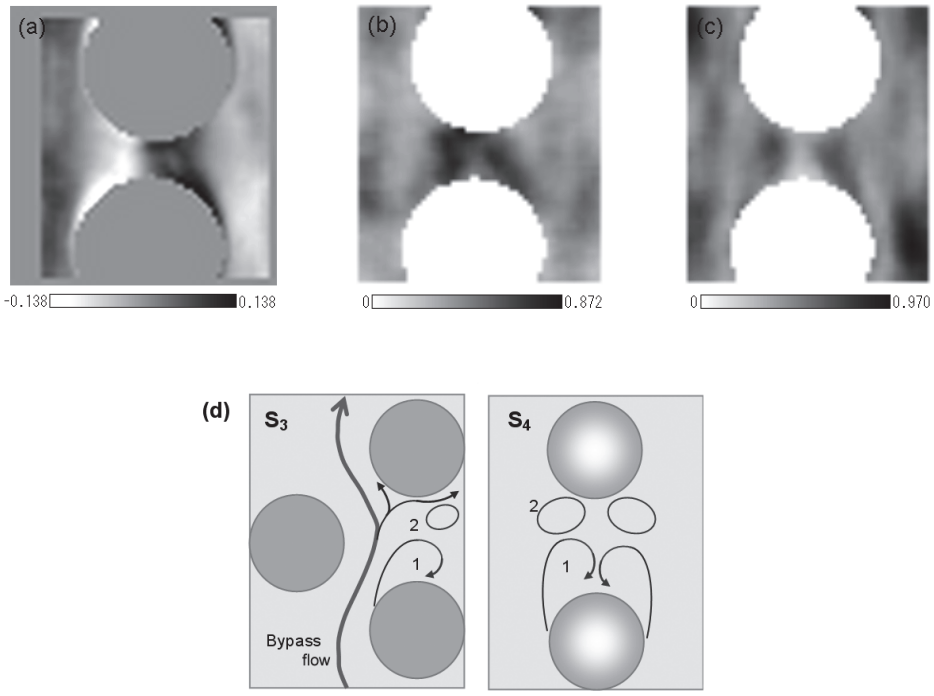


Fig. 16. Vorticity & velocity fluctuations (Re_d=4900), and illustration of flow structure in S_3 & S_4

4.4 Flow structure in a cross section of circular pipe

Finally, the flow fields in a lateral cross section of the pipe, i.e. a secondary flow, are visualized in order to investigate the three-dimensional structure of the flow field in the SPP. The matched refractive-index PIV method only enables to observe the flow field in this cross section. Figure 17 shows a time-series of the secondary flows at the Re_d numbers of 800 and 4900 in the lateral cross section, S_5, perpendicular to the longitudinal section S_1 (see Fig. 17 on the lower right). These figures clearly show that apparent twin vortices with strong variation exist in the lateral cross section. The four circled areas in the time-

averaged flow field at Re_d=800 in Fig. 17 on the upper right are the ones for the bypass flow passing through. As already mentioned, the bypass flow meanders along the sphere geometry and a part of it flows into the large gap area between the upstream and downstream spheres, so that it is obvious that the bypass flow becomes a significant driving flow for the twin vortices. Additionally, the strong flow toward the wall, which is in the middle of the twin-vortices, corresponds to the impinging flow confirmed in cross section S_1. In other words, the two bypass flows heading toward the large gap area, after passing through the upstream sphere, generate the twin-vortices that correspond to the upstream separation vortices. Since there is one more pair of twin vortices downstream due to the colliding effect of the two bypass flows with the downstream sphere, these two different twin vortices form three-dimensional structures in the gap area. On the other hand, the velocity vectors of the spouting flow from the pipe center observed in the cross section S_1 have not been confirmed. It is predicted that the bypass flow toward the gap area pushes back the spouting flow toward the pipe center again in this section S_5. As a result, the low-velocity area in the upstream of the impinging flow is formed in the circled areas. Although the spouting flow tries to spread out also in a circumferential direction, it heads toward the opposite direction against the rotating direction of the twin-vortices.

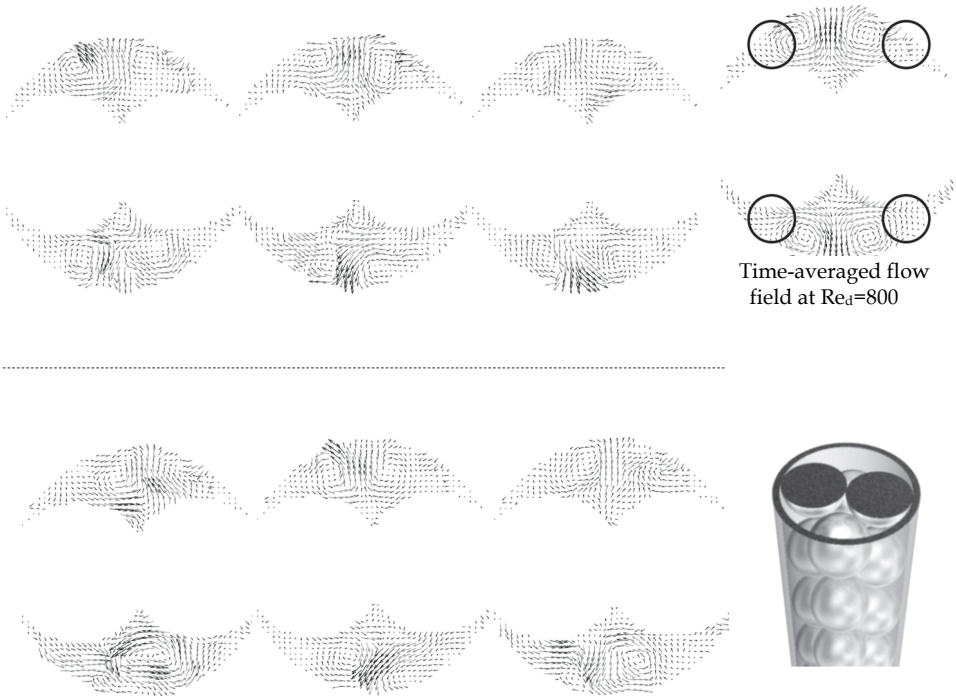


Fig. 17. Time-series of secondary flows (Upper: Re_d=800, Lower: Re_d=4900, Time interval=0.07sec.)

The multiple vortices with the above mentioned complex structure could have strong effects on the fluid mixing and heat transport from the heating wall, which suggests that control of the bypass flow plays an important role in heat transfer augmentation of the SPP.

5. Other application of the matched refractive-index PIV method

The matched refractive-index PIV method easily enables to visualize various kinds of complex flow structures. As the second application of this method, we focus on a flow separation formed in a curved pipe with a low curvature ratio. Figure 18 shows the longitudinal flow fields in the curved pipe and its downstream circular pipe region (Yuki et al., 2009). The Re number is approximately 50,000 (inlet mean velocity is 0.78 m/s). This flow condition corresponds to a subcritical regime defined by Idelchik (1986). The test section was made from an acrylic rectangular block where the curved flow-channel is manufactured by 3D-NC cutting. The pipe diameter is 56mm. The entrance section to the curved pipe is a circular straight pipe with a length of 50 times as long as the pipe inner diameter so that the fluid flow going into the test section becomes an almost fully developed turbulent flow. The curvature ratio of the curved pipe is 1.0 (=R/D : where R is the curvature radius). Here, we introduced an infrared ray Diode laser, with an output of 10mJ and a wavelength of 808nm. This indicates that the refractive-index of the NaI solution needs to be re-adjusted with the method shown in Figure 1. Nylon particles with the average diameter of 20μm and a specific gravity of 1.03 are used as the tracer particle. Figures 18(a), (b), (c), and (d) show the instantaneous flow fields, time-averaged flow field, velocity contour lines map, and the RMS distribution of velocity fluctuation in the longitudinal section of the curved pipe and its downstream, respectively. As the accuracy of PIV measurement near the outlet area of the curved pipe drastically degrades due to existence of an O-ring, a rectangular white box is put on it. The high velocity flow produced by a centrifugal force oscillates in an unsteady manner, and this oscillation is conceivably self-induced one due to flow instability by loosing a balance between a centrifugal force and a pressure gradient especially near the inlet area of the curved pipe. The highest velocity is confirmed near the inner wall of the inlet, due to the favorable pressure gradient toward the streamwise direction. The maximum velocity reaches 1.12m/s which is 1.44 times as the inlet mean velocity. Moreover, a low-velocity area on the inner wall side that is considered to be a part of the flow separation develops from around 40 degree of angle from the elbow inlet (see the time-averaged flow field). The instantaneous flow fields also show a counter flow structure that can be regarded as a part of the separation vortex. The low-velocity area develops toward the downstream of the curved pipe, and along with that, it is confirmed that the flow field in the central area is accelerated again in the outlet region. The separation point exists at a more upstream position, compared with the previous research results (Shiraishi et al., 2006). This is because the turbulent flow with a fully developed boundary layer flows into the curved pipe, unlike the previous research whose inlet flow has an almost uniform velocity profile. Focusing on the flow structure of the low velocity region, Figure 18 clearly show that the low-velocity area with a complicated unsteady flow exists on the inner wall side of the curved pipe and develops toward the downstream circular pipe region. Furthermore, comparatively small transverse vortices are repeatedly generated and disappeared in the upstream shear layer of the low-velocity area (hereafter, a separation region), which increases the RMS of velocity fluctuation as shown in Figure 18(d). In

addition, much larger flow structures exists in the downstream of the low velocity region. In particular, this larger flow structure significantly and unsteadily changes near the wall of the circular pipe (the RMS value is around 0.4), which is likely to reflect a part of the shedding behavior of separation vortices that have been pointed out by the previous research. In fact, the Strouhal number, fD/Um ~0.59 (f=8.2Hz), is almost the same as that estimated by Iwamoto et al. (2008). The time-average flow field shows that the velocity becomes higher again in this strong fluctuating region (hereafter, a velocity recovery region). The starting point of the velocity recovery is approximately 0.33D. Namely, since the high-velocity fluids transferred from the central area of the curved pipe by the secondary flow through the side walls collide each other around this area, the unsteadiness of the flow field becomes much more significant.

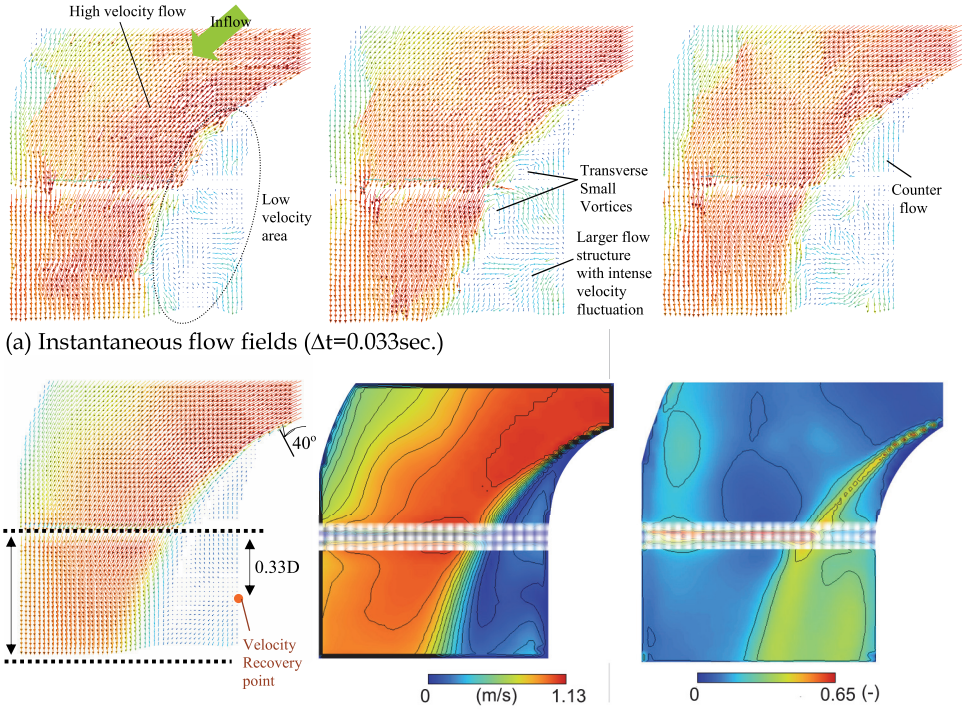
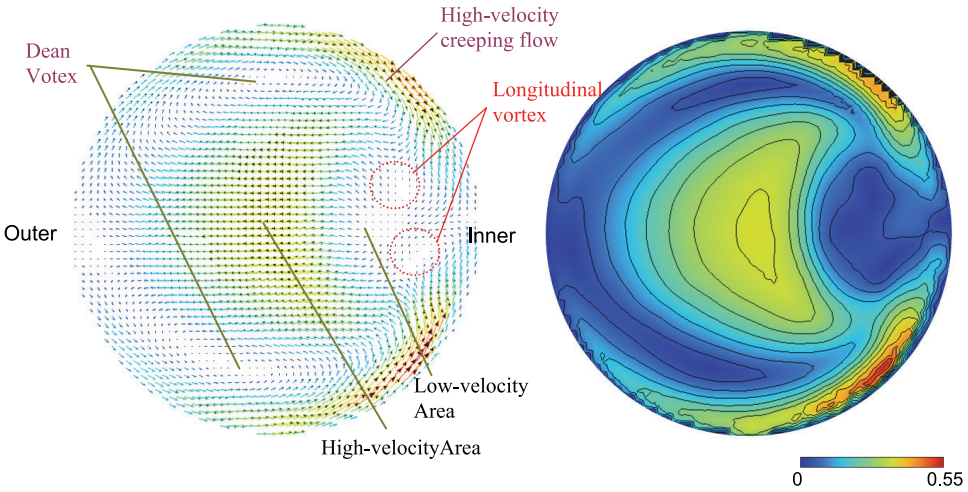
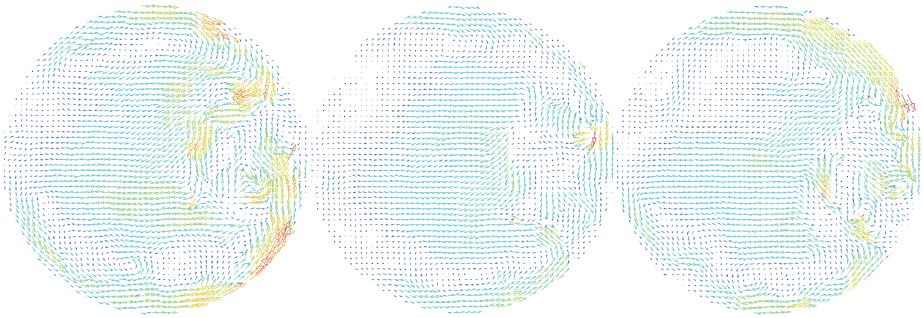


Fig. 18. Flow structure in a curved pipe and its downstream close-up of low velocity region

To clarify the three-dimensional structure of the separation region, the secondary flow is visualized at the 0.18D downstream of the outlet of the curved pipe (see Fig. 19). A low-velocity area exists between the high-velocity area formed in the pipe center and the inner wall. Furthermore, two pairs of characteristic twin vortices can be confirmed. The larger twin vortices are Dean vortices generated by the centrifugal force, whereas the smaller twin vortices formed near the inner wall represent the lateral structure of the separation region. In other words, the flow structure at this location has not only the transverse



(a) Time-averaged flow field



(b) Instantaneous flow fields (Δt=0.33 seconds)

Fig. 19. Secondary flow structure shortly after 1st elbow

vortices but also two longitudinal vortices. Moreover, a high-velocity flow that exceeds 70% of the inlet mean flow velocity exists around the side walls (hereafter, a high-velocity creeping flow). This high-velocity creeping flow alternately flows into the separation region, so that the flow field in the velocity recovery region could become more unsteadily, accompanied with complicated large vortices. Although further investigation should be conducted to clarify in what part of the elbow this high-velocity creeping flow is formed, there could be the possibility that this flow is formed to complement the fluid defect along with the shedding of the separation vortices. Whichever the case may be, there is no doubt that this high-velocity creeping flow has a significant impact on the flow

structure of the velocity recovery region and the dynamic behavior of the vortex shedding as well.

6. Conclusions

In this chapter, I introduced the matched refractive-index PIV measurement to visualize complex flow structures. The visualization was conducted by matching refractive-index of working fluid to that of an obstacle and/or the piping material. Here, flow structures in a sphere-packed pipe and a curved pipe were focused on as the examples, the interesting flow structure became obvious. Although the NaI solution was utilized here, some other solutions are also available or special materials with the same refractive index of the working fluid could be useful. For instance, FEP material has the same refractive index as that of water. The matched refractive-index visualization technique enables to understand not only a 3-dimentional time-series of complex flow structures with complex geometry by utilizing a stereo type of dynamic PIV technique but also a two-phase flow in the near future.

7. References

K. Yuki, et al. (2007). Heat transfer performance for high Prandtl and high temperature molten salt flow in sphere-packed pipes, *Fusion Science & Technology*, vol. 47, no. 4, pp. 1089-1093.

Kazuhisa Yuki, et al. (2008). Heat transfer enhancement in sphere-packed pipes under high Reynolds number conditions, *Fusion Engineering and Design*, vol. 83, Issues 7-9, pp. 1102-1107.

R. M. Fand & R. Thinakaran (1990). The influence of the wall on flow through pipes packed with spheres, *Journal of Fluids engineering*, vol. 112, pp. 84-88.

R. M. Fand, M. Sundaram, M. Varahasamy (1993). Incompressible fluid flow through pipes packed with spheres at low dimension ratios, *Journal of Fluids engineering*, vol. 115, pp. 169-172.

R. M. Fand, M. Varahasamy, L. S. Greer (1993). Empirical correlation equations for heat transfer by forced convection from cylinders embedded in porous media that account for the wall effect and dispersion, *International journal of heat mass transfer*, vol. 36, no. 18, pp. 4407-4418.

K. Yuki, et al. (2005). Evaluation of flow structure in packed-bed tube by visualization experiment, *Fusion science and technology*, vol. 47, no. 4, pp. 1089-1093.

Kazuhisa Yuki, et al. (2009), Unsteady Hydraulic Characteristics in Large-Diameter Pipings with Elbow for JSFR, (3) Flow Structure in a 3-Dimentionally Connected Dual Elbow Simulating Cold-Leg Piping in JSFR, *Proceedings of the 13th International Topical Meeting on Nuclear Reactor Thermal Hydraulics*, N13P1167.

Idelchik, I. E. (1986). *Handbook of Hydraulic Resistance, 2nd edition.*, Hemisphere, New York, p.271.

T. Shiraishi et al. (2006). Resistance and Fluctuating Pressures of a Large Elbow in High Reynolds Numbers, *Journal of Fluids Engineering*, vol. 128, pp. 1063-1073.

Y. Iwamoto et al. (2008). Study on Flow-Induced-Vibration Evaluation of Large-Diameter Pipings in a Sodium-Cooled Fast Reactor (4) Experiments on the 1/10-Scale Hot Leg Test Facility in Reynolds Number of 50000 and 320000, *Proceedings of Japan-Korea Symposium on Nuclear Thermal Hydraulics and Safety (NTHAS6)*, N6P1010.

2

Plasma Flow Control

Ying-hong Li, Yun Wu, Hui-min Song, Hua Liang and Min Jia
Air Force Engineering University of China,
People's Republic of China

1. Introduction

Plasma flow control, based on the plasma aerodynamic actuation, is a novel active flow control technique to improve aircrafts' aerodynamic characteristics and propulsion efficiency. Plasma flow control has drawn considerable attention and been used in boundary layer acceleration, airfoil separation control, forebody separation control, turbine blade separation control, axial compressor stability extension, heat transfer and high speed jet control. Plasma aerodynamic actuator has many advantageous features including robustness, simplicity, low power consumption and ability for real-time control at high frequency.

In this chapter, the principle of plasma aerodynamic actuation and its application in subsonic and supersonic flow control was summarized. In order to better understand the underlying physical mechanism of plasma flow control and optimize the geometric configuration of the actuator, the characteristics of the plasma aerodynamic actuation, including gas temperature, electron density and temperature, induced body force, velocity and vorticity were investigated. Both wind tunnel experiments and computations were performed to investigate the flow control capability of plasma aerodynamic actuation for airfoil separation control, corner separation control, axial compressor stability extension and shock wave control.

2. Plasma aerodynamic actuation

The mechanism for plasma flow control can be summarized as momentum effect, shock effect, and chemical effect. Momentum effect induces near-surface flow velocity of 1-6 m/s. Shock effect induces local gas pressure or temperature rise near the electrode. Chemical effect adds new particles, such as ions, electrons, and excited particles, into the flow field. Surface dielectric barrier discharge plasma aerodynamic actuation is typically used in subsonic plasma flow control, while arc discharge plasma aerodynamic actuation is mostly used in supersonic and hypersonic plasma flow control. Plasma aerodynamic actuation generated by corona discharge, radiofrequency discharge, microwave discharge are also widely investigated.

2.1 Surface dielectric barrier discharge plasma aerodynamic actuation

A schematic of the dielectric barrier discharge plasma aerodynamic actuator is shown in Fig. 1. Most of the previous experimental and simulation studies mainly focused on the asymmetric dielectric barrier discharge plasma aerodynamic actuation excited by sinusoidal or sawtooth voltage waveforms at amplitude of 2-20 kV and frequency of 1-100 kHz, which can be named as the microsecond discharge plasma aerodynamic actuation. In recent years,

the plasma aerodynamic actuation generated by nanosecond pulsed dielectric barrier discharge has become a hot topic.

Synthetic measurements and analysis of the characteristics of the asymmetric surface dielectric barrier discharge plasma aerodynamic actuation were performed. Fig. 1 shows the schematic diagram of the asymmetric surface dielectric barrier discharge plasma aerodynamic actuator. The dielectric layer used is a RO4350B (Rogers Corporation) plate with a relative permittivity constant of 3.48. The electrodes are made of copper, covered with a thin layer of lead-tin film. The plasma aerodynamic actuator is driven by a high frequency high voltage power supply (CTP-2000K, Suman Electronics). The output voltage range and frequency range of the power supply are 0~40 kV and 6~40 kHz respectively.

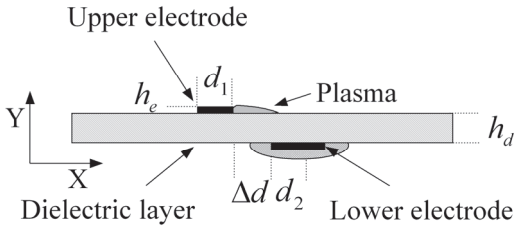


Fig. 1. A schematic of the asymmetric surface dielectric barrier discharge plasma aerodynamic actuator

The experimental arrangement is shown in Fig. 2. Synthetic measurements are applied to measure both the plasma and the induced flow characteristics of the plasma aerodynamic actuation. The applied voltage and the total discharge current are measured by a high voltage probe (P6015A, Tektronix Inc.) and a current probe (TCP312+TCPA300, Tektronix Inc.). Signals are recorded on an oscilloscope (DPO4104, Tektronix Inc.). Optical emission spectroscopy is measured by a 0.5m monochrometer (TRIAX550, Jobin Yvon Inc.) through an optical fiber collector, 1 cm over the surface of the dielectric layer. The detector of the monochrometer is a set of photon counting system (Model 76915, Oriel Inc.). The slit width and the calibrated resolution in this case are 20μm and 0.05nm. The emission intensity is averaged temporally and spatially. Body force induced by the plasma aerodynamic actuation is measured by a electronic balance(ABS 204-S, Mettler Toledo Inc.). Velocity and vorticity induced by the plasma aerodynamic actuation is measured by Particle Image Velocimetry (Lavision). The air is seeded by vaporization of mineral oil with a mean size of about 0.3 μm.

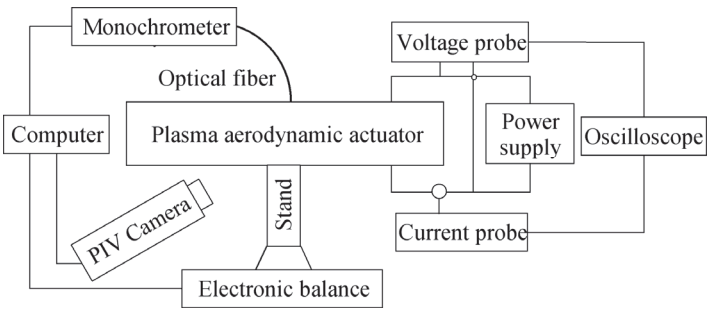


Fig. 2. Schematic diagram of the experimental arrangement

2.1.1 Characteristics of the plasma

Optical emission spectroscopy(OES), as a simple and non-intrusive diagnostic method, is playing a very important role in providing qualitative and quantitative information of plasma properties atmospheric pressure air plasmas. Several plasma parameters, including gas temperature, vibrational temperature, electron temperature and density, have been successfully characterized by OES method in this study. The major spectra come from the second positive system (SPS) of $N_2(C^3\Pi_u \rightarrow B^3\Pi_g)$, and the first negative system (FNS) of $N_2^+(B^2\Sigma_u^+ \rightarrow X^2\Sigma_g^+)$. The relative concentration of $N_2(C^3\Pi_u)$ is much greater than that of $N_2^+(B^2\Sigma_u^+)$. The emitting species of $N_2(C^3\Pi_u)$ and $N_2^+(B^2\Sigma_u^+)$ mainly come from the following excitation processes from the ground state of N_2.

$$e + N_2(X) \rightarrow e + N_2(C) \quad (E_{th}=11 \text{ eV}) \tag{1}$$

$$e + N_2(X) \rightarrow e + e + N_2^+(B) \quad (E_{th}=19 \text{ eV}) \tag{2}$$

Other main dynamics/kinetics processes in this surface discharge include[24-28]:

$$e + N_2(X,\nu) \rightarrow e + N_2(X,\omega) \tag{3}$$

$$N_2(X,\nu) + N_2(X,\omega) \rightarrow N_2(X,\nu-1) + N_2(X,\omega+1) \tag{4}$$

$$N_2(X,\nu) + N_2 \rightarrow N_2(X,\nu-1) + N_2 \tag{5}$$

$$N_2(C) + N_2, O_2 \rightarrow N_2, O_2 + product \tag{6}$$

$$N_2(C) \rightarrow N_2(B) + h\nu \tag{7}$$

$$N_2^+(B) + N_2, O_2 \rightarrow N_2, O_2 + product \tag{8}$$

$$N_2^+(B) \rightarrow N_2^+(X) + h\nu \tag{9}$$

In atmospheric discharge condition, the gas temperature can be estimated by the rotational temperature of a molecular because the rotational energy levels are closely spaced(10^{-3} eV) and allow for rapid energy transfer between the two energy modes. The population distribution in the rotational energy level of N_2 is also believed to fit Boltzmann distribution due to the dense collisions between molecules. Here, we obtained the rotational temperature of N_2 through fitting the N_2 second positive system band from 378 nm to 381 nm for the 380.5 nm ($\nu'=0$, $\nu''=2$) N_2 line. By assuming a rotational temperature and considering the dipole radiation probability and the response function of the monochrometer, one can calculate the profile of a certain emission band. The actual rotational temperature(T_r) can be determined through comparing the experimental measurement and theoretical calculation.

In this surface discharge, the vibrational excitation of nitrogen seems to be the most important. $N_2(C^3\Pi_u)$, which is not a metastable state, is generated from the ground-state electron impact excitation and the cascading effect is not important for the $N_2(C^3\Pi_u)$ state population. Therefore, the vibrational temperature can be determined according to the ratio of two lines in the $N_2(C)$ second positive system. The spectra lines at 371.1 nm and 380.5 nm are selected to calculate the vibrational temperature(T_v), as follows:

$$\frac{I_{371.1nm}}{I_{380.5nm}} = 1.1384 \cdot \exp(-0.4952 / T_v) \tag{10}$$

When the applied voltage amplitude and driving frequency are 10kV (peak to peak) and 23kHz respectively, the rotational and vibrational temperatures of $N_2(C^3\Pi_u)$ are 500 K and 0.22 eV respectively. The rotational temperature is insensitive to the applied voltage and the driving frequency. The vibrational temperature also shows minor dependence on the applied voltage and the driving frequency, as shown in Fig. 3.

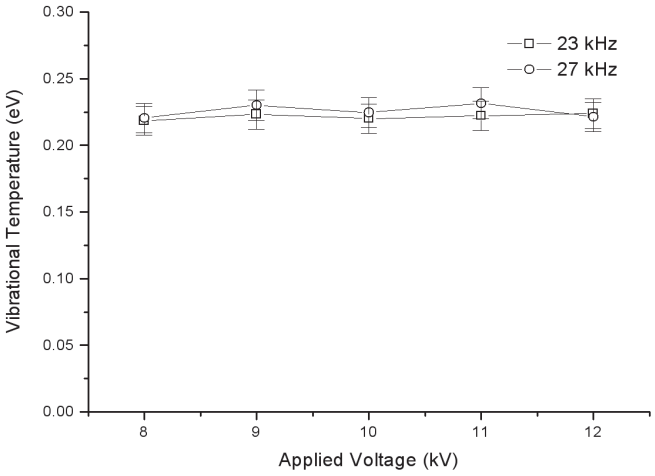


Fig. 3. Vibrational temperature versus the applied voltage

The 0D rate balance equation for the concentration, n_v, of molecules in the vth vibrational level of the ground state is

$$\begin{aligned}\frac{dn_v}{dt} = {} & n_e \sum_{w\neq v} n_w C_w^v - n_e n_v \sum_{w\neq v} C_v^w + n_{v-1}\sum_w n_{w+1} Q_{v,v-1}^{w+1,w} \\ & + n_{v+1}\sum_w n_w Q_{v+1,v}^{w,w+1} - n_v\Big(\sum_w n_{w+1} Q_{v,v+1}^{w+1,w} + \sum_w n_w Q_{v,v-1}^{w,w+1}\Big) \\ & + n_{v+1}\sum_w n_w Q_{v+1}^{v} - n_v \sum_w n_w Q_v^{v-1}\end{aligned} \tag{11}$$

Here Q and C are rate coefficients of collision.
The rate balance equation for each excited state is, for state $N_2(C)$,

$$\frac{dn_C}{dt} = n_e n_{N_2} Q_C - A_C n_C - n_C n_{N_2} Q_{N_2} - n_C n_{O_2} Q_{O_2} \tag{12}$$

for state $N_2^+(B)$,

$$\frac{dn_{B+}}{dt} = n_e n_{N_2} Q_{B+} - A_{B+} n_{B+} - n_{B+} n_{N_2} Q_{N_2} - n_{B+} n_{O_2} Q_{O_2} \tag{13}$$

Here A is the Einstein coefficient.
The intensity ratio of optical emission line from FNS to SPS is strongly dependent on the electron temperature or the energy distribution function for the electrons whose energy is more than 11 eV, because the excited thresholds of the two processes have a difference of about 8 eV. According to equation (12) and (13), electron temperature(T_e) can be calculated using the intensity ratio of 391.4 nm and 380.5 nm:

$$\frac{I_{391.4nm}}{I_{380.5nm}} = K_0 \cdot (T_e)^{C_0} \cdot \exp(-\frac{E_0}{T_e}) \tag{14}$$

As shown in equation (12) and (13), the relative intensity of different vibrational levels of $N_2(X)$ is mainly determined by the electron density. According to Frank-Condon principle, the vibrational energy level distribution of $N_2(C)$ can be revealed using the changes in vibrational energy level distribution of $N_2(X)$. Therefore, the ionization rate(n) can be calculated using the intensity ratio of 371.1 nm and 380.5 nm:

$$\frac{I_{371.1nm}}{I_{380.5nm}} = C_0 + C_1 \log_{10}(n) + C_2 (\log_{10} n)^2 \tag{15}$$

The electron temperature and electron density are 1.63 eV and $1.1 \times 10^{11} \mathrm{cm}^{-3}$ respectively with the applied voltage of 10 kV. The driving frequency is 23 kHz. At one atmospheric pressure, electron temperature is usually 1~2 eV because of the frequent collisions between electrons and molecules. The variations of average electron temperature and density with the applied voltage are shown in Fig. 4. It can be concluded that the average electron density and temperature have minor dependence on the applied voltage and its frequency. The electron temperature of a dielectric barrier discharge is strongly affected by the gas pressure because the collisionless free path of the electrons mainly determines the energy obtained by the electron from the electric field. The frequent collisions between electrons and molecules govern the discharge process at one atmosphere.

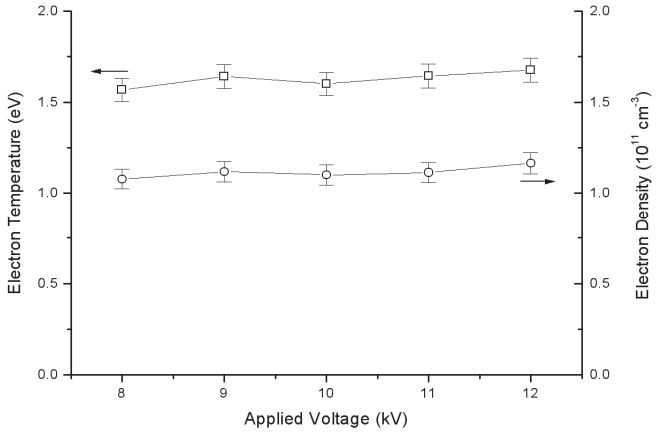


Fig. 4. Electron temperature and density versus the applied voltage

2.1.2 Characteristics of the induced flow

The induced body force in X direction versus the applied voltage is shown in Fig. 5. The driving frequency is 23 kHz. When the applied voltage increases from 8 kV to 12 kV, the body force increases from 11 mN to 65 mN.

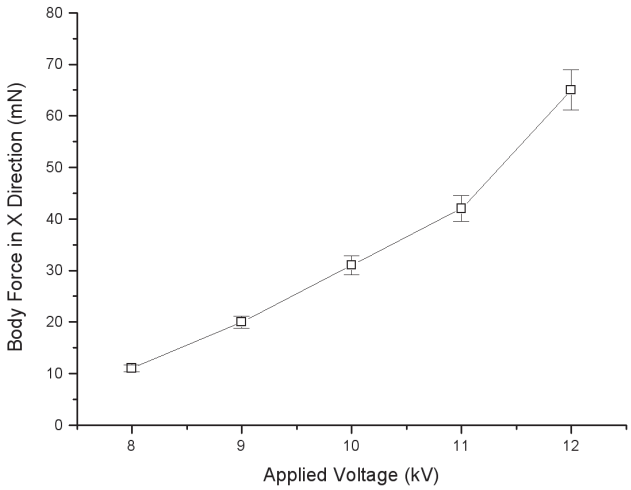
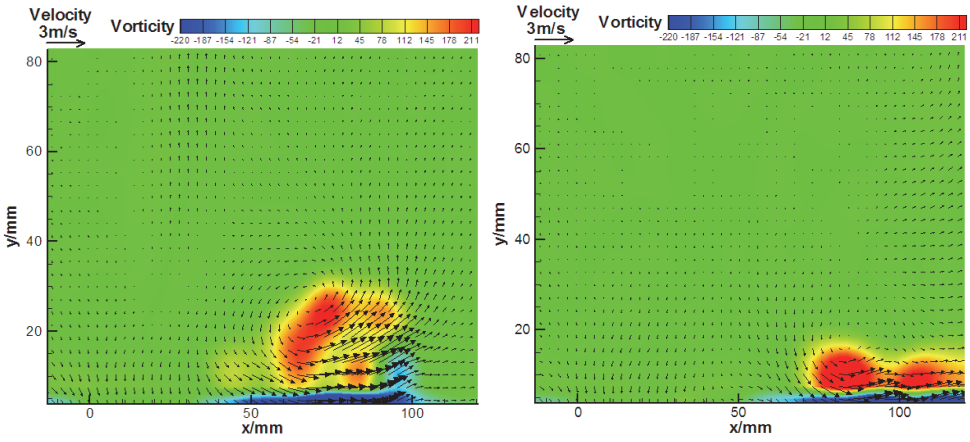


Fig. 5. Body force induced by the plasma aerodynamic actuation

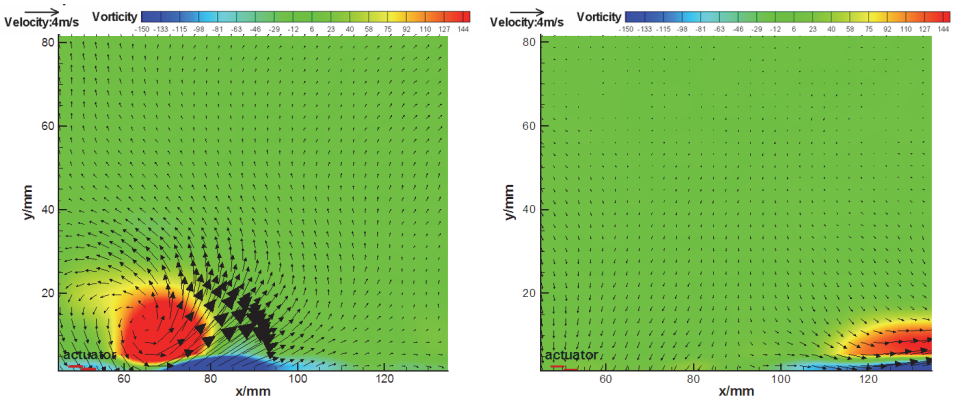
Starting vortex and directed wall jet induced by the steady plasma aerodynamic actuation in static air are shown in Fig. 6. The starting vortex exists for about 1s and then directed wall jet is formed, about 70mm downstream of the upper electrode. The induced flow velocity is 3 m/s when the applied voltage and the driving frequency are 10 kV and 23 kHz respectively.



(a) 0.3s (b) 2.5s

Fig. 6. Velocity and vorticity of starting vortex induced by steady plasma aerodynamic actuation

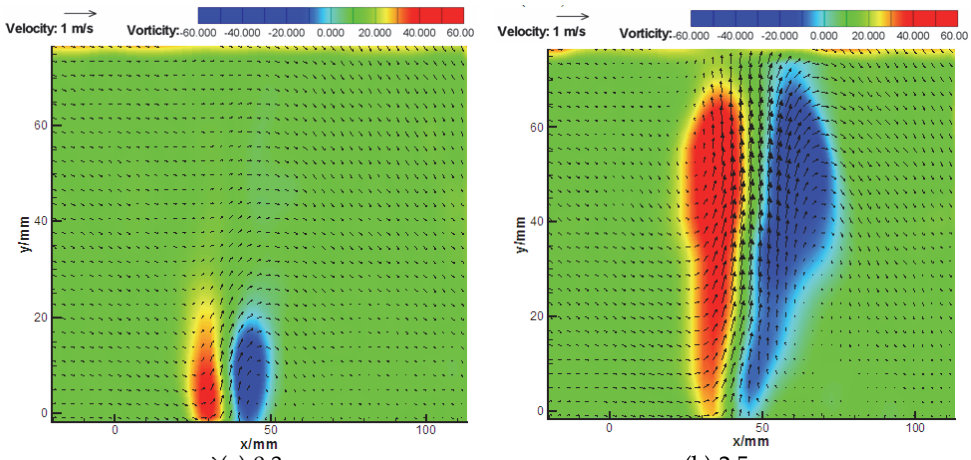
When the unsteady plasma aerodynamic actuation is on, the starting vortex is stronger and exists for a longer time, which is about 4m/s. Then the directed wall jet along with many small vortexes is also formed, about 50mm downstream of the upper electrode, as shown in Fig. 7. The duty cycle and excitation frequency are 70% and 190Hz respectively. The unsteady plasma aerodynamic actuation induces much more vortex into the flow field.



(a) 0.3s (b) 2.5s

Fig. 7. Velocity and vorticity of starting vortex induced by unsteady plasma aerodynamic actuation

When the voltage waveform is a nanosecond pulse, not the sinusoidal pulse, the induced flow direction changes remarkably. The induced flow direction by nanosecond discharge plasma aerodynamic actuation is not parallel, but vertical to the dielectric layer surface(see Fig. 8). The voltage pulse has a full wave at half maximum (FWHM) of 190 ns and a rise time of 450 ns. The peak voltage and frequency are 10 kV and 1 kHz, respectively.



`(a) 0.3s (b) 2.5s

Fig. 8. Velocity and vorticity of starting vortex induced by nanosecond discharge plasma aerodynamic actuation

2.2 Arc discharge plasma aerodynamic actuation

A schematic of the arc discharge plasma aerodynamic actuator is shown in Fig. 9. A dc surface arc discharge is generated between the anode and the cathode. In order to generate a nonuniform electric field that can reduce gas breakdown voltage, a cylindrical structure is designed for the electrodes. The electrodes are oriented in the spanwise direction, which means that the flow direction is perpendicular to the discharge current direction.

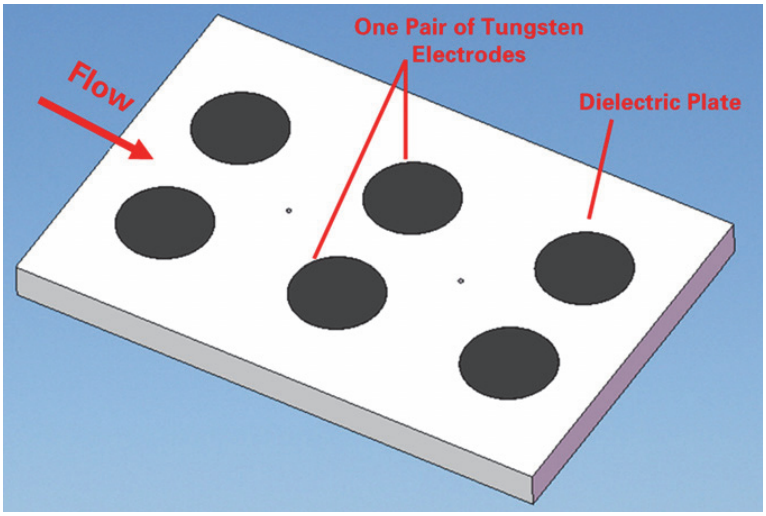


Fig. 9. A schematic of the arc discharge plasma aerodynamic actuator

The power supply consists of a high-voltage pulse circuit and a high-voltage dc circuit. The output voltage of the pulse circuit can reach 90 kV, which is used for electrical breakdown of the gas. The dc circuit is the 3 kV-4 kW power source, which is used to ignite the arc discharge. The plasma aerodynamic actuator consists of graphite electrodes and boron-nitride (BN) ceramic dielectric material. Three pairs of graphite electrodes are designed with the cathode-anode interval of 5mm and the individual electrode is designed as a cylindrical structure which is embedded in the BN ceramic. The arc discharge voltage and current are monitored by a voltage probe (P6015A, Tektronix Inc.) and a current probe with a signal amplifier (TCP312+TCPA300, Tektronix Inc.), respectively. The two signals are measured by a four-channel digital oscilloscope (DPO4104, Tektronix Inc.).

The voltage, current and power measurements are shown in Fig. 10. The gas breakdown voltage between the graphite electrodes is about 2 kV and when the input voltage provided by the power supply exceeds this value, arc discharge happens. At the instant of gas breakdown, voltage decreases from 2 kV to about 300 V and current increases to about 1 A. The discharge power is calculated as 300 W. Then the voltage holds at 300 V, but the current decreases gradually. After about 0.5 s, the current sustains at about 440 mA and the discharge power holds at about 130 W. The arc discharge can be separated into two phases, which correspond to the strong pulsed breakdown process and the steady discharge process, respectively.

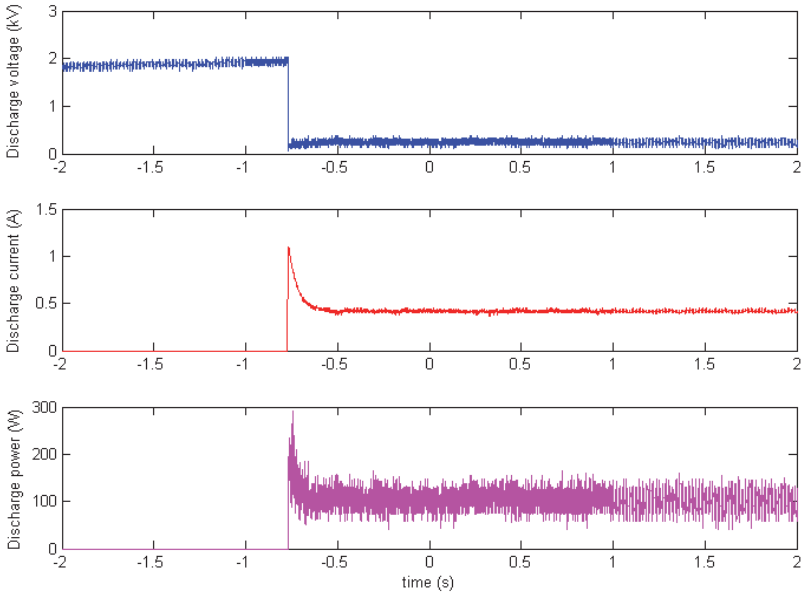


Fig. 10. Electrical characteristics of arc discharge

3. Subsonic plasma flow control

Surface dielectric barrier discharge was proved effective in subsonic plasma flow control. A great number of papers devoted to subsonic plasma flow control have appeared in the past ten years. The use of dielectric barrier discharge for flow control has been demonstrated in many applications. Examples include boundary layer acceleration, transition delay, lift augmentation on wings, separation control for low-pressure turbine blades, jet mixing enhancement, plasma flaps and slats, leading-edge separation control on wing sections, phased plasma arrays for unsteady flow control, and control of the dynamic stall vortex on oscillating airfoils.

3.1 Airfoil flow separation control

More than 70% lift force of aircraft is produced by wings. The lift-to-drag ratio and stall characteristic of the wing is of vital importance to the takeoff distance and climbing speed and the flight quality of the aircrafts. In order to enhance the manoeuvrability and flexibility of the aircrafts, large angle of attack is used frequently. New technology should be employed into the development of aircrafts of the next generation. Active flow control technologies are considered to be the most promising technology in the 21th century.

3.1.1 Flow separation control using microsecond and nanosecond discharge

Flow separation control by microsecond and nanosecond discharge plasma aerodynamic actuation was presented. The control effects influenced by various actuation parameters were investigated.

The airfoil used was a NACA 0015. This shape was chosen because it exhibits well-known and documented steady characteristics as well as leading-edge separation at large angles of attack. The airfoil had a 12 cm chord and a 20 cm span. The airfoil was made of Plexiglas. Twelve pressure ports were used to obtain the pressure distribution along the model surface. Fig. 11 shows location of the pressure ports on the model's surface. Three pairs of plasma aerodynamic actuators were mounted on the suction side of the airfoil. The actuators were positioned 2% and 20% and 45% cord length of the airfoil. The plasma aerodynamic actuators were made from two 0.018mm thick copper electrodes separated by 1mm thick Kapton film layer. The electrodes were 4mm in width and 120mm in length. They were arranged just in the asymmetric arrangement. A 1mm recess was molded into the model to secure the actuator flush to the surface. The pressure distribution along the airfoil surface was obtained by a Scanivalve with 96 channels having a range of ±11 kPa. A pitot static probe was mounted on the traversing mechanism. This was located at different positions downstream of the airfoil, on its spanwise centerline. Discrete points were sampled across the wake to determine the mean-velocity profile. The uncertainty of the measurement was calculated to be less than 1.5%.

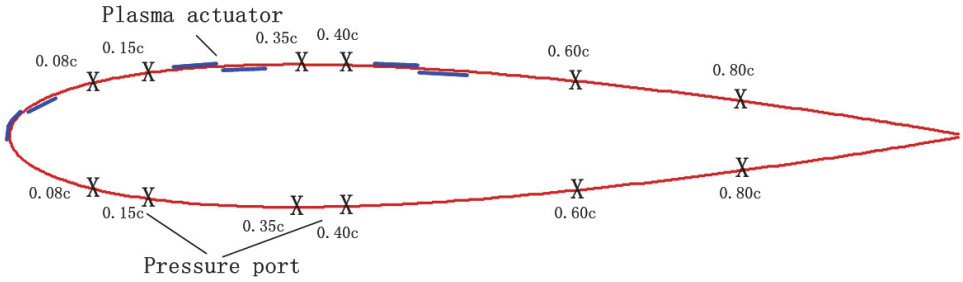


Fig. 11. A schematic of NACA 0015 airfoil with dielectric barrier discharge plasma aerodynamic actuator

The power supply used for microsecond discharge is 0-40 kV and 6-40 kHz, respectively. The output voltage and the frequency range of the power supply used for nanosecond discharge are 5-80 kV and 0.1-2 kHz, respectively. The rise time and full width half maximum (FWHM) are 190ns and 450ns, respectively.

The plasma aerodynamic actuation strength, which is related to the discharge voltage, is an important parameter in plasma flow control experiments. The flow control effects influenced by discharge voltage were investigated. Flow separates at the leading edge of the airfoil without discharge. The pressure distribution has a plateau from leading edge to trailing edge which corresponds to global separation from the leading edge. When the microsecond discharge voltage is 13 kV and 14 kV, the flow separation can not be suppressed. As the microsecond discharge voltage increases to 15 kV, the actuation intensity increases and the flow separation is suppressed. There is a 34.0% lift force increase and a 25.3% drag force decrease when the discharge voltage is 15 kV. When the millisecond discharge voltage increases to 16 kV, there is a 35.1% lift force increase and a 25.5% drag force decrease. The control effects for discharge voltage of 15 kV and 16 kV are approximately the same. Thus, a threshold voltage exists for plasma aerodynamic actuation of different time scale. The flow separation can't be suppressed if the discharge voltage is

less than the threshold voltage. When the flow separation is suppressed, the lift and drag almost unchanged when the discharge voltage increases. The initial actuation strength is of vital importance in plasma flow control. Once the flow separation is suppressed with a initial discharge voltage higher than the threshold voltage, the flow reattachment can be sustained even the discharge voltage was reduced to a value less than the threshold voltage, that is to say, the voltage to sustain the flow reattachment is lower than the voltage to suppress the flow separation in the same conditions. We can make use of the results by managing the discharge voltage properly. A higher discharge voltage can be used to suppress the separation in the beginning, and then we can use a much lower discharge voltage to sustain the flow reattachment later. Not only the power consumption can be reduced obviously, but also the life-span and the reliability of the actuator can be increased greatly.

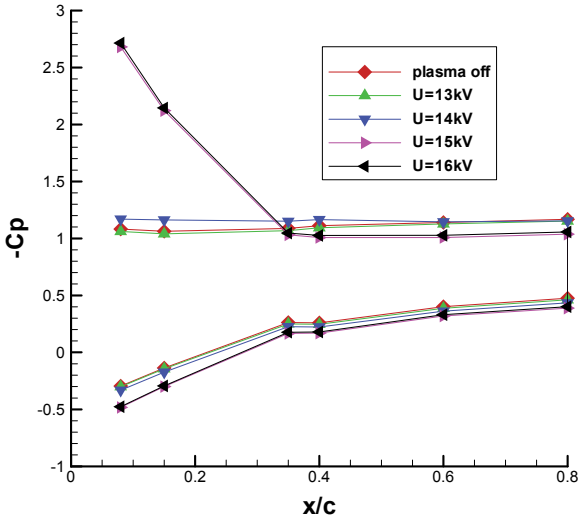


Fig. 12. Pressure distribution for microsecond discharge of different voltage
(α=20°, V_∞=72 m/s, Re=5.8×10^5)

The frequency of nanosecond discharge is believed to be optimum when the Strouhal number $S_{tr} = fc_{sep}/v_\infty$ is near unity. The separation region length and inflow velocity are 100% chord length and 100m/s respectively. The Strouhal number is 1 when the pulse frequency is 830 Hz. Experiments of different pulse frequency were made to determine if such an optimum frequency exists for the unsteady actuation used in controlling the airfoil flow separation.

The experimental results are shown in Fig. 13. It is found that there's an optimum pulse frequency in controlling the airfoil flow separation. The inflow velocity and the angle of attack are 100 m/s and 25° respectively. The duty cycle is fixed at 50%. All three electrodes are switched on. The threshold voltage for different discharge frequency was shown Fig. 14. When the pulse frequency is 830 Hz, the threshold voltage to suppress the flow separation is only 10 kV which is the lowest. When the pulse frequency is 200 Hz and 1500 Hz, the threshold voltage is 13 kV and 12 kV respectively.

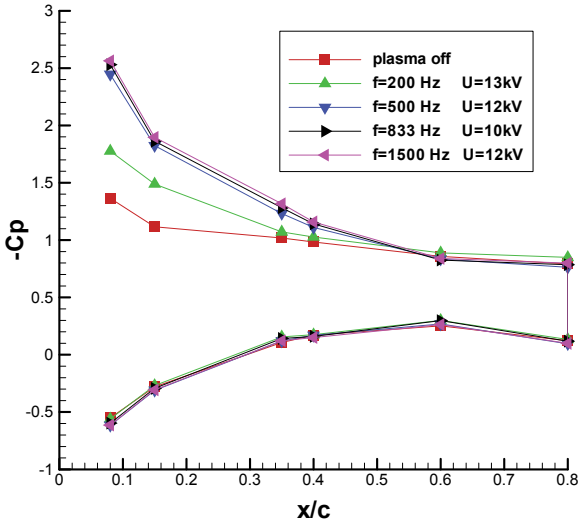


Fig. 13. Pressure distribution for nanosecond discharge of different frequency (α=20°, V_∞=100 m/s, Re=8.1×10^5)

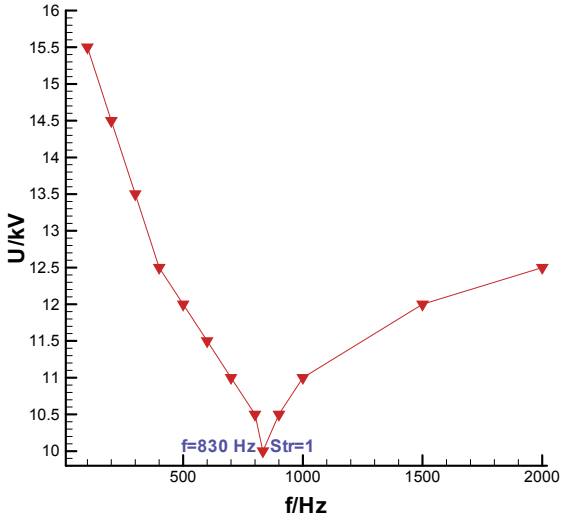


Fig. 14. The threshold voltage at different frequencies for nanosecond discharge (V_∞=100 m/s, a=22°, Re=8.1×10^5)

Plasma aerodynamic actuation of different time scales was used for flow separation control. The flow control ability for microsecond discharge and nanosecond discharge were analyzed. The pressure distribution along airfoil surface obtained in experiments for inflow velocity of 150 m/s (Re=12.2×10^5) are presented in Fig. 15. The angle of attack is 25°, which is approximately 5° past the critical angle of attack at the inflow velocity of 150m/s

(Re=12.2×10^5). The discharge frequency is fixed at 1600 Hz. The discharge voltage for microsecond and nanosecond discharge is 17 kV and 12 kV respectively. When the nanosecond discharge is on, the flow is fully attached at the leading edge. The lift force increases by 22.1% and the drag force decreases by 17.4% with the actuation on. But the microsecond discharge can not suppress the flow separation. The flow still separates at the leading edge with microsecond plasma aerodynamic actuation. It indicates that the flow control ability for nanosecond discharge is stronger than that of the microsecond discharge. The nanosecond discharge is much more effective in leading edge separation control than microsecond discharge.

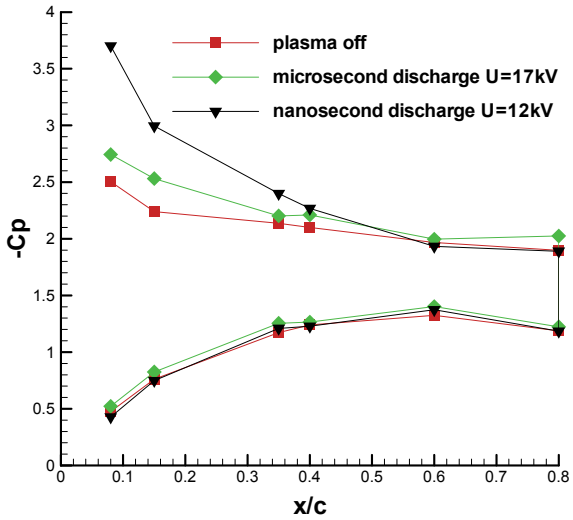


Fig. 15. Experimental results for microsecond and nanosecond discharge (V_∞=150 m/s , a=25°, Re=12.2×10^5)

3.1.2 Flow separation control by spanwise nanosecond discharge

The model used in this study was a NACA 0015 airfoil. Fig. 16 shows the geometry of the airfoil and the actuators. The actuator was made from two 0.018mm thick copper electrodes separated by 1mm thick Kapton film layer. The electrodes were 4mm in width and 60mm in length. They were arranged just in the asymmetric arrangement.

Experimental results for different angle of attacks (a) at the inflow velocity of 72 m/s (Re=5.8×10^5) are shown in Fig. 17. The discharge voltage and frequency of the nanosecond power supply were fixed at 13 kV and 1000 Hz respectively. Experimental results show that spanwise nanosecond discharge aerodynamic actuation can suppress the flow separation effectively. The lift and drag coefficient are nearly unchanged with actuation when the angle of attack is less than 18° or more than 24°. When the angle of attack is less than the critical value, there is nearly no flow separation on the airfoil surface. The effect of spanwise nanosecond discharge aerodynamic actuation can is not obvious. When the angle of attack is more than 24°, the flow separation on the airfoil surface is so aggressive that spanwise nanosecond discharge aerodynamic actuation can not suppress the flow separation on the suction side of the airfoil. So the lift and drag coefficients nearly the same. There is an

obvious lift augmentation and drag reduction after actuation when the angle of attack is between 18° and 24°. The lift coefficient is increased from 0.814 to 1.099 and the drag coefficient is decreased from 0.460 to 0.328 after actuation at the angle of attack 24°. The critical stall angle of attack for NACA 0015 airfoil increased from 18° to 24°. When the angle of attack is 24°, there is a lift force augmentation of 30.2% and a drag force reduction of 22.1% after actuation.

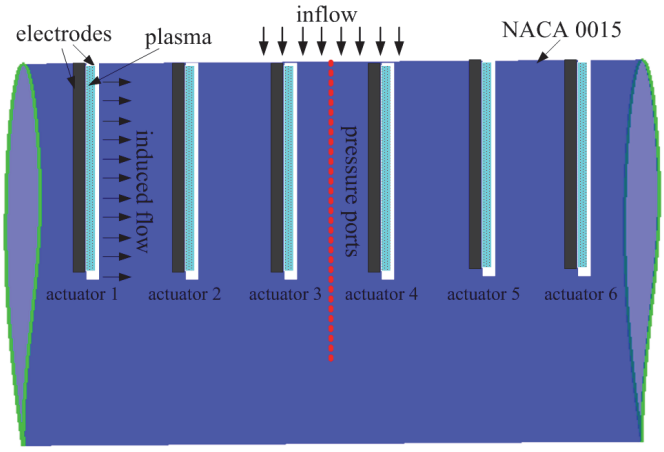
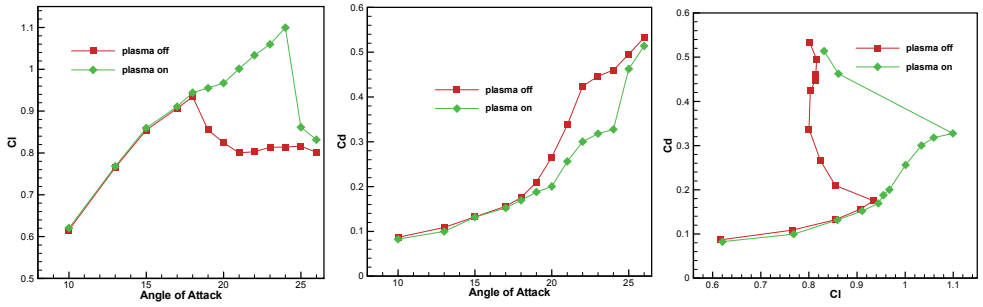


Fig. 16. Schematic drawing of the actuators on the airfoil



(a) Results of lift coefficient (b) Results of drag coefficient (c) Results of lift-to-drag ratio

Fig. 17. Experimental results at different angles of attack (V_∞=72 m/s, Re=5.8×10^5)

The discharge frequency for microsecond discharge is in the orders of kilo hertz. Spanwise plasma aerodynamic actuation of different time scales was used for flow separation control. The flow control ability for microsecond discharge and nanosecond discharge were analyzed. The pressure distribution along airfoil surface obtained in experiments for inflow velocity of 66 m/s (Re=5.3×10^5) and 100 m/s (Re=8.1×10^5) are presented in Fig. 18 and Fig. 19. At the angle of attack 22° and inflow velocity of 66 m/s (Fig. 18), there is initial separated flow on the suction surface of the airfoil without discharge. The discharge voltage for microsecond and nanosecond discharge is 7 kV and

12 kV respectively. The discharge frequency is 1000 Hz. The flow separation on the suction surface can be suppressed by both microsecond and nanosecond discharge actuation. The control effects are nearly the same for microsecond and nanosecond discharge. The spanwise plasma aerodynamic actuations result in a lift augmentation of 23.6% and a drag reduction of 25.6%.

In Fig. 19, the angle of attack is 24°, which is approximately 4° past the critical angle of attack at the inflow velocity of 100m/s (Re=5.8×10^5). The discharge frequency is fixed at 1000 Hz. The discharge voltage for microsecond and nanosecond discharge is 8.5 kV and 12 kV respectively. When the nanosecond discharge is on, the flow is fully attached at the leading edge. The lift force increases by 25.3% and the drag force decreases by 20.1% with the actuation on. But the microsecond discharge can not suppress the flow separation. The flow still separates at the leading edge with microsecond plasma aerodynamic actuation. It indicates that the flow control ability for nanosecond discharge is stronger than that of the microsecond discharge. The nanosecond discharge actuation is much more effective in leading edge separation control than microsecond discharge actuation.

The dielectric layer will be destroyed when the discharge voltage is strong enough. Kapton is used as the dielectric in our experiments. The threshold voltage to destroy the Kapton layer is 8.5kV for microsecond discharge in our experiments. The actuators will be destroyed when the discharge voltage is more than 8.5kV for microsecond discharge. The threshold voltage to destroy the Kapton layer is 17 kV for nanosecond discharge in our experiments.The instantaneous actuation intensity for nanosecond discharge is much stronger than microsecond discharge. So nanosecond discharge is more effective in flow control than microsecond discharge.

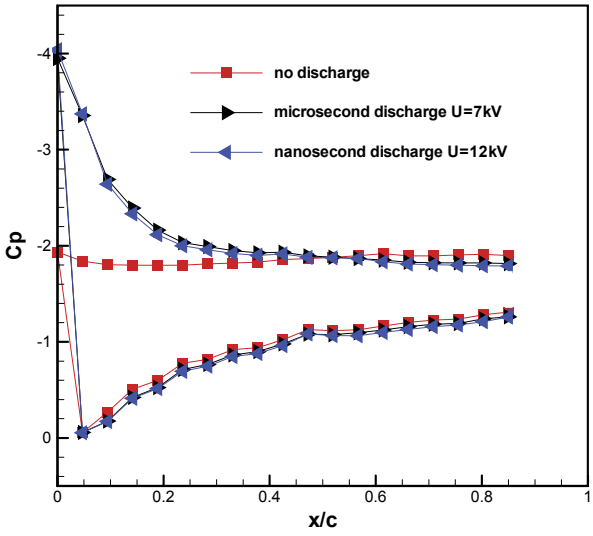


Fig. 18. Experimental results for microsecond and nanosecond discharge (V_∞=66 m/s and α=22° Re=5.3×10^5)

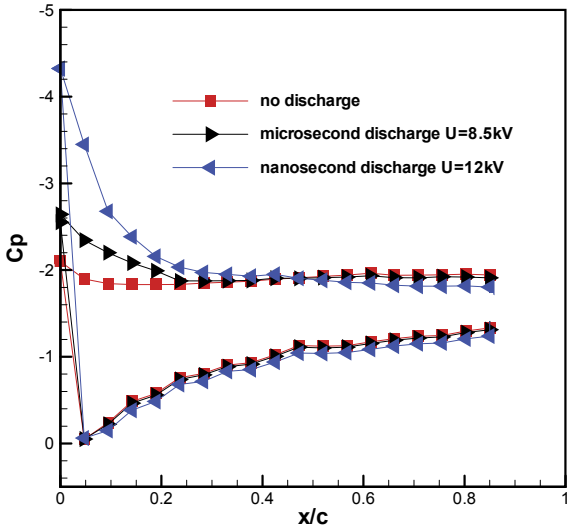


Fig. 19. Experimental results for microsecond and nanosecond discharge (V_∞=100 m/s, a=24°, Re=8.1×10^5)

3.1.3 The mechanism of plasma shock flow control

Based on our works, the principle of "plasma-shock-based flow control" was proposed. Energy should be released in extremely short time to intensify the instantaneous actuation strength, such as nanosecond discharge. Nanosecond discharge yields strong turbulence even shock waves which are act on the boundary layer. Shock wave produces stronger turbulent mixing of the flow, which can enhance momentum and energy exchange between the boundary layer and inflow greatly. High momentum fluid was brought into the boundary layer intermittently, enabling the flow to withstand the adverse pressure gradient without flow separation .The spirits of "plasma-shock-based flow control" lay in three aspects. Firstly, "Shock Actuation", nanosecond discharge should be used to increase the instantaneous discharge power. Nanosecond discharge induces strong local pressure or temperature rise in the boundary. Pressure or temperature rise result in strong pulse disturbance or shock waves in the boundary. Secondly, "Vortex control", shock wave disturbance induces vortex in the process of propagation. Vortex enhances energy and momentum mixing between boundary layer and inflow. The velocity of the boundary layer increase and the flow separation is suppressed. Thirdly, "Frequency Coupling", adjust the discharge frequency to the optimal response frequency in flow control. The optimal response frequency is the one which makes the Strouhal number equal to 1. The plasma aerodynamic actuation work best at the optimal response frequency. Nanosecond discharge can increase the capability of plasma flow control effectively while its energy consumption can be reduced greatly.

For microsecond plasma aerodynamic actuation, the momentum effect may be the dominant mechanism. Microsecond plasma aerodynamic actuation induces near-surface boundary layer acceleration. Energy and momentum is added into the boundary layer, which enhances the ability to resist flow separation caused by adverse pressure gradient for boundary. But the maximum induced velocity for microsecond discharge is less than 10m/s.

The actuators will be destroyed if the discharge voltage is too high. The momentum added into the boundary layer by microsecond discharge is quite limited. The microsecond plasma aerodynamic actuation can only work effectively when the inflow velocity is several tens of meters per second.

The main mechanism for nanosecond discharge plasma flow control may be not momentum effect, since the induced velocity is less than 1m/s. The velocity and vorticity measurements by the Particle Image Velocimetry show that, the flow direction is vertical, not parallel to the dielectric layer surface. The induce flow is likely to be formed by temperature and pressure gradient caused by nanosecond discharge other than energy exchange between charged and neutral particles. Thus, the main flow control mechanism for nanosecond plasma aerodynamic actuation is local fast heating due to high reduced electric field, which then induces shock wave and vortex near the electrode.

Experimental results indicate that nanosecond discharge is more effective in flow control than microsecond discharge. The latest study showed that nanosecond discharges have demonstrated an extremely high efficiency of operation for aerodynamic plasma actuators over a very wide velocity range (Ma= 0.03-0.75). So shock effect is more important than momentum effect in plasma flow control.

3.2 Corner separation control in a compressor cascade

Control of the corner separation is one of the important ways of improving axial compressor stability and efficiency. Our approach to control the corner separation is based on the use of plasma aerodynamic actuation. Experiments were carried out on a low speed compressor cascade facility. Main cascade parameters are shown in Fig. 20. Only the middle blade was laid with the plasma aerodynamic actuator.

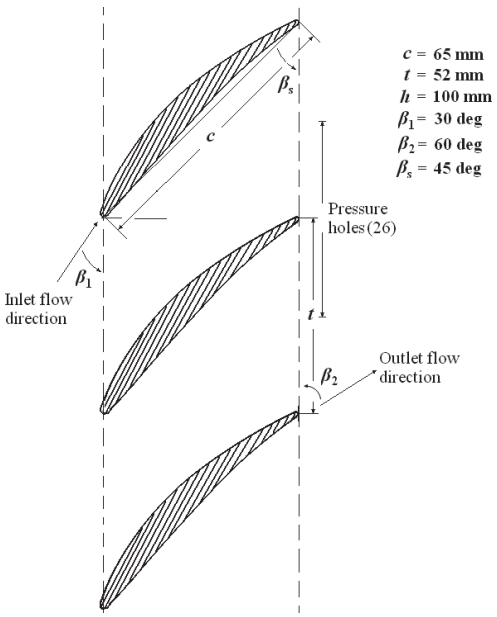


Fig. 20. Compressor cascade parameters

Total pressure distributions at 10mm, which is 15% of the chord length, downstream of the blade trailing edge along the pitch direction at 50%, 60% and 70% blade spans were measured with and without the plasma aerodynamic actuation. A three-hole probe calibrated for pitch and yaw was used to measure the total pressure at the cascade exit. Two parameters, total pressure recovery coefficient σ and the relative reduction of the total pressure loss coefficient $\delta(\omega)$, were used to quantify the performance improvement due to the plasma aerodynamic actuation.

The plasma aerodynamic actuator used in the present experiments consists of four electrode pairs, located at 5%, 25%, 50% and 75% of the chord length, respectively. The electrode pair at 5% of chord length is named as the 1st electrode pair. A sketch of a blade with the actuator on the surface is shown in Fig. 21. The electrode thickness is not to scale in the figure.

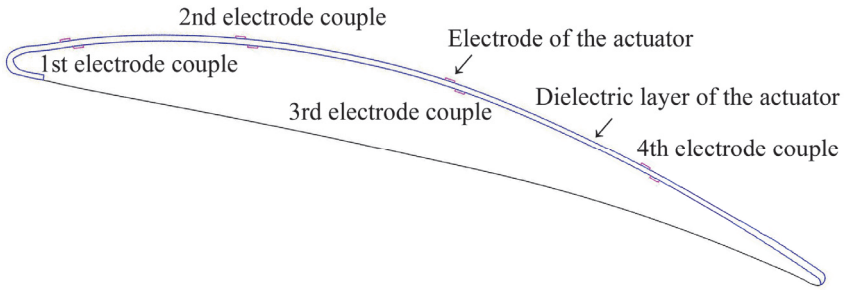


Fig. 21. A sketch of a blade with plasma aerodynamic actuator

The plasma aerodynamic actuator is driven by a high frequency high voltage power supply (CTP-2000M+, Suman Electronics). The output waveform is sine wave. The output ranges of the peak-to-peak voltage and the driving frequency of the power supply are $V_{p\text{-}p}$ = 0~40 kV and F = 6~40 kHz, respectively. The driving frequency is fixed at 23 kHz in the experiments. The plasma aerodynamic actuator works at steady or unsteady mode in the experiments. In the steady mode, the actuator is operated at the ac frequency. In the unsteady mode of operation, the ac voltage is cycled off and on. Fig. 22 shows a typical signal sent to the plasma aerodynamic actuator during the unsteady actuation. Two important parameters of the unsteady plasma aerodynamic actuation are the excitation frequency f, and the duty cycle a, respectively.

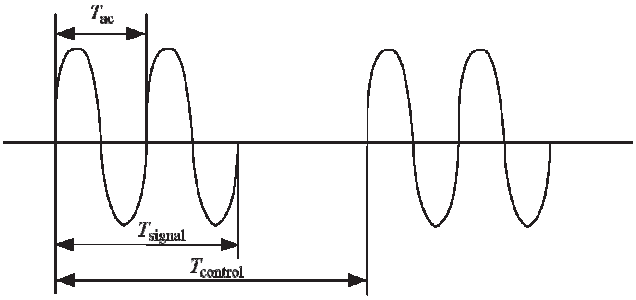


Fig. 22. The signal sent to the plasma aerodynamic actuator during unsteady excitation

3.2.1 Steady plasma flow control experiment results

The mechanism of steady plasma aerodynamic actuation to control the corner separation may be that the actuation induces a time-averaged body force on the flow due to that the flow can't respond to such high frequency (23 kHz in the experiments) disturbances. A wall jet, which is oriented in the mean flow direction, is produced to add momentum to the near-wall boundary layer near the flow separation location. The energized flow is able to withstand the adverse pressure gradient without separation. The directed wall jet governs the flow control effect of steady plasma aerodynamic actuation. When the electrode length is enlarged, the consumed power increases nonlinearly.

The location of the plasma aerodynamic actuation is a key parameter in plasma flow control experiments. Total pressure recovery coefficients with steady actuation at different locations are shown in Fig. 23.

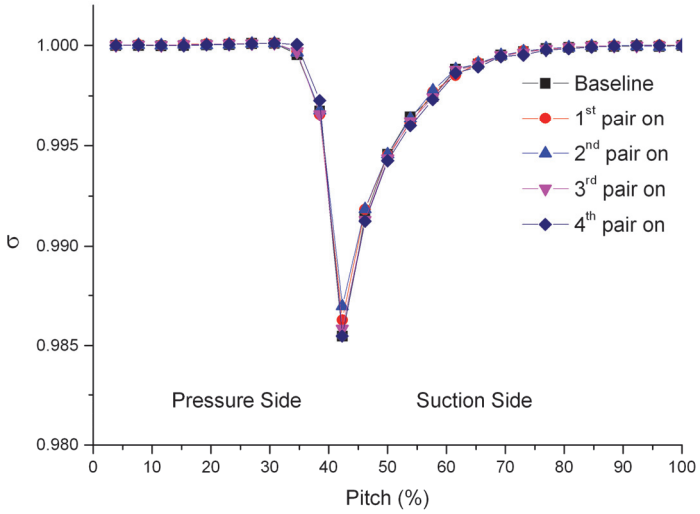


Fig. 23. Total pressure recovery coefficients with steady actuation at different locations (v_∞ = 50 m/s, i = 0 deg, $V_{p\text{-}p}$ = 10 kV, F = 23 kHz, 70% Span)

The applied peak-to-peak voltage and driving frequency are $V_{p\text{-}p}$ = 10 kV and F = 23 kHz, respectively. $\delta(\omega)_{max}$ is 5.5%, 10.3%, 2.4% and 0.07% when the 1st, 2nd, 3rd and 4th electrode pair is switched on, respectively. The 2nd electrode pair at 25% chord length is most effective and the control effect is as the same as that obtained by all four electrode pairs. The power dissipated by the 2nd electrode pair is just 18.4W, about half of the power dissipated by all four electrode pairs. Therefore, the actuation location is vital to the control effect in corner separation control. In corner separation control by tailored boundary layer suction, the optimum slot should be long enough to be sure to remove the limiting streamline and the suction upstream of the corner separation location at the suction surface is most important for the control effect. Therefore, it can be inferred that the location of the 2nd electrode pair is just upstream of the corner separation.

The plasma aerodynamic actuation strength is another important parameter in plasma flow control experiments. The body force increases with the voltage amplitude in proportion to

the volume of plasma (ionized air) and the strength of the electric field gradient. As the applied peak to peak voltage increases from 8 kV to 12 kV, $\delta(\omega)_{max}$ increases from to 2.7% to 11.1%, as shown in Fig. 24. The 2nd electrode pair at 25% chord length is switched on and the driving frequency is 23 kHz. The power dissipation increases from 8.4 W to 23.5 W when the applied peak to peak voltage increases from 8 kV to 12 kV. When the applied voltage is less than 9 kV, the control effect is very tiny. When the applied voltage is higher than 10 kV, the control effect saturates and further increases in the voltage amplitude shows no evident benefit. Furthermore, higher voltage may lead to earlier destruction of the dielectric material, which is not desirable in the experiments.

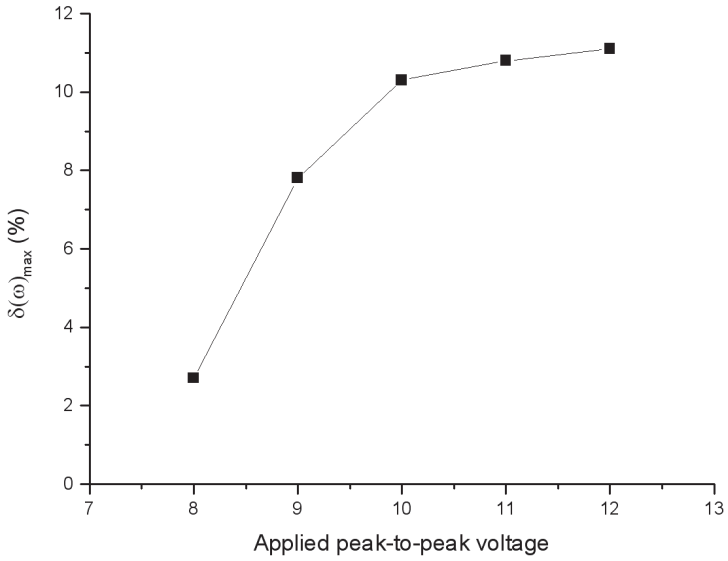


Fig. 24. Control effect with steady actuation of different applied voltages
(v_∞ = 50 m/s, i = 0 deg, F = 23 kHz, 70% Span)

3.2.2 Unsteady plasma flow control experiment results

Optimization of the excitation mode based on coupling between the plasma aerodynamic actuation and the separated flow is one of the important ways of improving plasma flow control effect. It has been shown in the literature that the introduction of unsteady disturbances near the separation location can cause the generation of large coherent vortical structures that could prevent or delay the onset of flow separation. These structures are thought to intermittently bring high momentum fluid to the surface, enabling the flow to withstand the adverse pressure gradient without separation.

A sensitive study is performed to determine if such an optimum frequency exists for the unsteady actuation used in controlling the corner separation. Fig. 25 documents the relative reductions of maximum total pressure loss coefficient at 70% blade span for a range of excitation frequencies from 100 Hz to 1000 Hz when the duty cycle is fixed at 60%. All four electrode pairs are switched on. The applied peak-to-peak voltage and driving frequency are $V_{p\text{-}p}$ = 10 kV and F = 23 kHz, respectively.

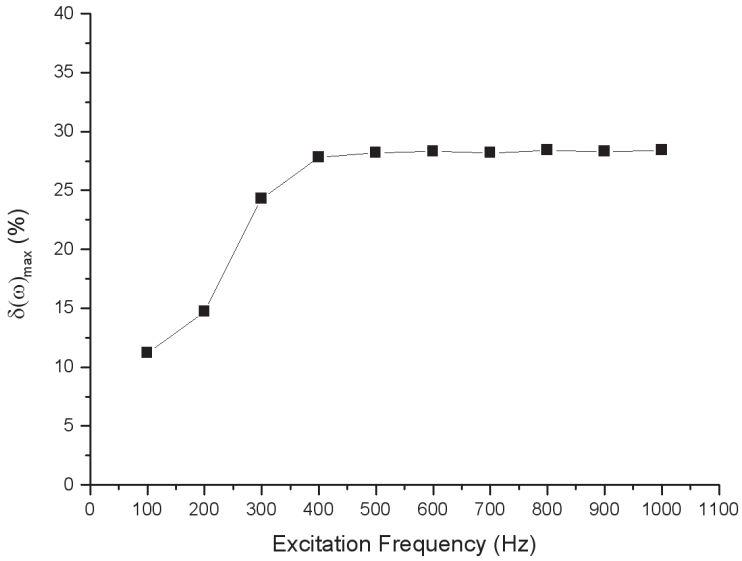


Fig. 25. Maximum relative reductions of total pressure loss coefficient with unsteady actuation of different duty cycles
(v_∞ = 50 m/s, i = 0 deg, $V_{p\text{-}p}$ = 10 kV, F = 23 kHz, 70% Span)

When the excitation is 100 Hz, $\delta(\omega)_{max}$ is just 11.2%, which is almost as same as the steady control effect that is 10.7%. Along with the excitation frequency increasing, the control effect increases. When the excitation frequency is 400 Hz, $\delta(\omega)_{max}$ increases to 28%. Thus, compared with the steady actuation, the unsteady actuation is much more effective and requires less power. When the excitation frequency is higher than 400 Hz, the control effect saturates and further increases in the excitation frequency shows no evident benefit. The difference between steady and unsteady plasma aerodynamic actuation may be that, the unsteady pulsed operation allows the continuously generation of vortical structures, while the steady operation can't. Vortical structures in the flow field promote momentum transfer in the boundary layer in order to withstand separation. Under different duty cycles and excitation frequencies, the coupling between actuation and flow field leads to different flow control effects.

Each electrode pair is switched on to study the effect of the actuation location. The control effect of all four electrode pairs is almost as same as that obtained by the 2nd electrode pair. The saturation frequency is also 400 Hz. For the 2nd electrode pair, the characteristic length is the remaining chord length downstream of the actuator, which is 75% chord length. Thus, the Strouhal number $Sr = f \times C/v_\infty$ is 0.4 when the frequency and freestream velocity are f = 400 Hz and v_∞ = 50 m/s, respectively. When the Strouhal number exceeds 0.4, the control effect saturates in the unsteady plasma flow control experiments. In the separation control above a NACA 0015 airfoil with unsteady plasma aerodynamic actuation (Benard et al. 2009), the most effective actuation was performed with a Strouhal number of Sr ranging from 0.2 to 1.The optimum excitation frequency depends much on the flow separation state. Under different flow conditions, the optimum excitation frequency is also different.

Fig. 26 documents the maximum relative reductions of total pressure loss coefficient for a range of unsteady duty cycles from 5% to 100% when the excitation frequency is fixed at 400 Hz. All four electrode pairs are switched on. The applied peak-to-peak voltage and driving frequency are $V_{p\text{-}p}$ = 10 kV and F = 23 kHz, respectively.

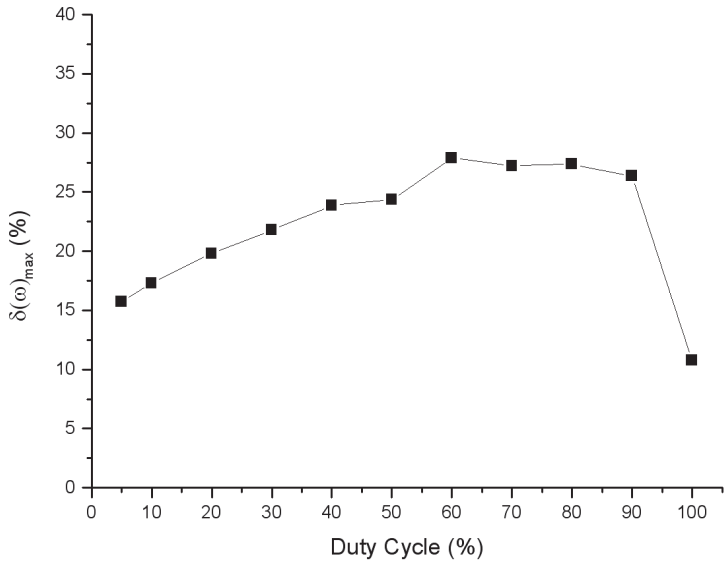


Fig. 26. Maximum relative reductions of total pressure loss coefficient with unsteady actuation of different duty cycles
(ν_∞ = 50 m/s, i = 0 deg, $V_{p\text{-}p}$ = 10 kV, F = 23 kHz, 70% Span)

It is found that there's also a duty cycle threshold in controlling the corner separation. When the duty cycle is less than 60%, the control effect increases along with the duty cycle increasing. $\delta(\omega)_{max}$ is 28% at the duty cycle of 60%. Even when the duty cycle is 5%, $\delta(\omega)_{max}$ is 15.7%, much more effective than the steady actuation. When the duty cycle is higher than 60%, the control effect saturates along with the duty cycle increasing. Thus it can be inferred that when the duty cycle is less than 60%, the injected energy is not sufficient to control the corner separation. In the separation control above a NACA 0015 airfoil with unsteady plasma aerodynamic actuation, the most effective duty cycle values range from 10% to 60%. In the separation control of low-pressure turbine blades with unsteady plasma aerodynamic actuation, the lowest plasma duty cycle (10%) was as effective as the highest plasma duty cycle (50%) at the same excitation frequency. Thus, the optimum duty cycle also depends much on the flow separation state.

3.3 Low speed axial compressor stability extension

This series of tests were carried out using a low speed axial compressor test rig at Institute of Engineering Thermophysics, Chinese Academy of Sciences. The tested compressor rotor was isolated from the stator to avoid interaction effects generated by the presence of a downstream stator blade row. The isolated compressor rotor selected for this investigation is

actually the rotor of the first stage of a low-speed three-stage axial compressor test rig, which has been used for a number of research programs for the flow instability in compression system. The blading is typical of high-pressure ratio compressor design. Previous work indicates that the isolated rotor is prone to tip stall behavior, which is suitable for flow control methods in the end wall flow regions.

The overall compressor performance in terms of pressure rise coefficient Ψ and mass flow coefficient Φ was measured with eight static pressure taps on casing around the annulus in both the inlet and the outlet of the compressor. The measurement uncertainties were: static pressure, ±60N/m^2. Errors in calculated Ψ and Φ were estimated at ±0.2% maximum, as far as relative comparison between the results for a certain condition is concerned.

The basic principle of using plasma actuation reated caseing(PATC) to improve compressor stability range is shown in Fig. 27. When the PATC is energized, plasma forms and induces airflow along the direction of compressor inflow in the end wall flow region.

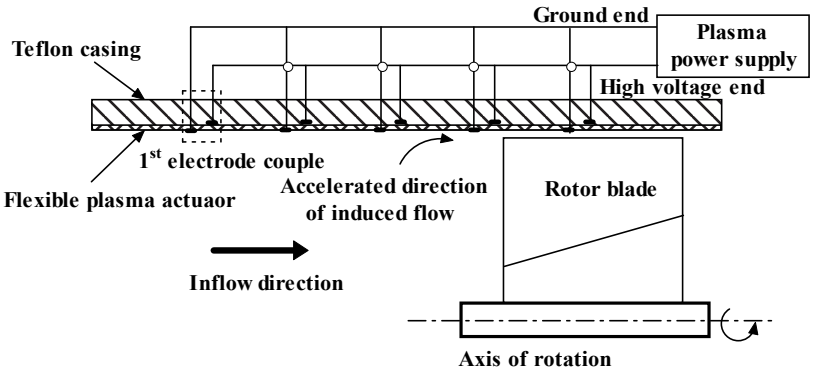


Fig. 27. Sketch map of using PATC to improve compressor stability range

The basic mechanism for plasma actuation to extend compressor's stability can be classified to three effects. The first is that plasma actuation induces air acceleration along with the inflow direction in the blade tip end wall region. Energy is added to the low-energy flow in the end wall region, which can increase mass flux at blade leading edge, inhibit development of blade tip secondary flow and leakage flow, and enhance circulating ability in the end wall region. Thus the accumulation of flow build up is minished. The second is that due to the end wall flow acceleration induced by plasma actuation, velocity in flow direction at blade tip channel is enhanced and inflow attack angle is reduced. Thus flow separation at blade suction surface is inhibited. The last effect is that plasma actuation is non-stationary and non-linear actuation, which can enhance mixture among flow with different momentum in the end wall region. Thus flow separation due to low energy is inhibited and compressor stability is extended. Since plasma actuation can minish flow build up extent in the blade tip end wall region, inhibit secondary flow and leakage flow, and enhance circulating ability, compressor pressure rise ability is improved.

PATC consists of a flexible plasma actuator and a casing. The plasma actuator, layout of which is asymmetrical, consists of 5 electrode couples. The 4th electrode couple is located at 3mm away from the blade leading edge, while the 5th couple is located at the 40% blade tip chord. The thickness of teflon layer, h, is 0.5mm. The electrode is 0.035×2mm copper layer.

The horizontal displacement between upper and lower electrode for each couple, Δd, is 1 mm. The distance between adjacent electrode couples, D, is 10 mm. The casing is also made of teflon. Fig. 28 and Fig. 29 show the PATC and the low speed axial compressor with PATC, respectively. For the PATC and tested rotor, tip clearance is 0.6 mm, which is 1.65% of the blade chord length.

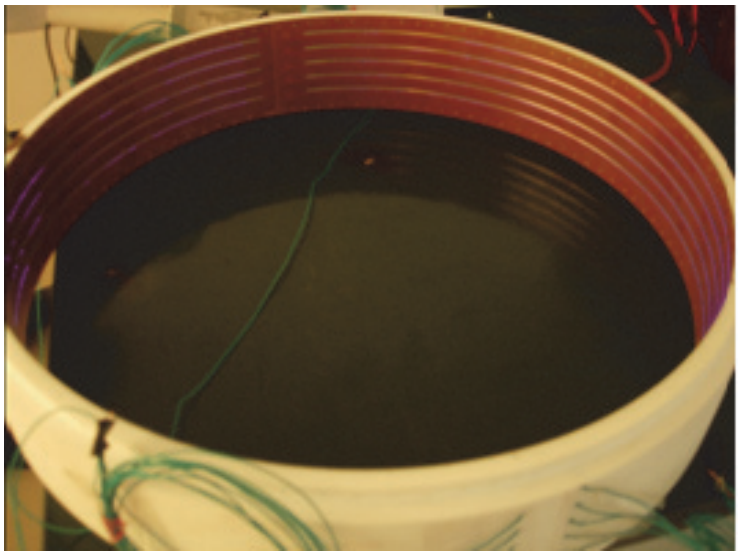


Fig. 28. Plasma actuation treated casing



Fig. 29. Low speed axial compressor test rig with PATC

Plasma actuation casing is energized by a high voltage power supply. The output of the power supply is sine wave. The amplitude and frequency range is 0-30 kV and 6-40 kHz, respectively, which can be adjusted continuously.

The compressor throttling was throttled by the exit rotary cone valve mounted on the shaft and regulated manually when stall was approached. Wall static pressure was collected to calculate pressure rise coefficient Ψ and flow coefficient Φ, which are adopted as representatives of compressor performance and stability with and without plasma actuation at a constant rotor speed.

The effect of plasma actuation on the compressor performance and stability range is displayed in Fig. 30 at the rotor speed of 900 rpm. The 4th electrode couple is actuated and the actuation voltage is 9 kV.

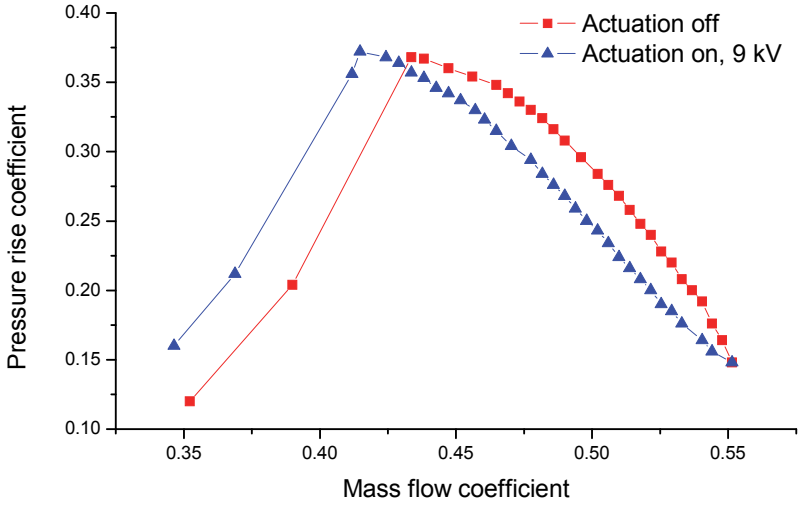


Fig. 30. Test results with and without plasma actuation(rotor speed: 900 rpm)

The changes of maximum pressure rise coefficient, Ψ_{max} and mass flow coefficient near stall, Φ_{ns} are summarized in table 1. The Φ_{ns} decreases by 5.2%, while the Ψ_{max} increases by 1.08%.

	Ψ_{max}	$\Delta\Psi_{max}/\Psi_{max}$	Φ_{ns}	$\Delta\Phi_{ns}/\Phi_{ns}$
0	0.3721		0.4426	
1	0.3761	1.08%	0.4196	-5.2%

0: PATC off, 1: 4th electrode couple on, 9 kV.

Table 1. The effect of plasma actuation on compressor performance and stability range

Fig. 31 illustrates the test results with and without plasma actuation at the rotor speed of 1080 rpm. When the 2nd and 3rd electrode couples are switched on, Φ_{ns} decreases by 1.42% and 5.07% when the actuation voltage is 9 kV and 12 kV respectively. Ψ_{max} decreases by 2.21% and 0.74% respectively.

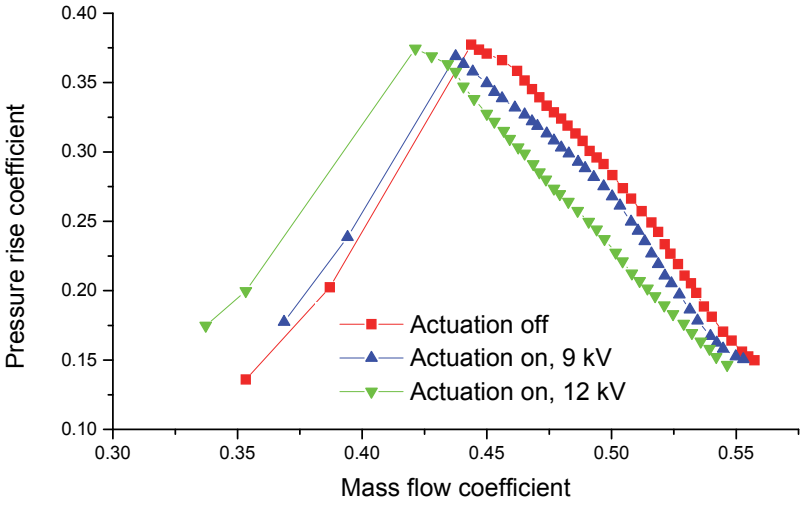


Fig. 31. Test results with and without plasma actuation(rotor speed: 1080 rpm)

Along with the actuation voltage ascending, plasma actuation strength increases while plasma actuator's layout form and material remain same. Thus the velocity of induced flow acceleration increases, which can better enhance the circulation ability in end wall region, reduce inflow attack angle and promote flow mixture. As a result, the compressor stability range is much wider. Therefore plasma actuation strength, which can be improved by adjusting actuator layout form or increasing actuation voltage, is one key factor in plasma based stability extension.

Fig. 32 represents the effect of plasma actuation location on the compressor performance and stability range when the rotor speed equals 1080 rpm. When 3rd and 4th electrode couples are switched on at 12 kV, Φ_{ns} decreases by 1.42% and Ψ_{max} decreases by 1.47%. When the 2nd and 3rd electrode couples are switched on, Φ_{ns} and Ψ_{max} decrease by 5.07% and 0.74%, respectively. Therefore, different actuation location results in different stability range extension effect. One possible reason is that the 4th electrode couple is just 3mm(8.3% of axial chord) away from the rotor blade leading edge, where flow build up is very serious and flow separation has well developed in blade tip end wall region at near stall state. Thus plasma actuation at this location can't control the flow field well and the stability extension effect is limited. When the 2nd and 3rd electrode couple is on, because the 3rd electrode couple is 18mm(49.5% of axial chord) away from the rotor blade leading edge, plasma actuation can accelerate the flow boundary layer before flow separation and build up in well development, which can inhibit the end all separation flow, secondary flow and leakage vortex better. Therefore stability extension effect is much better.

The changes of Ψ_{max} and Φ_{ns} are summarized in table 2. Ψ_{max} decreases at every case when plasma actuation is on. The Ψ_{max} decrease is least when the Φ_{ns} decrease is most. So there is no contradiction between stability range extension and pressure rise coefficient improvement. When the ability for plasma actuation to control the blade tip end wall region flow becomes stronger, the stability extension effect is better and the pressure rise ability almost remains same.

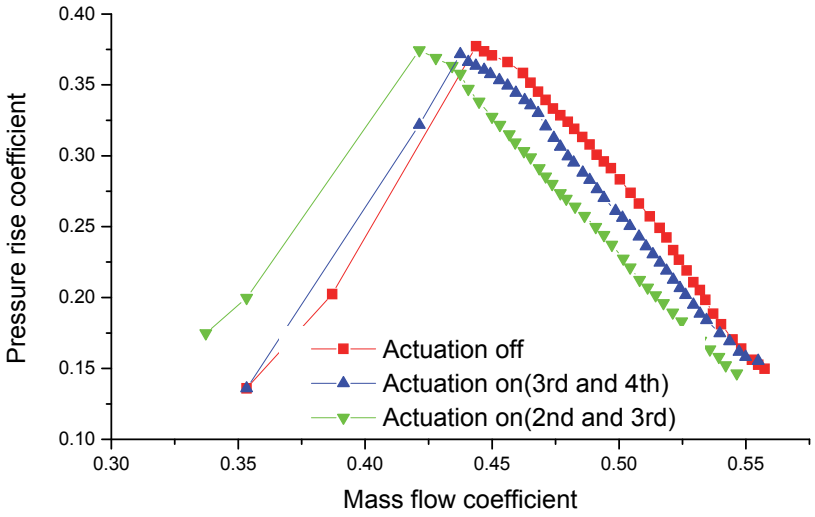


Fig. 32. Test results with and without plasma actuation (rotor speed: 1080 rpm)

	Ψ_{max}	$\Delta\Psi_{max}/\Psi_{max}$	Φ_{ns}	$\Delta\Phi_{ns}/\Phi_{ns}$
0	0.3772		0.4438	
1	0.3689	-2.21%	0.4375	-1.42%
2	0.3744	-0.74%	0.4213	-5.07%
3	0.3717	-1.47%	0.4375	-1.42%

0: PATC off.
1: 2nd and 3rd electrode couples on, 9kV.
2: 2nd and 3rd electrode couples on, 12kV.
3: 3rd and 4th electrode couples on, 12kV.

Table 2. The effect of plasma actuation on compressor performance and stability range

4. Supersonic plasma flow control

Based on plasma aerodynamic actuation, plasma flow control is a novel active flow control technique and has important applications in the field of supersonic flow control. Shock waves are typical aerodynamic phenomena in supersonic flow. If they are controlled effectively, the aerodynamic performance of both flight vehicles and aeroengines will be greatly enhanced. Conventional mechanical or gasdynamic control methods have disadvantages of complex structure and slow response. Novel plasma flow control method has advantages of simple structure, fast response and wide actuation frequency range. Therefore, plasma flow control method has become a newly-rising focus in the field of shock wave control.

4.1 Experimental principle and arrangement

Fig. 33 shows the MHD flow control experimental principle. The high density plasma column which primarily consists of ions and electrons was generated between a pair of electrodes

through pulsed DC discharge. There were three pairs of electrodes and an oblique shock wave appeared in front of the ramp in low-temperature supersonic flow. The alphabet "I" and "B" represented the current and magnetic field. The arrows gave their directions.

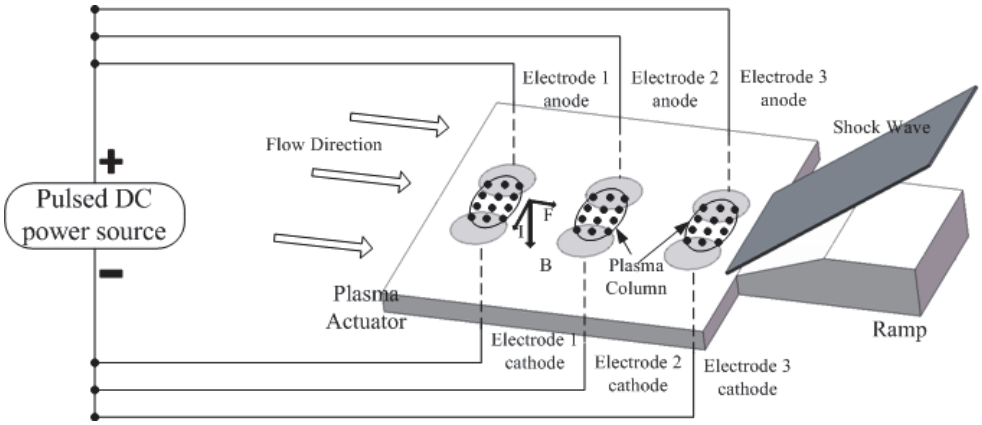


Fig. 33. The experimental principle

When magnetic field, normal to the surface, was imposed on the plasma column created in the boundary layer, it affected both the plasma and, through the Lorentz body force (j×B body force), the flow. The direction of Lorentz body force was determined by the directions of current and magnetic field. The alphabet "F" represented the Lorentz body force which could accelerate the flow.

The plasma column was produced by pulsed DC discharge. Therefore the plasma would be influenced by electric field force, magnetic field force and the airflow inertial force. The magnetic field force and the airflow inertial force were dominant. When the direction of magnetic field force was same as that of airflow inertial force and the velocity of plasma was faster than that of the neutral gas molecules, the plasma would strike the neutral gas molecules to transfer momentum and accelerate the flow in the boundary layer. Otherwise, when the direction of magnetic field force was against with that of airflow inertial force, the plasma would strike the neutral gas molecules to transfer momentum and decelerate the flow in the boundary layer.

MHD flow control system consisted of low-temperature supersonic wind tunnel, plasma actuation system, experimental ramp, magnetic field generator, parameter measurement system and schlieren optical system. The inlet total pressure of low-temperature supersonic wind tunnel was about 5-7atm. The stagnation conditions for the tunnel were atmospheric pressure and room temperature. The run time could reach up to 60 seconds dependent on the inlet total pressure. The experimental duct was 115mm(length)×80mm(width) and the designed Mach number was 2.2. The static pressure was 0.5-0.7atm and the static temperature was 152K.

The plasma actuation system included pulsed DC power source, plasma actuator, insulating acrylic base. Pulsed DC power source was the critical equipment which consisted of high voltage pulsed circuit, high voltage DC circuit and feedback circuit. It could provide 0-90kV selected high voltage pulse and 0-3kV selected high voltage direct current. The electrodes were made of plumbago, and were flush-mounted on the top wall of the insulating

dielectric. The diameter of the electrode was 10mm. The insulating dielectric was made of BN ceramic. Two kinds of arrangements for the plasma actuator were adopted according to the distance between a pair of electrodes (D=5mm or D=8mm).

The experimental ramp was also constructed of insulating acrylic material. It was installed on the insulating acrylic base. The dimension of the ramp was 34mm(length) ×25mm(width) ×6mm(height). One angle of ramp was A=15°and another was A=20°. As illustrated in Fig. 34, 10 static pressure measurement holes were drilled on the acrylic base, the plasma actuator and the ramp. The holes were numbered with k1-k10 from upstream to downstream. Holes k2-k8 were drilled on the plasma actuator, and holes k9-k10 were drilled on the ramp. The diameter of k1-k10 was 0.5mm. On the plasma actuator the distance between adjacent holes was 10mm except that the distance between k2 and k3 was 7.5mm and the distance between k7 and k8 was also 7.5mm. On the ramp the distance between k9 and plasma actuator edge was 6mm and the distance between k9 and k10 was 6mm.

Nd-Fe-B rare-earth permanent magnets were used as the magnetic field generator which was located normal to experimental duct. The direction of magnetic field was perpendicular to the flow direction and the electric field direction. The magnetic field strength was about 0.3T between two magnets.

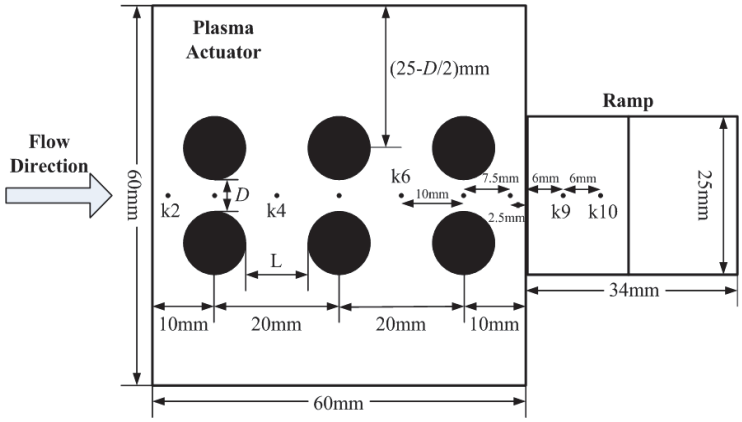


Fig. 34. The dimensions of the plasma actuator and the ramp

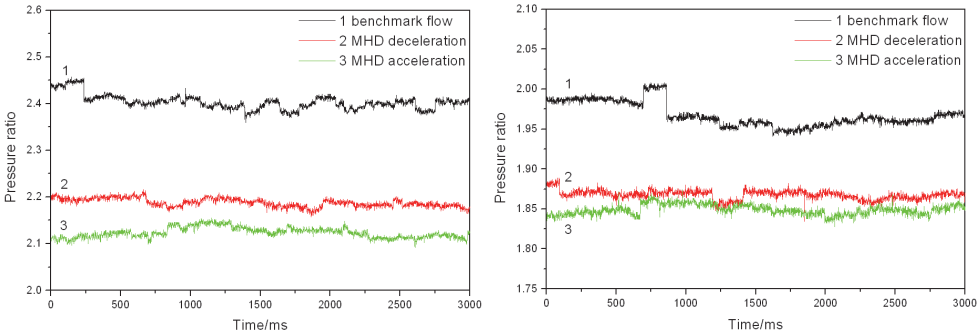
Parameter measurement system consisted of electric parameter and flow characteristic measurement systems. Electric parameter measurement system included oscillograph (DPO4104, Tektronix Inc.), high voltage probe (P6015A, Tektronix Inc.) and current probe(TCP312+TCPA300, Tektronix Inc.). Flow characteristic measurement system included 10 static pressure sensors and a data collection apparatus. Because the run time of the wind tunnel was above 10 seconds in every experiment, the inlet total pressure decreased slowly during the experimental time. Therefore, in this study the ratio of Pitot pressure after shock-wave to that before shock-wave was adopted to compare the flow characteristic of the airflow around the ramp that was illustrated as P_{k10}/P_{k7}.

Schlieren optical system consisted of a high-speed camera and a storage computer. The high-speed camera was an Optronis® high-speed camera, and the maximum frame frequency was 200k frames per second(FPS). In this study, the schlieren pictures were taken at 8kFPS. The exposure time was 100μs and the run time was 8s.

4.2 Experimental results

In terms of different magnetic field directions, distances between electrodes, ramp angles and DC Voltage, the results of these four kinds of MHD flow control were compared and analyzed.

Through changing the magnetic field direction, MHD acceleration and MHD deceleration experiments were carried out. Fig. 35 represented two typical flow characteristics.



(a) D=5mm, A=20° (b) D=8mm, A=15°

Fig. 35. Two typical flow characteristics

The time-averaged pressure ratio decreased by 6.04% and 5.09% with MHD acceleration and MHD deceleration respectively. Thus, MHD flow control could drastically weaken the oblique shock wave strength and change the flow characteristic of the airflow around the ramp. MHD acceleration was more effective than MHD deceleration.

MHD acceleration experiments were carried out at ramp angle A=20°with D=8mm or 5mm. Fig.36 showed the flow characteristics with different distances. At D=8mm the time-averaged pressure ratio decreased by 19.66% with MHD acceleration. At D=5mm the time-averaged pressure ratio decreased by 11.64% with MHD acceleration. Thus, MHD acceleration was more effective to weaken the shock wave strength when D increased, but there existed a maximum value D_{max} exceeding which the airflow could not be ionized restricted by power source.

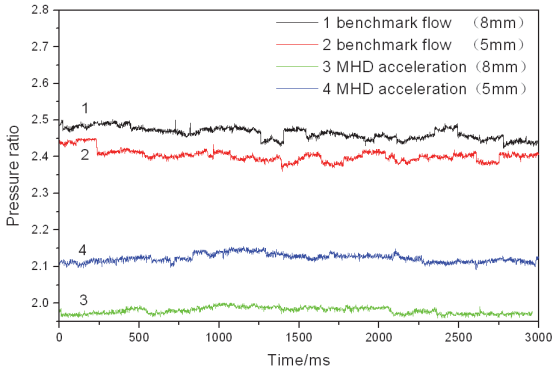


Fig. 36. Flow characteristics with different distances

MHD acceleration experiments were carried out at D=8mm with different ramp angles A=15°or 20°. Fig. 37 showed the flow characteristics with different ramp angles. At A=15° the time-averaged pressure ratio decreased by 6.04% with MHD acceleration. At A=20° the time-averaged pressure ratio decreased by 19.66% with MHD acceleration. Thus, MHD acceleration was more effective to weaken the shock wave strength when A increased, but there existed a maximum A_{max} exceeding which oblique shock wave strength would be too strong to be changed.

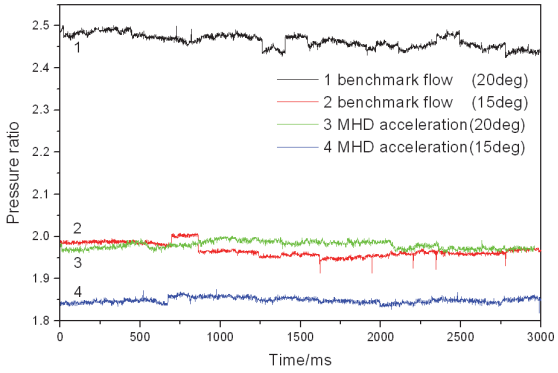
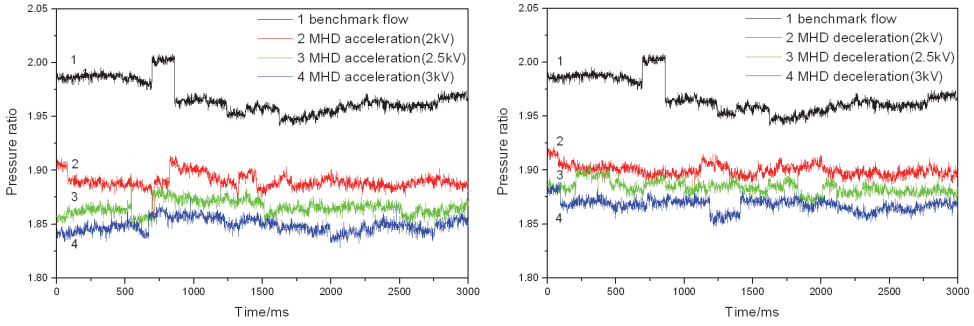


Fig. 37. Flow characteristics with different ramp angles

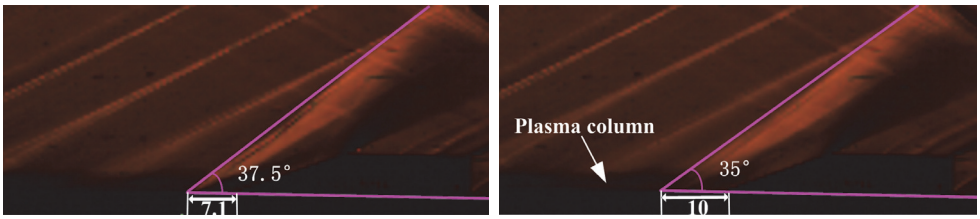
Through changing DC voltage V_{DC}(2kV, 2.5kV, 3kV), MHD acceleration and MHD deceleration experiments were carried out at ramp angle D=8mm, A=15°. Fig. 38 showed the flow characteristics with different V_{DC}. Fig. 38(a) showed the static pressure ratio varying with MHD acceleration. The time-averaged pressure ratio decreased by 3.95%, 5.19% and 6.04% with V_{DC}=2kV, 2.5kV and 3kV respectively. Fig. 38(b) showed the static pressure ratio varying with MHD deceleration. The time-averaged pressure ratio decreased by 3.44%, 4.26% and 5.09% with V_{DC}=2kV, 2.5kV and 3kV respectively. Thus, MHD flow control could drastically weaken the oblique shock wave strength and change the flow characteristic of the airflow around the ramp. MHD interaction was more effective when V_{DC} increased.



(a) MHD acceleration (b) MHD deceleration

Fig. 38. Flow characteristics with different V_{DC}

The schlieren pictures were captured during the time of flow characteristic measurement at all conditions. The typical schieren pictures were illustrated in Fig. 39 at D=8mm, A=20° with MHD acceleration. Fig. 39(a) showed the benchmark flow with no electric field and no magnetic field while Fig. 39(b) showed the flow with electric field and magnetic field in the case of MHD acceleration. Four shock waves before the ramp were shown in the benchmark picture, three of which were produced by the coarse interface between electrodes and ceramic and the last shock wave was produced by the ramp. Without MHD acceleration the shock wave angle near the ramp was about 37.5°, and the distance between the shock wave location and the ramp was about 7.1mm. With MHD acceleration the shock wave angle near the ramp was reduced to 35°, and the distance was increased to 10mm which meant the shock wave was moved upstream by 2.9mm. Therefore, MHD flow control could change shock wave location, convert a strong shock wave into many weak shock waves, weaken the shock wave strength and change the flow characteristic near the ramp.



(a) Benchmark flow (b) MHD acceleration flow

Fig. 39. Benchmark and MHD acceleration flow at a typical condition

5. Conclusion

The principle of plasma aerodynamic actuation and its application in subsonic and supersonic flow control was summarized. The mechanism for plasma flow control can be summarized as momentum effect, shock effect, and chemical effect. Both the plasma and flow characteristics of the plasma aerodynamic actuation were investigated. Plasma flow control used in airfoil separation control, corner separation control, axial compressor stability extension and shock wave control were studied.

6. Acknowledgment

The authors thank Yikang Pu, Shouguo Wang, Junqiang Zhu and Chaoqun Nie for the help in the experiment and analysis. Support for this work was received from the National Natural Science Foundation of China (50906100, 10972236, 51007095) is gratefully acknowledged.

7. References

Corke, T., Post, M. & Orlov, D. Dielectric Barrier Discharge Plasma Actuators for Flow Control, *Progress in Aerospace Sciences*, Vol.43, (January 2010), pp. 505-529, ISSN 0376-0421

Moreau, E. Airflow Control by Non-thermal Plasma Actuators, *Journal of Physics D: Applied Physics*, Vol.40, No.3, (January 2007), pp. 605-636, ISSN 0022-3727

Roth, J. Aerodynamic Flow Acceleration Using Paraelectric and Peristaltic Electrohydrodynamic Effects of a One Atmosphere Uniform Glow Discharge Plasma, *Physics of Plasmas*, Vol.10, No.5, (May 2003), pp. 2117-2126, ISSN 1070-664X

Bletzinger, P., Ganguly, B. & VanWie, D. (2005). Plasmas in High Speed Aerodynamics, *Journal of Physics D: Applied Physics,* Vol.38, No.4, (April 2005), pp. R33–57, ISSN 0022-3727

Wu, Y., Li, Y., Jia, M., Song, H., Guo, Z. G., Zhu, X. & Pu, Y. Influence of Operating Pressure on Surface Dielectric Barrier Discharge Plasma Aerodynamic Actuation Characteristics, *Applied Physics Letters*, Vol.93, No.3, (July 2008), pp. 031503(3p), ISSN 0003-6951

Li, Y., Wu, Y., Zhou, M., Su, C., Zhang, X. & Zhu, J. Control of the Corner Separation in a Compressor Cascade by Steady and Unsteady Plasma Aerodynamic Actuation, *Experiments in Fluids*, Vol.48, No.6, (June 2010), pp. 1015-1023, ISSN 0723-4864

Li, Y., Wu, Y., Jia, M., Zhou, Z., Guo, Z., & Pu, Y. Optical Emission Spectroscopy Investigation of a Surface Dielectric Barrier Discharge Plasma Aerodynamic Actuator, *Chinese Physics Letters*, Vol.25, No.11, (November 2008), pp. 4068-4071, ISSN 0256-307X

Wu, Y., Li, Y., Jia, M., Song, H., Su, C. & Pu, Y. Experimental Investigation into Characteristics of Plasma Aerodynamic Actuation Generated by Dielectric Barrier Discharge, *Chinese Journal of Aeronautics*, Vol.23, No.1, (January 2010), pp. 39-45, ISSN 1000-9361

Li, Y., Wang, J., Wang, C., An, Z., Hou, S. & Xing, F. Properties of Surface Arc Discharge in a Supersonic Airflow, *Plasma Sources Science and Technology*, Vol.19, No.2, (March 2010), pp.025016(9p), ISSN 0963-0252

Wang, J., Li, Y. & Xing, F. Investigation on Oblique Shock Wave Control by Arc Discharge Plasma in Supersonic Airflow, *Journal of Applied Physics*, Vol.106, No.7, (October 2009), pp.073307(7p) , ISSN 0021-8979

Wang, J., Li, Y., Cheng, B., Su, C., Song, H. & Wu, Y. Effects of Plasma Aerodynamic Actuation on Oblique Shock Wave in a Cold Supersonic Flow, *Journal of Physics D: Applied Physics*, Vol.42, No.16, (June 2009), pp. 165503(8p), ISSN 0022-3727

Su, C., Li, Y., Cheng, B., Wang, J., Cao, J. & Li, Y. Experimental Investigation of MHD Flow Control for the Oblique Shock Wave Around the Ramp in Low-temperature Supersonic Flow, *Chinese Journal of Aeronautics*, Vol.23, No.1, (January 2010), pp. 22-32, ISSN 1000-9361

Wu, Y., Li, Y., Zhu, J., Su, C., Liang, H. & Li, G. (2007). Experimental Investigation of a Subsonic Compressor with Plasma Actuation Treated Casing, *37th AIAA Fluid* Dynamics Conference and Exhibit, pp. 1-8, Miami, FL, USA, June 25-28, 2007

Li, Y., Wu, Y., Liang, H., Song, H. & Jia, M. (2010), The mechanism of plasma shock flow control for enhancing flow separation control capability, Chinese Sci Bull (Chinese Ver), vol. 55, No. 31, (November 2010), pp. 3060- 3068, ISSN 0023-074X

Li, Y., Liang, H., Ma, Q., Wu, Y., Song, H., & Wu, W.(2008). Experimental investigation on airfoil suction side flow separation suppression by pulse plasma aerodynamic actuation, Acta Aeronautica et Astronautica Sinica (Chinese Ver), vol. 29, No. 6, (November 2008), pp. 1429-1435, ISSN 1000-6893

3

Nonequilibrium Plasma Aerodynamics

Andrey Starikovskiy[1] and Nickolay Aleksandrov[2]
[1]*Princeton University*
[2]*Moscow Institute of Physics and Technology*
[1]*USA*
[2]*Russia*

1. Introduction

Currently, the problem of flow active control by low-temperature plasma is considered to be one of the most booming realms of aerodynamics [Bletzinger et al, 2005, Moreau, 2007]. One of the main practical advantages of plasma-assisted flow control is its zero reaction time. This, in principle, provides control over a wide range of frequencies as well as covering a wide range of processes beginning from stationary to separated and turbulent flows. Two other important advantages of plasma systems are their low weights and sizes. If combined with relatively low energy consumption, the features allow the possibility of principally developing new systems of flight control at high velocities [Nikipelov, 2009].

Promising applications of plasma systems to control flow involve the management of laminar-to-turbulent transition in the boundary layer, the management of flow detachment or attachment to the surface, and the resultant management of lift and drag force of an airfoil. This also provides management of noise and vibrations, and control of shock wave patterns as well as their interactions with the boundary layer.

One of the first attempts of flow control using nonequilibrium plasma was performed more than 40 years ago [Mhitaryan et al, 1964]. The authors investigated the flow separation control by surface dielectric barrier discharge with sinusoidal high-voltage power supply in the frequency range f = 50 – 570 Hz. Figure 1 shows the scheme of the experimental installation [Mhitaryan et al, 1964].

The authors [Mhitaryan et al, 1961, 1964] demonstrated up to 30% drag force decrease and 40% lift force increase for airfoil with 460 mm chord length (AoA = 8^0), thickness of 13% and flap length 30% (AoA = 43^0) of the chord in velocity range from 8.75 to 20.4 m/s. They demonstrate that the parameter $c_{\mu_e} = \dfrac{2P_a \cdot l}{\rho V^2 \cdot b \cdot U \cdot S \cdot \cos\varphi}$ (discharge power density) controls the discharge efficiency. Here P_a is active power of the discharge, l – plasma layer length, V – free stream velocity, ρ - gas density, b – ions mobility, U – discharge voltage, S – airfoil surface, φ - phase shift. The authors made a conclusion that the actuator affects the flow through ionic wind mechanism when induced velocity was in the order of 20-25% from the velocity of free stream.

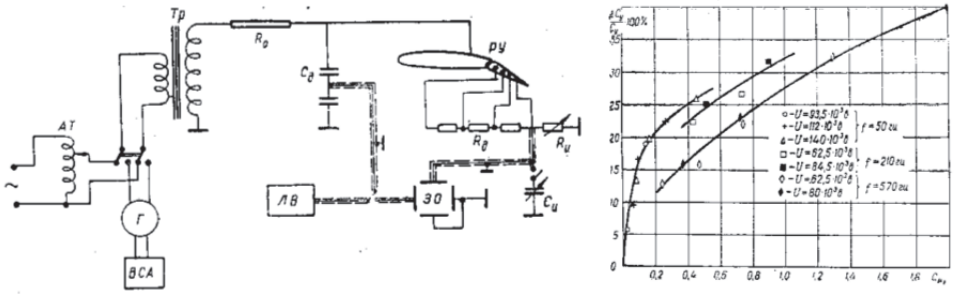


Fig. 1. a) Scheme of the experimental installation [Mhitaryan et al, 1964].
Tp – high-voltage transformer; R_0 – limiting resistor; C_d – capacitive voltage attenuator; R_d – reactive attenuator; PY – plasma generation region; R_u – resistor and C_u – capacitor for electrical measurements; ЗО – oscillograph; ПВ – voltmeter; Г – generator; BCA – reference DC power supply; AT – autotransformer.
b) – relative lift force increase.

The thermal management constraints mean that hypersonic flight at Mach numbers greater than 10 can be achieved only at very high altitudes. That is why the plasma mechanisms of external flow modification, shock wave reconfiguration and drag force reduction were considered as an important alternative to traditional solutions for thermal management [Bletzinger et al, 2005].
The paper [Klimov et al, 1982] was one of the first articles devoted to the impact of plasma effects on shock wave propagation. The authors observed an increasing velocity of shock waves propagating in the gas discharge tube with simultaneous amplitude decrease (Figure 2). They compared the observed increase of shock wave velocity with the calculated values corresponding to the discharge heat release. The difference obtained (1200-1300 m/s instead of 900 m/s) was explained by a possible release of vibrational energy, quenching of electronically excited states and double layer formation in front of the SW.

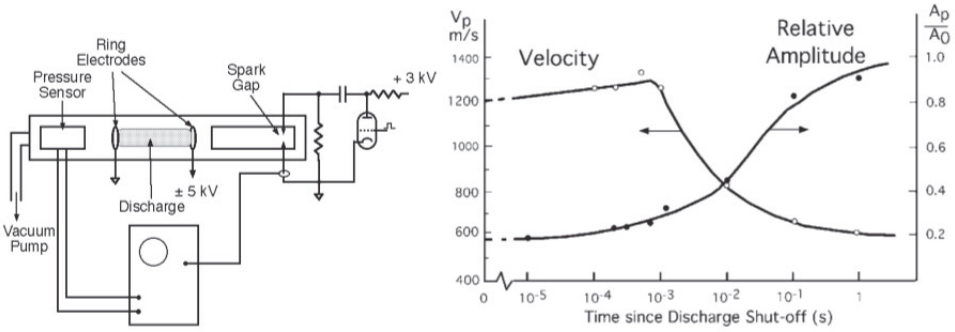


Fig. 2. a) Schematic of experiment of [Klimov et al, 1982]; b) Dependence of the velocity and relative amplitude of the shock wave in the decaying plasma on the delay time since switch-off of the discharge. A_p and A_0 are the amplitudes of the shock waves in air and in the plasma, respectively.

A number of imaginative schemes were proposed for modifying and controlling the flow around a hypersonic vehicle. These schemes included novel approaches for plasma generation, magnetohydrodynamic (MHD) flow control and power generation and new hot gas counter-flows and other purely thermal approaches [Bletzinger et al, 2005]. The first of these highly innovative concepts of the aircraft called AJAX (Ayaks) was developed by Vladimir Freishtadt in the late 1980s. The project of this new hypersonic, air-breathing aircraft is currently under development by the Leninetz Holding Company in St. Petersburg, Russia [Gurijanov&Harsha 1996]. Schematics of AJAX concept is shown in Figure 3.

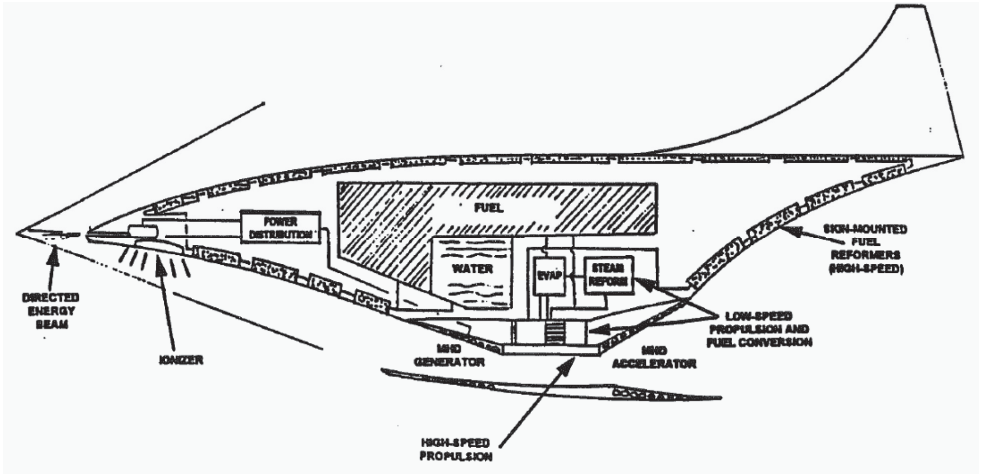


Fig. 3. AJAX concept [Gurijanov&Harsha 1996].

The AJAX was projected to employ a novel engine that uses some form of MHD generator to collect and slow down highly ionized and rarefied air. The air is mixed with fuel into the mixture that burns in the plasma-stabilized combustor, while the electricity produced by the inlet's MHD generator is used in the magneto-plasmadynamic thruster to provide additional thrust. The plasma funnel developed over the air inlet from the magneto-hydrodynamics forces greatly increases the ability of the engine to collect air, increasing the effective diameter of the air inlet.

Further increase of interest to plasma assisted flow control took place in 1993 when Roth and co-workers re-invented surface discharge application for boundary layer separation control. The paper [Roth, 2003] demonstrates the AC dielectric surface discharge ability to suppress the flow separation in low-speed regimes. Figure 4 demonstrates the reattachment of an air flow at AoA = 8^0 and free stream velocity of 2.85 m/s.

Another exciting experiment was presented by Bedin and Mishin in [Bedin&Mishin, 1995]. Ballistic range tests were conducted to measure the drag of a sphere flying through gas discharge plasmas. Experiments were conducted using 15 mm polyethylene spheres fired at speeds between 200 and 1350 ms^{-1} in air at 15 Torr. A RF generator operating at 440 kHz was used to produce weak plasma ($a \le 10^{-6}$) over a length of 3 m. The gas temperature in the discharge was approximately 1140 K and the electron temperature was estimated as T_e~2–4 eV. The authors report anomalous shock standoff distances for the spheres within the discharge zone at distances far from the thermal nonuniformities near the entrance to

the plasma. Time-of-flight measurements were made to determine the drag coefficient of the sphere with an accuracy of 1% in ambient air and 3% in plasma. The resulting drag coefficient *(CD)* measurements, shown in Figure 5, indicate a substantial decrease in the drag coefficient at subsonic speeds, with a small increase at supersonic speeds. The decrease in *CD* observed at subsonic speeds in the presence of the plasma is inconsistent with that expected with the higher temperatures and lower Reynolds number obtained in these experiments. The authors indicated that the observed effects were due to a higher velocity with disturbances transferred throughout the flowfield in the presence of plasma.

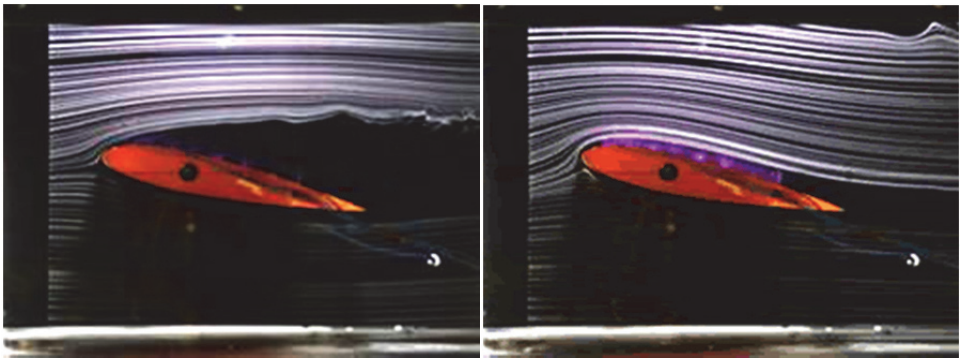


Fig. 4. Low-speed flow separation control by surface dielectric barrier discharge plasma. NACA-0015 airfoil, V = 2.85 m/sec, AoA = 8^0 [Roth, 2003].

Flow control opportunities by plasma could be summarized into two groups: 1) main flow control, including shock wave configuration in supersonic and hypersonic regimes; 2) boundary layer control.

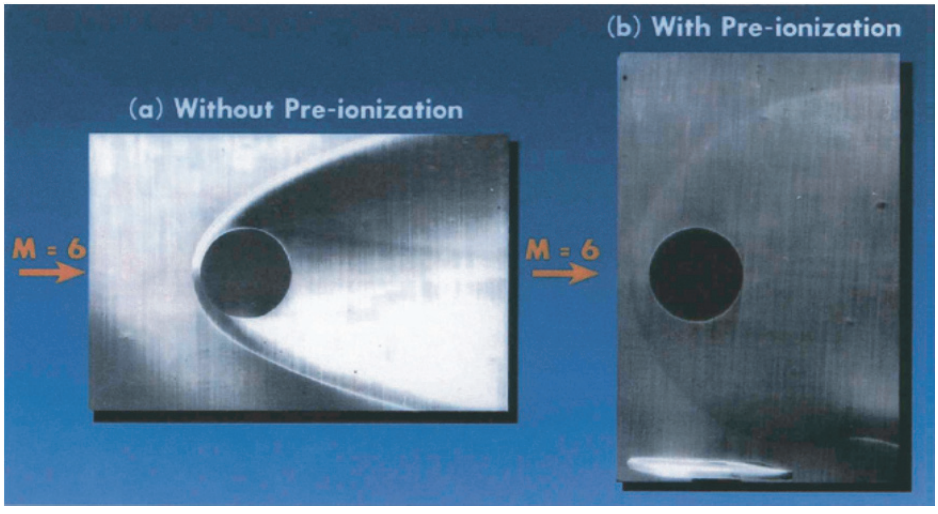


Fig. 5. Schlieren photographs of flow around sphere: (*a*) flight through air without discharge and (*b*) flight through the positive column of a glow discharge in air [Bedin&Mishin, 1995].

Main flow control includes shock wave pattern control; aerodynamic breaking; drag reduction; heat mitigation; flow vectorization, acceleration and deceleration; MHD power extraction and breaking.
Boundary layer control could be subdivided into laminar-turbulent transition control; boundary layer separation control; lift and drag force control; acoustic noise control; mixing enhancement.
Nonequilibrium plasma also may be very efficient in ignition and flame stabilization control; engine performance enhancement, including possibility of fast initiation of detonation waves and transition from Braiton ($P = const$) to more efficient Humphrey ($V = const$) thermodynamic cycle [Starikovskii, 2005; Starikovskii et al, 2008].
Three different physical mechanisms control the efficiency of plasma aerodynamics. There are: 1) gas heating; 2) electrostatic momentum transfer to the gas; 3) magneto-hydrodynamic effects, including MHD flow acceleration and on-board electricity generation using gas flow kinetic energy. Sometimes it is necessary to take into account gas dissociation and change of average molecular weight of the gas, but strong dissociation or ionization of the gas require huge amount of the energy. That is why for aerospace applications we generally trying to keep the gas excitation on the minimal possible level, which decreases the role of variation of gas composition in the discharge.
Increased interest to plasma assisted flow caused an increased publications flux. There were a lot of impressive demonstrations of plasma-assisted flow control during last 10 years. A lot of them were summarized in several topical reviews published recently. For example, possibilities of supersonic flow control (including shock wave attenuation and shock configuration change) by plasma were discussed in [Bletzinger et al, 2005]. Drag reduction at high speed by energy deposition was discussed in paper [Knight, 2008]. Boundary layer separation control by non-equilibrium non-thermal plasma was reviewed in [Zouzou et al, 2006] and [Corke et al, 2010]. In addition, the recent review [Miles, 2011] includes a lot of results of all mechanisms (thermal, electrostatic and MHD) applications for flow control for subsonic and supersonic regimes.
That is why in our current review we will only briefly mention the most important results in plasma assisted aerodynamics and will focus on the physical mechanisms which limit or promote our abilities to control the flow field by plasma.

2. Supersonic flow and shock waves control

On the whole, two approaches can be distinguished to explain the nature of the supersonic plasma-flow interaction. The former describes the process in terms of the thermal mechanism of interaction connected with the energy release in the discharge. Meanwhile, the latter considers the momentum transfer from the electric field to gas flow to be the main mechanism of the process, this is a pure plasma mechanism due to charged components; it is caused by uncompensated spatial charge and is related to volume forces.
The above-mentioned paper [Klimov et al, 1982] demonstrated an increasing velocity of shock waves propagating in the gas discharge tube with simultaneous amplitude decrease. The authors compared the observed increase of shock wave velocity with the calculated values corresponding to the discharge heat release. The difference obtained (1200-1300 m/s instead of 900 m/s) was explained by a possible release of vibrational energy, quenching of electronically excited states and double layer formation in front of the SW. However, later works questioned such interpretation [Meyer et al, 2003; Miles et al, 2001]. Gas heating in discharge was considered the main effect responsible for the changes of flow parameters (Figure 6).

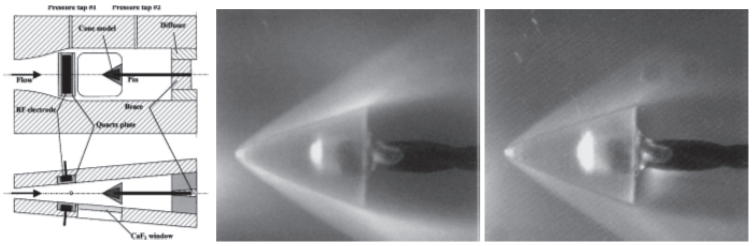


Fig. 6. a) Schematic of the supersonic RF discharge/test section: side view and top view. b-c) Photographs of the supersonic flow over a cone visualized by plasma generated by DC and RF discharges in a 30% N_2 –70% He flow at P_0 = 250 Torr: c) only 175 W DC discharge is on and d) both 175 W DC and 250 W RF discharge are on [Meyer et al, 2003].

In the paper [Meyer et al, 2003] an experimental study of shock modification in an M = 2.5 supersonic flow of nonequilibrium plasma over a cone is discussed. The experiments are conducted in a nonequilibrium plasma supersonic wind tunnel using a supersonic plasma flow over a quasi-two-dimensional wedge. The experiments showed that an oblique shock can be considerably weakened by transverse RF discharge plasma. The observed shock weakening has been found consistent with a temperature rise in the boundary layers heated by the discharge.

Spark-generated shock waves were studied in [Miles et al, 2001] in glow discharges in argon and argon–nitrogen mixtures. A direct proof of the thermal mechanism of shock wave-plasma interaction was obtained for pulsed discharges. With a sub-millisecond delay between starting the discharge and shock launch, plasma parameters reach their steady-state values, but because of low discharge power the temperature increase was low during several milliseconds. It was shown that laser schlieren signals are virtually identical to those without the discharge, differing dramatically from the signals in discharges with fully established temperature profiles (Figure 7).

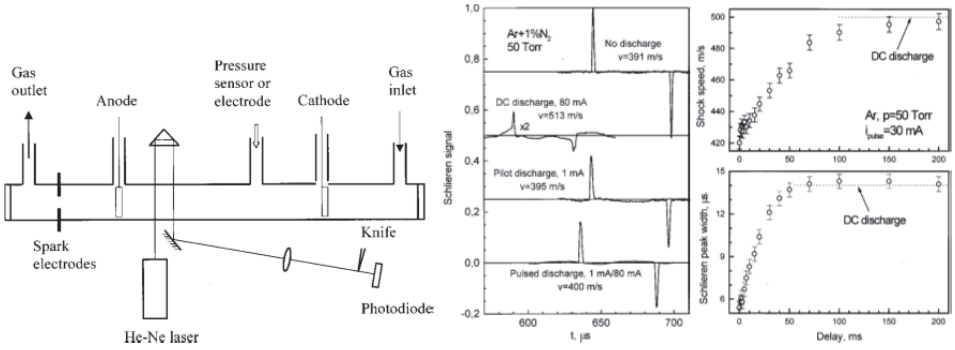


Fig. 7. a) Schematic of the experimental setup for studies of shock propagation in glow discharges; b) Schlieren signals and shock velocities in a glow discharge in Ar + 11%N2 at 50 Torr: the discharge off; 80-mA continuous dc discharge, weak (I ~ 1 mA) pilot discharge, and a pulsed discharge; c) Evolution of shock velocity and the width of the schlieren signal at a fixed location inside the discharge with time delay between the beginning of the discharge and the shock launch [Macheret et al, 2001].

Nevertheless, an important question about possible peculiarities of the discharge and plasma properties in molecular gases still remained. Strong excitation of internal degrees of freedom of molecular gases, electrons attachment and lower electron temperatures (strong influence of electrical double layer under low ionic Mach number) could lead to significant influence of non-equilibrium excitation on plasma-shock interaction. The clarification was done in papers [Khorunzhenko et al, 2003; Anokhin et al, 2004].

Papers [Khorunzhenko et al, 2002,2003] describe the results of measurements of the structure of the strong shock waves in plasma with strong electric fields (E/n ~ 10 kTd). The authors obtained profiles of gas rotational temperature, distribution of electron number density and electric field in air flow at M = 8.2. Comparison of the experimental results and numerical modeling showed that drag change can be totally explained by gas heating during energy thermalization under plasma relaxation (Figure 8).

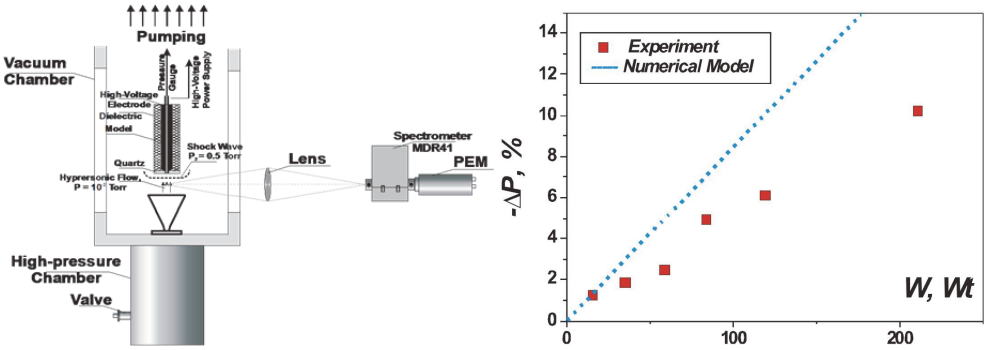


Fig. 8. a) Organization of hypersonic plasma flow. b) Comparison of the experimental results and numerical modeling of stagnation pressure variation vs discharge power in M = 8.2 air flow [Khorunzhenko et al, 2003].

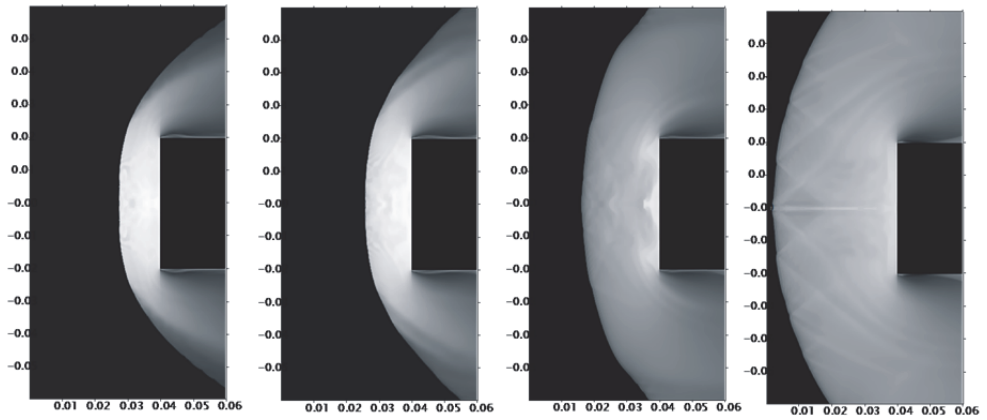


Fig. 9. Flow field calculation for different discharge energy deposition. a) discharge is OFF; b) E = 0.044 eV/mol; c) E = 0.049 eV/mol; d) E = 0.054 eV/mol. Air, M = 8.2, T_0 = 300 K, P_0 = 35 Torr, T_1 = 21 K, P_1 = 5·10^{-3} Torr [Khorunzhenko et al, 2003].

Figure 9 demonstrates the computed flow field change for different discharge power. Main energy release occurs behind shock wave front due to higher gas number density, and, consequently, higher energy deposition by discharge, in electron-ion recombination processes.

Paper [Anokhin et al, 2004] presented the investigations of shock wave propagation through the pulsed nanosecond discharge decaying air plasma (E/n ~ 600-800 Td). It was discovered that the velocity of the shock wave propagation increases as the delay time increases. The delay time here means the time between discharge switching off and shock wave entering into the plasma region. This observation made it possible to prove the thermal nature of such interaction and estimate the plasma thermalization time (Figure 10).

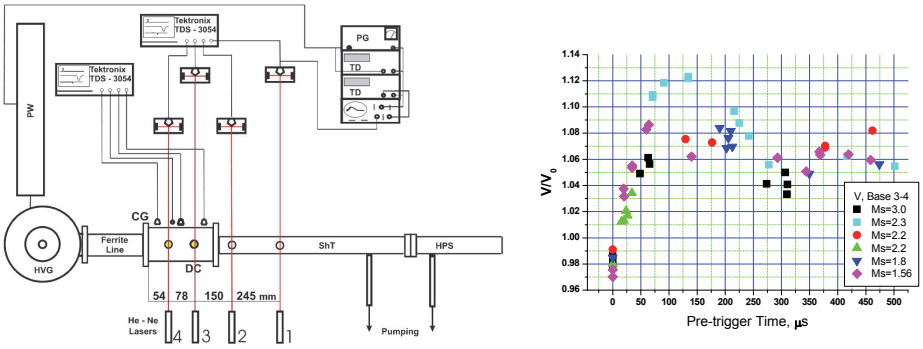


Fig. 10. Plasma shock tube schematics a) and shock wave acceleration in the plasma region in dependence on the plasma relaxation time [Anokhin et al, 2004].

Thus the results of papers [Miles et al, 2001, Meyer et al, 2003, Khorunzhenko et al, 2003; Anokhin et al, 2004] clear demonstrate the thermal nature of the shock wave – plasma interaction. Theory of shock wave interaction with hot layers (and, in fact, with entropy layers of any nature) was developed long ago.

Publications by group of Nemchinov (see, for example, [Nemchinov et al, 1989]) show that the length of the pre-cursor for sufficient overheating of hot layer depends linearly on the shock wave's run-up distance along the hot layer:

$$l \approx U_s t$$

This relation was checked up to $U_s\ t$ = (80-90)h for planar thermal layer (2D flow), and $U_s\ t$ = (300 – 400)d for thermal channel (axisymmetric). This means that the mechanism should exist which will provide the gas accumulation in the precursor zone both from the hot layer and undisturbed flow region. One of possible variants of such a structure is the hot layer detachment from the axis or plane of symmetry and vortex structure formation.

In [Nemchinov et al, 1989] it was found that up to overheat value of $T/T' \cong 2.5$ the ratio of the precursor length and the shock wave run-up distance along the hot layer is in the range of ξ = 0.02÷0.05 (M = 3, γ = 1.4), and only starting from the overheat value of $T/T' > 2.5$ demonstrated the dependence from the overheat value, reaching the value ξ = 0.29 at $T/T' > 5.5$.

Initially the criteria of the flow rearrangement has been proposed by Mark [Mark, 1957] and Griffiths [Griffiths, 1957] right after the first tests of the ultra-powerful bombs, where the emission from the epicenter causes a hot layer formation above the Earth surface and

intensive blast wave interaction with this layer. A Mark-Griffiths criterion was based on the comparison of the stagnation pressure of the gas from entropy layer and static pressure of the gas from main flow. It was assumed that shock wave velocity in entropy layer and outside the layer are the same at least until the moment when critical conditions appear. One can see that the condition of equal velocity of the shock wave front in all regions de-facto forbids the flow rearrangement even when the Mark-Griffiths criterion will meet. It indicates on the internal contradictions in the model. Indeed, the Mark-Griffiths criterion at M = 3, γ = 1.4 estimates the threshold of a global flow restructuring close to $T/T' > 1.18$, which is significantly different from the values observed ($T/T'_{crit} \cong 2.5$).

Paper [Lashin&Starikovskii, 1995] suggests a modification of the Mark-Griffiths criterion. As criterion of flow modification from initial stable shock wave propagation to self-similar precursor growth to conditions should be satisfied:

1. Static pressure behind the SW in the entropy layer should be equal to the static pressure behind SW in the main flow;
2. Gas velocity behind the SW in the entropy layer should be equal to the velocity of SW propagation in the main flow.

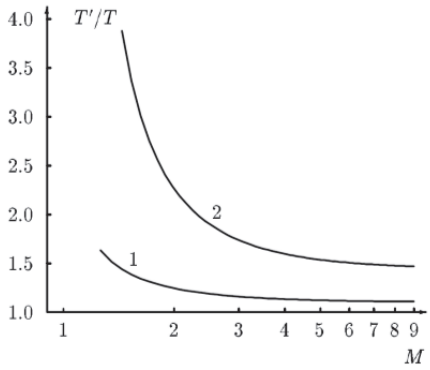


Fig. 11. Dependence of the critical layer overheating $(T/T')_{crit}$ on the shock wave Mach number. γ = 1.4. 1) – Mark-Griffiths criterion, 2) – modified criterion [Lashin&Starikovskii, 1995].

The conditions 1) and 2) means the complete deceleration of gas from the entropy layer with respect to the main shock wave front in the plane shock wave. Mathematical details can be found elsewhere [Lashin&Starikovskii, 1995], while here we give the final expression for the critical overheating of the longitudinal thermal layer, which initiates a self-similar growth of the shock wave front precursor:

$$\left(\frac{T'}{T}\right)_{crit} = \left(\frac{M_s^2}{M_s^2 - 1}\frac{(\gamma + 1)}{2}\right)^2$$

This criterion quantitatively reproduces existing experimental results on the shock-wave – hot layer interaction and allows to perform an analysis of the energy needed for supersonic flows control with the help of hot layers (Figure 11).

Another important question is the minimal possible thickness of the entropy layer to achieve a stable interaction. Minimization of the layer thickness means the energy consumption minimization and is extremely important from the point of view of practical

flow control. Based on the results [Nemchinov et al, 1989] the self-similar solution exists at least up to run-up distance dx = (300 – 400)d with the dynamics of precursor length increase close to l ~ 0.3 dx. These relations show that one can expect the shock wave pattern change even using the entropy layer diameter of about 1% of the object. Figure 12 shows the example of interaction of the entropy layer of different size with a bow shock. For relative layer thickness of 10, 100% and excitation of the entire flow the stagnation pressure decrease relative to unexcited flow was 21, 34 and 29% respectively. It is clear that the decrease of the entropy layer size allows significantly increase the efficiency of flow control.

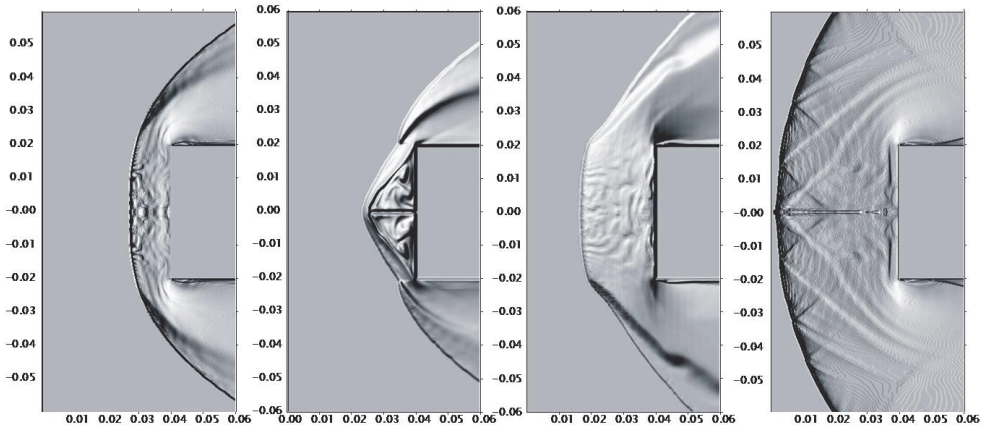


Fig. 12. Numerical model of hot layer – shosk wave interaction. The influence of layer thickness. Air, M = 8.2, T_0 = 300 K, P_0 = 35 Torr. Excitation energy is E = 0.054 eV/mol, layer thickness is a) 0%; b)10%; c) 100% and d) entire flow [Opaits et al, 2004].

Approximately equal effect of thin and thick entropy layers on the shock waves configuration allows to use focused and relatively weak excitation of the gas ahead of the shock wave front. That is why many experiments on drag reduction control by plasma were performed with thin filaments generation [Azarova et al, 2011]. As examples we will mention the pulsed laser energy deposition [Adelgren et al, 2001], MW filaments generation [Kolesnichenko et al, 2004], counter-flow jet injection [Shang, 2002].

Papers [Adelgren et al, 2001, 2001, 2002] examined the effect of pulsed energy deposition on the flow past a sphere at M_1 = 3-4.5. A Nd:YAG laser (532 nm, 10 ns pulse duration, 10 Hz repetition rate) was used in all experiments with energy levels from 12 to 300 mJ. The first configuration is an isolated sphere. A single laser pulse generates a thermal spot upstream of the bow shock. The interaction of the thermal spot with the bow shock is shown in Fig. 13.

In Ref. [Shang, 2002] it was presented a possibility of control of shock wave structure and decrease of the drag due to thermal gas heating by injecting a hot plasma jet toward the main flow. Drag measurements were made with both fluid injection and plasma jet injection. Experiments were conducted at Mach 6. A plasma torch was used to inject a counter-flow at Mach 3.28. Plasma jet injection from a spherically blunted cylinder can significantly reduce drag. Schlieren images of the experimental and calculated flow patterns are shown in Figure 14.

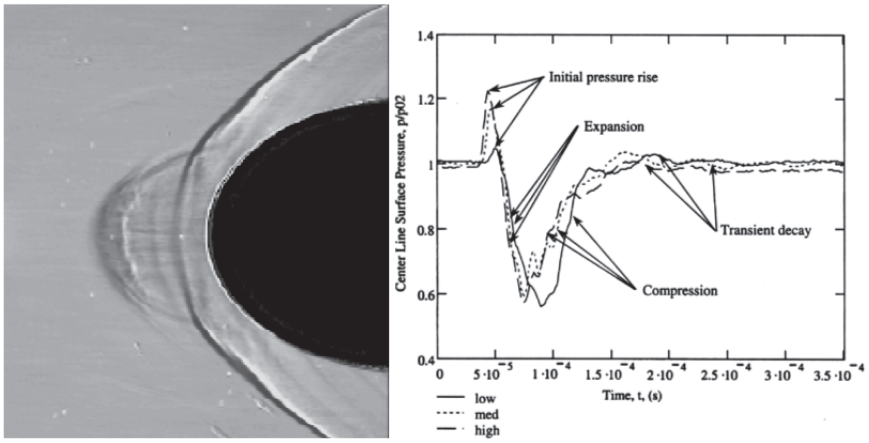


Fig. 13. a) Bow shock interaction with a laser-generated hot spot. b) Temporal behavior of the stagnation pressure on the body centerline for energy pulses energy 13 mJ (low), 127 mJ (med) and 258 mJ (high) [Adelgren et al, 2002].

Figure 13,b confirms very weak dependence of the interaction dynamics for all perturbations above the strong interaction threshold (see equation (1)).

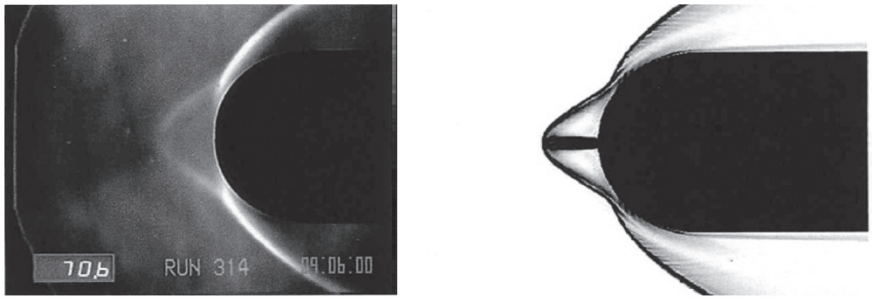


Fig. 14. a) Schlieren picture of counter-flow experiment. b) Calculated flow patterns of counter-flow configuration [Shang, 2002].

Paper [Kolesnichenko et al, 2004] presented the results of controlling the bow shock wave structure with a thin plasma layer formation by applying microwave discharge. The filament, which is formed in 1–2 ns and has a radial dimension of less than 3×10^{-3} cm, has a specific energy of approximately 7 eV per particle with a peak $n_e \approx 5 \times 10^{16}$ cm^{-3}. The temperature rise in the filament is approximately 2800 K with a heating rate of ~2000–3000 K/μs. Time-resolved measurements show that the discharge significantly alters the bow shock configuration due to hot filament formation. Figure 15 demonstrates the bow shock propagation forward into the hot zone, forming the same structure as it is observed for typical hot layer – shock wave interaction (Figure 12). The interaction of the filaments with the bow shock waves after 60 and 85 μs is shown in the Schlieren images (Figure 15, left), and the vortices can be seen as a bright regions in luminosity distribution fields (Figure 15, right).

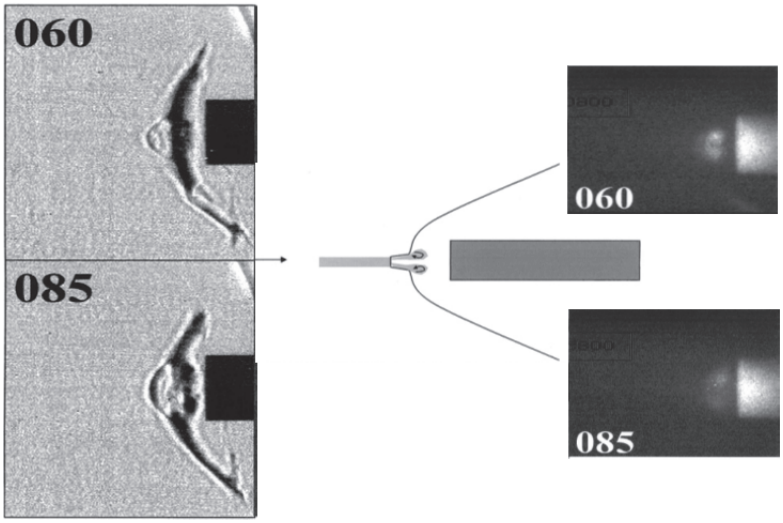


Fig. 15. Interaction of microwave filament with blunt body flow. Left - Schlieren images; right - luminosity distribution fields. 60 and 85 μs after the MW discharge [Kolesnichenko et al, 2004].

In [Kolesnichenko et al, 2004] it was shown that the energy efficiency of the stagnation pressure reduction is directly proportional to *(D/d)* ratio, where *D* and *d* are the model and discharge filament diameter, respectively. The validity of this scaling relationship was proved for model diameters between 8 and 30 mm. These results confirm the conclusion of [Nemchinov et al, 1982; Lashin&Starikovskii, 1995] about interaction efficiency increase and for reducing aerodynamic drag when the hot layer's thickness decreases.

As a summary for this part we can conclude that the experiments demonstrate high efficiency in generating small areas of heated gas and controlling configuration of the shock waves and flow pattern. Plasma effects, such as spatial charges and double-layer formations, do not seem to significantly impact the propagation of strong shock waves.

The main problem of supersonic and hypersonic flow control is the technology of hot filament generation ahead of the leading shock wave. It is necessary to generate a plasma filament with controlled properties and location in the low-density region in front of the shock wave. The problem is clearly demonstrated in recent paper [Erdem et al, 2011]. The paper presents the effect of energy deposition via arc discharge on 15^0 half-angle truncated cone-cylinder configurations and a truncated cone with flare configuration in Mach 5 flow. To initiate the arc discharge the authors should place high voltage tungsten electrode ahead of the model (Figure 16). The arc discharge influence is much weaker than the influence of the electrode in front of model. Besides, the electrode system will be blown away in almost all practical conditions.

Current methods under consideration include DC electric fields, RF fields, microwaves, lasers, electron beams and short pulsed high voltage electric fields [Macheret et al, 2001]. In all cases we have to aviod the energy release behind shock wave front to avoid an additional heat fluxes. A combination of several methods such as the use of jets, laser beams or e-beams to localize MW or pulsed DC energy release may be considered.

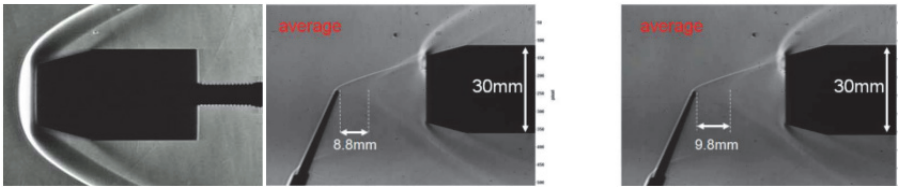


Fig. 16. Comparison of a) base; b) discharge-off and c) b) discharge-on cases in terms of averaged Schlieren images over a truncated cone-cylinder [Erdem et al, 2011].

3. MHD flow control

The history of MHD concept for flow control actually was started even earlier than AJAX concept. A wide range of MHD flow-control opportunities has been suggested: MHD flow control for inlet mass flow and shock position control, MHD flow turning and separation control; MHD power generation; MHD heat transfer reduction at leading edges; MHD flow control for control surfaces [Van Wie et al, 2004].

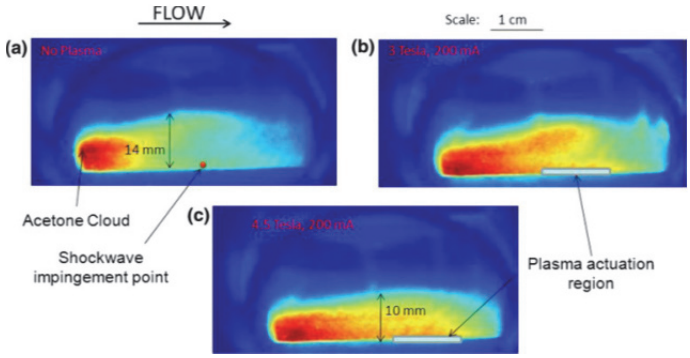


Fig. 17. Time averaged side view of the separation as visualized by acetone PLS with the shock impingement location moved downstream by 2d, (a) base line case with no plasma actuation, (b) plasma actuation at 3 Tesla and (c) plasma actuation at 4.5 Tesla [Karla et al, 2010].

The practical implementation of MHD technologies for aerospace applications faces three major problems: magnetic field generation; gas ionization; plasma recombination. That is why despite of very clear theoretical basis the experimental demonstrations of MHD flow control are very limited.

The paper [Karla et al, 2010] demonstrates the potential for shockwave–turbulent boundary layer interaction control in air using low current DC constricted surface discharges forced by moderate strength magnetic fields. Experiments are conducted in a Mach 2.6 in draft air tunnel with discharge currents up to 300 mA and magnetic field strengths up to 5 Tesla. The effect of plasma control on boundary layer separation due to interaction with shock wave depends on the direction of the Lorentz force. If shock-induced separation is present, it is observed that by using Lorentz force that pushes the discharge downstream, separation can be suppressed (Figure 17). Electron density was maintained by DC grow discharge on the level of $n_e \sim 2.9 \times 10^{12}$ cm^{-3}.

Extensive experimental work on the problem of MHD flow control has been conducted by a team of researchers at the Ioffe Physical Technical Institute led by Bobashev [Lapushkina et al, 2011]. The papers [Erofeev et al, 2010; Lapushkina et al, 2009] discusses the MHD influence on the flow (velocity ~2 km/s, Mach number = 4). Two models in the shape of a cone mated with a cylinder were used. At the cone vertex and along the mating line two electrodes were located, and inside the cylindrical parts magnetic inductors were housed. One of the models was equipped with a magnetic core inside the inductor (Figure 18). An appreciable difference between the influences on the gas flow about these two models was detected. Measurements of the heat flux toward the surfaces of the models were carried out. It was found that the induction of the magnetic field appreciably influence on the local heat load.

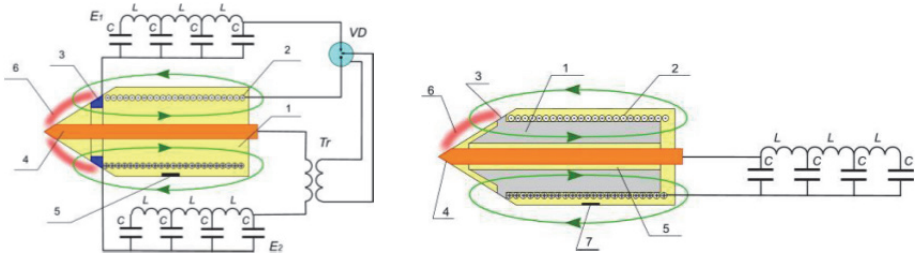


Fig. 18. Model geometry a) without; b) with magnetic core [Lapushkina et al, 2009].

Figure 19,a shows Schlieren patterns of the flow for the model with the core (left) and without core (right). The maximum magnitudes of the current and magnetic induction amounted to I_m = 1.7 kA and B_m = 0.3 T for the model without core, and I_m = 1.5 kA and B_m = 0.7 T for the model with the core, respectively. Figure 22,b shows signals of the heat sensor located on the cylindrical surface of the model containing the magnetic core. In the plot there are signals at various discharge currents and magnetic inductions (I_m =1.3 kA, B_m = 0.6 T and I_m =1.5 kA, B_m = 0.7 T) obtained at the positive and negative polarities of the ring electrode [Lapushkina et al, 2009].

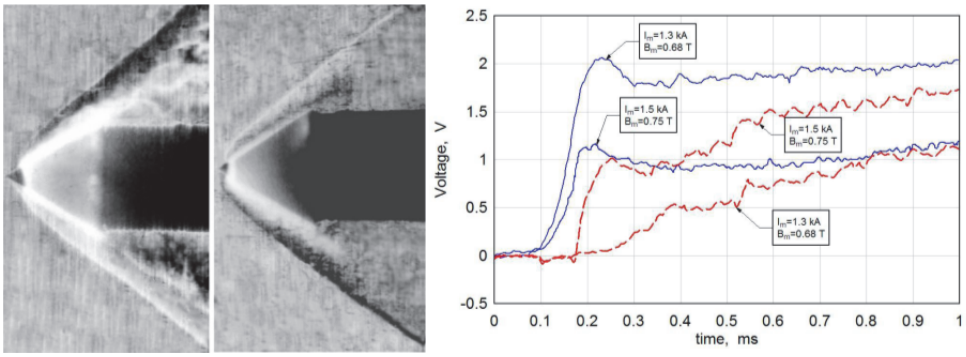


Fig. 19. a) Schlieren patterns of the flow about the model with the magnetic core (left) and without core (right). b) Oscillograms of the signals of the heat flux sensors installed on the model with the core at the positive (dashed lines) and negative (solid lines) polarities of the ring electrode and various discharge currents [Lapushkina et al, 2009].

When increasing the discharge current and magnetic induction the magnitude of the signal of the heat sensor varies differently: when the ring electrode is positive the mean magnitude of the signal increases, and at the negative polarity of the ring electrode the signal decreases (Figure 19,b).
Another series of experiments at the Ioffe Physical Technical Institute have been conducted using a shock tunnel (Figure 20) operating with rare gases (krypton, xenon and argon) to produce an ionized gas flow [Bobashev et al, 2006].

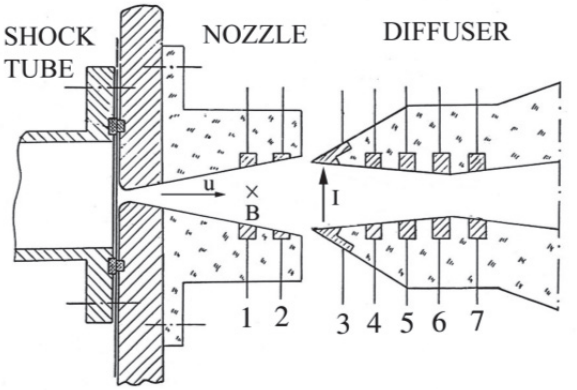


Fig. 20. Scheme of MHD channel with electrodes [Bobashev et al, 2006]. Figures are numbers of electrodes. *U* is flow velocity, B is magnetic field, I is current.

The experiments shown in Figure 21 were carried out in Xe. In this case the magnetic field influence on a change in the Mach number, when flow enters into the diffuser, should be predominated one at B > 0.8 T.

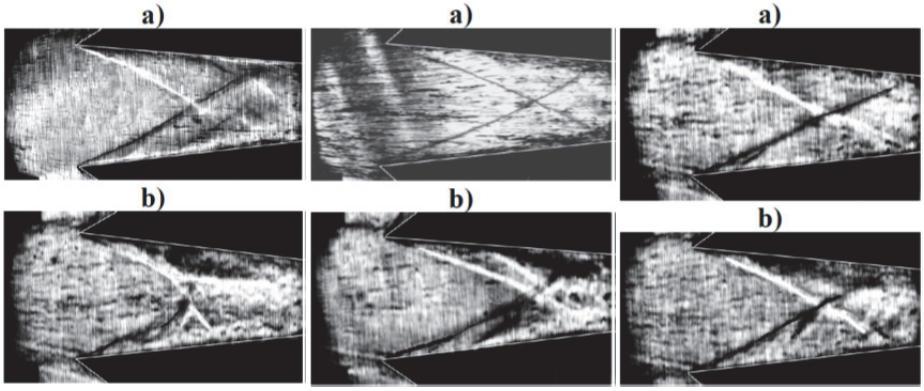


Fig. 21. Schlieren pictures of the flow in the case I, II and III (left to right). (a) V=110 V, B=0; (b) V = 110 V, B=1.3T [Bobashev et al, 2006].

In Figure 21 showed are the distinguished region of the diffuser functioning as the Faraday channel with the sectioned electrodes: I – a whole diffuser, the electrodes from 3rd to 7th pairs

functioning; II – a region of the diffuser, the inlet section excluded, a current goes via 4-7th pairs of the electrodes; III – the inlet section, A current goes only via 3rd pair of the electrodes. All the electrodes are supplied with an equal voltage V = 110V, the experiment was carried out at B = 1.3 T. In Figure 21 showed are the Schlieren images of a flow obtained at the different commutations of a current.

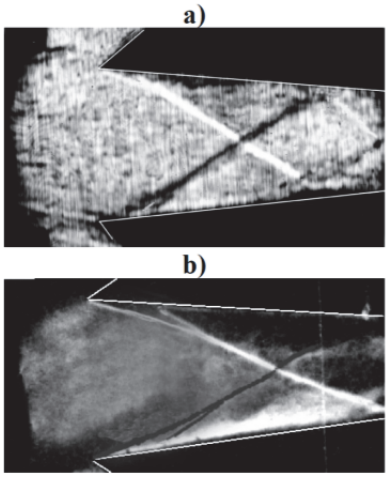


Fig. 22. Examples of variations in shock-wave configurations under the action of electric and magnetic fields. a) deceleration regime; b) acceleration regime
[Bobashev et al, 2006].

Experiments shown in Figure 21 revealed a strong effect of Joule heating [Bobashev et al, 2006]. The aim of experiment demonstrated in Figure 22 was to separate the action of ponderomotive force and Joule heating. In this series of experiments interaction with magnetic and electric fields was localized in a short inlet part of the diffuser, i.e., where the action of the fields is most efficient [Bobashev et al, 2006]. Authors [Bobashev et al, 2006] underlined that the hypersonic MHD experiments should be performed in air flow ionized by the external power sources, but at present air ionization in the diffusers is questioned and require additional investigations.
Below we will illustrate general principles and problems of MHD flow control using an example taken from the review [Van Wie, 2004]. A schematic of MHD inlet flow control system is shown in Figure 23. The concept proposed in [Van Wie et al, 2004] incorporates a large 5-m diameter magnet located in the forward end of the forebody to produce a 3-T field at the surface. A 1D array of e-beam guns is located within the magnet to inject high-energy electrons along the magnetic field lines. The e-beam energy is enough to provide sufficient ionization at a distance of 2.2-m from the surface. Electrodes are located on either side of the e-guns to collect the transverse MHD current. Figure 24 shows predicted flowfield of MHD-controlled M_{DES} = 5 inlet operating at Mach 10 [Schneider et al, 2004]. The temperature contours show that the MHD flow control is successful in repositioning the forebody shocks at the cowl lip. The narrow MHD interaction region is seen in the contours of the electron density.

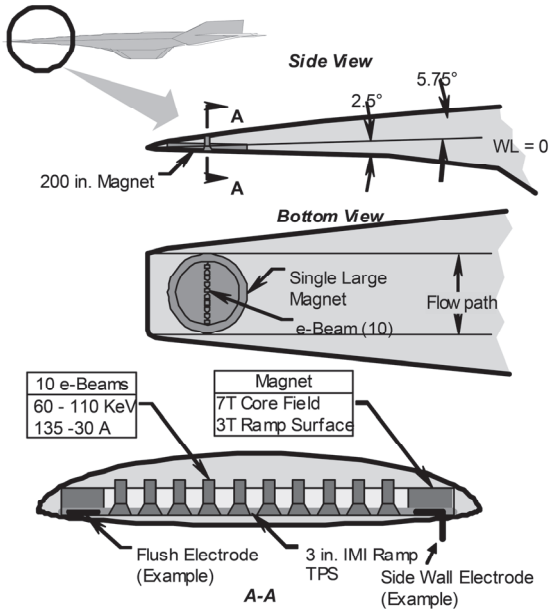


Fig. 23. MHD inlet control system [Van Wie, 2004].

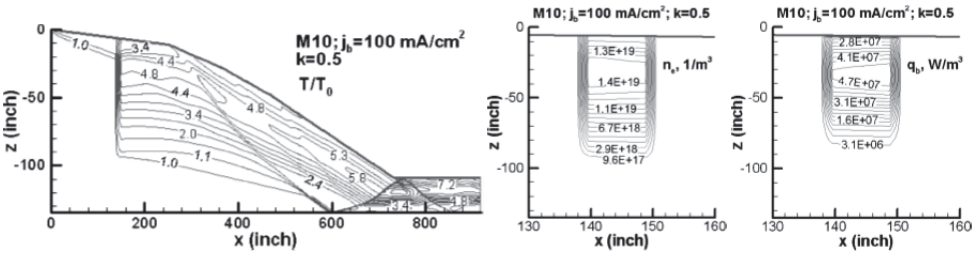


Fig. 24. Predicted flowfield of MHD-controlled M_{DES}=5 inlet operating at Mach 10. a) Temperature field; b) Electron density field; c) Beam power [Schneider et al, 2004].

Estimations of [Schneider et al, 2004] show that the flow control system can operate in a self-sustained mode with the ~76 MW/m power extracted, while a power required for the ionization system is less than 29 MW/m. This extremely important conclusion requires some additional comments. First, to achieve a high efficiency of MHD interaction extremely heavy 3.5-T magnets are proposed; second, the interaction efficiency is limited by the efficiency of gas ionization by e-beams (energy required is ~34 eV per electron-ion pair); and third, the region of interaction is limited by plasma life time – i.e., rate of nonequilibrium plasma recombination. It should be noted that in [Schneider et al, 2004] the only recombination channel, dissociative recombination with simple molecular ions, was taken into account (the rate coefficient $k = 2\times10^{-7}(300/T_e)^{1/2}$, where T_e is the electron temperature). The energy efficiency of gas ionization by high-energy e-beam is well-known. Energy threshold for nitrogen ionization is ~15.6 eV, and similar energy is spent on excitation and

dissociation of the molecules. As a result, the energy cost for electron-ion pair production in air under the action of high-energy electrons is 33-34 eV.

There are several mechanisms of electron loss that lead to a decrease in the conductivity of a nonequilibrium molecular plasma. They are dissociative electron- ion recombination, three-body electron-ion recombination, the third body being a molecule or electron, and electron attachment to molecules. Under the conditions typical for MHD applications, electron density is sufficiently high to neglect electron attachment as compared to electron-ion recombination.

In [Schneider et al, 2004], it was assumed that the dominant mechanism of electron loss is electron recombination with simple positive ions such as O_2^+ and N_2^+. This is not valid in an air plasma at room temperature at which simple ions are usually transformed to complex ions such as O_4^+ and N_4^+. The rates of dissociative recombination for complex ions are an order of magnitude higher than the rates of dissociative recombination for simple ions [Florescu-Mitchell&Mitchell, 2006]. Therefore, the lifetime of the plasma was overestimated in [Schneider et al, 2004] approximately by an order of magnitude. This follows also from direct measurements of the effective recombination rates in room temperature N_2, CO_2 and H_2O under conditions close to those for MHD-controlled inlets were performed in papers ([Zhukov et al, 2006; Aleksandrov et al, 2007a,2007b,2008,2009]), and in air in paper [Aleksandrov et al, 2011].

Discharge was initiated in a quartz tube of inner diameter 47 mm and outer diameter 50 mm, the metallic electrodes being at the ends of the tube. Observations were made for gas pressures between 1 and 10 Torr. Pulses of amplitude 11 kV in cable, duration 25 ns at half-height and rise time 5 ns were supplied to the electrodes (Figure 25). The time-resolved electron density was measured by a microwave interferometer for (f = 9.4 × 10^{10} Hz, a wavelength of 3 mm) initial electron densities in the range 8 × 10^{11} – 10^{12} cm^{-3} and the effective electron–ion recombination coefficient was determined. It was shown that this coefficient varies in time and depends on pressure. A numerical simulation was carried out to describe the temporal evolution of the densities of charged particles under the conditions considered. A good agreement was obtained between the calculated and the measured electron density histories. It was shown that the loss of electrons is governed by dissociative recombination with complex ions, their density being dependent on pressure.

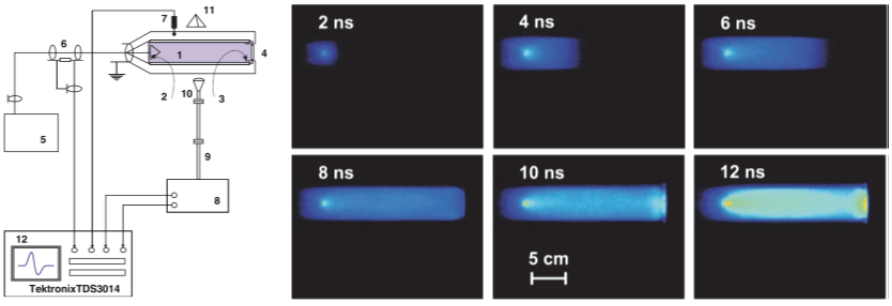


Fig. 25. a) schematic diagram of the experimental setup: (1) quartz discharge tube, (2) high-voltage electrode, (3) low-voltage electrode, (4) end CaF2 window, (5) high-voltage generator, (6) back-current shunt, (7) capacitive gauge, (8) main block of the interferometer, (9) wave guide, (10) horn antenna, (11) reflector and (12) oscillograph; b) ICCD images of nanosecond discharge in air. ICCD gate is equal to 1 ns, time moments from the discharge start are indicated. High voltage electrode is on the left hand side [Aleksandrov et al, 2007a].

The plasma life-time $\tau_{1/2}$ was determined at the beginning of the plasma decay or later, at the instant at which n_e decreases to 2×10^{11} cm^{-3} (Figure 27). In all gases considered, the coefficient a_{eff} varies in time in the afterglow and depends on pressure. Huge effective recombination coefficient a_{eff} (in comparison with dissociative recombination coefficient used in [Schneider 2004]) has been explained by extremely fast formation of complex ions. For example, in nitrogen we have [Aleksandrov et al, 2007a, 2007b]:

$$e + N_2^+ \Rightarrow N + N \qquad k_d \text{ (molecular ion)} = 2.8 \times 10^{-7}(300/T_e)^{1/2}$$

$$e + N_4^+ \Rightarrow N_2 + N_2 \qquad k_d \text{ (cluster ion)} = 2\times10^{-6}(300/T_e)^{1/2}$$

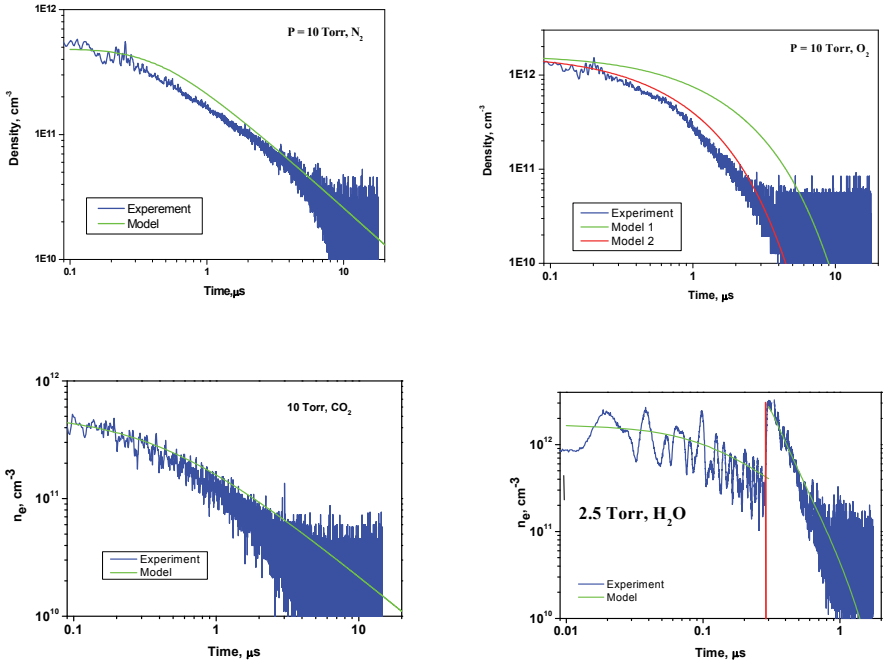


Fig. 26. Dynamics of electron density in plasma afterglow. T = 300 K; a) N_2; b) O_2; c) CO_2; d) H_2O [Aleksandrov et al, 2007a, 2007b].

Figure 26 shows typical electron density histories measured, respectively, in N_2, O_2, CO_2 and H_2O at a discharge repetitive frequency of 2 Hz.

The positive ion composition can be dominated by simple O_2^+ ions in a high-voltage nanosecond discharge in room-temperature air (see calculations in [Aleksandrov et al, 2011]). In this case, O_4^+ ions have no time to form from O_2^+ ions in the discharge phase and in the discharge afterglow. However, measurements [Aleksandrov et al, 2011] showed that in this case the predominance of O_2^+ ions does not necessarily lead to increasing the lifetime of the air plasma. Let us consider this point in more detail.

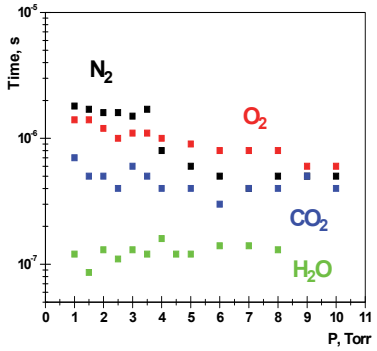


Fig. 27. Effective plasma life time in different gases [Aleksandrov et al, 2007a, 2007b].

Figure 28 compares the evolution in time of the electron density measured in [Aleksandrov et al, 2011] during the discharge afterglow and that of the electron density calculated using the generally accepted rate constants for electron loss [Kossyi et al, 1992]. The difference between the measurements and calculations reached a factor of three, much higher than the experimental error of the electron density measurements that was around 20-30%. The analysis of the kinetic scheme and rate constants used showed that all rate constants were taken from measurements, with the exception of the rate of three-body electron-ion recombination

$$e + O_2^+ + e \rightarrow \text{neutral products} + e.$$

The rate coefficient of this reaction has been measured only at $T_e = T > 1500$ K and only for atomic ions. It was shown in a model calculation [Collins 1965] that the rate of three-body recombination for molecular ions can be an order of magnitude higher than the rate of three-body recombination for atomic ions. The calculations with the rate of this reaction increased according to [Collins 1965] led to good agreement with the measurements (see Figure 28).

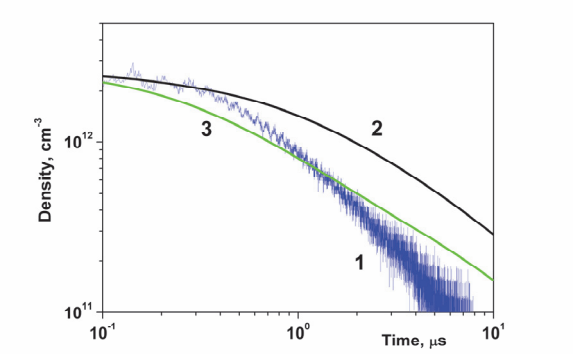


Fig. 28. The evolution in time of the electron density in the nanosecond discharge afterglow in air for 8 Torr [Aleksandrov et al, 2011]. Curve 1 corresponds to measurements. Calculations were carried out (curve 2) with the generally accepted rate constants and (curve 3) when the rate of three-body electron-ion recombination was increased by analogy with [Collins, 1965].

It may be concluded that the lifetime of room-temperature nonequilibrium air plasma could be an order of magnitude shorter than that used in [Schneider et al, 2004] to estimate air plasma conductivity even when the dominant ion species is O_2^+. This means that the power required for the ionization system of MHD inlet actually is 10 times higher than estimations of [Schneider et al, 2004] and close to ~290 MW/m while the power extracted remains the same ~76 MW/m. Power budget of MHD inlet control becomes negative and clearly demonstrates the importance of detailed kinetic mechanisms for analysis of plasma applications.

Plasma lifetime could be lengthened by an increase in the electron temperature. This occurs in the plasma decay at elevated gas temperatures. In paper [Aleksandrov et al, 2008] the results of plasma decay in air and N_2:O_2:CO_2:H_2O mixtures (model mixtures for GTE's outlet) at elevated gas temperatures were presented. Plasma decay after a high-voltage nanosecond discharge has been studied experimentally and numerically behind incident and reflected shock waves in high temperature (600–2400 K) air and N_2:O_2:CO_2 mixtures for pressures between 0.05 and 1.2 atm (Figure 29,a). Time-resolved electron density history was measured by a microwave interferometer for initial electron densities in the range $(1–3)\times10^{12}$ cm^{-3} and the effective electron–ion recombination coefficient was determined.

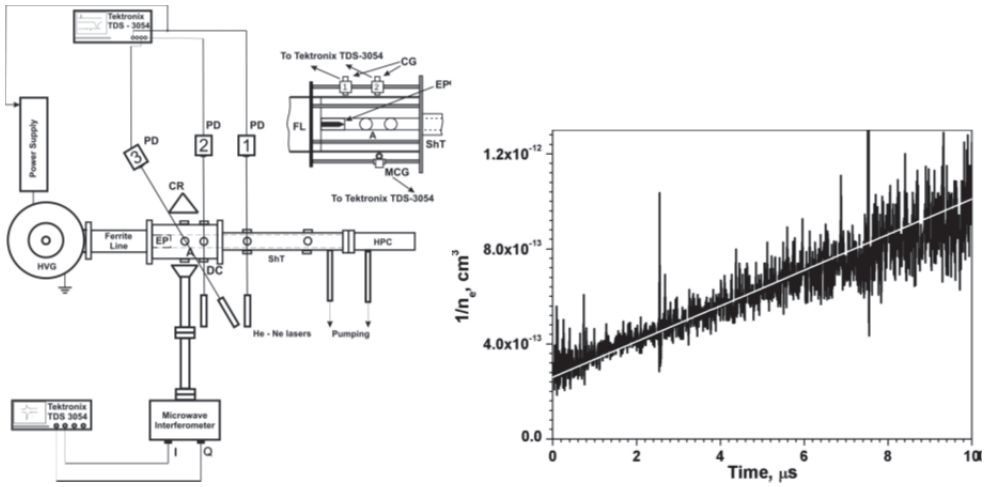


Fig. 29. A) schematic diagram of the experimental setup: (ShT) shock tube; (DC) discharge cell, (A) cross section of measurement, (EP) end plate, (HPC) high pressure cell, (HVG) high voltage generator, (PD) photodiodes, (CR) corner reflector, (CG) capacitance gauge and (MCG) magnetic current gauge. The insert shows the discharge cell on an enlarged scale. B) Typical [$1/n_e$ against time] plot in air at 0.22 atm and 1026 K. The white straight line corresponds to the approximation used to determine the effective recombination coefficient [Aleksandrov et al, 2008].

A numerical simulation was carried out to describe the temporal evolution of the densities of charged and neutral particles. It was shown that the loss of electrons in this case is determined by dissociative recombination with O_2^+ ions, whereas the effect of complex ions and that of three-body recombination are negligible. Electron attachment to O_2 to form negative ions is not important because of fast electron detachment in collisions with O atoms produced in the discharge. In the absence of O atoms the electron density could decay as if the loss of charged particles were governed by electron–ion recombination with the effective rate coefficient being much higher than the dissociative recombination coefficient.
It follows from the measurements [Aleksandrov et al, 2008] in the CO_2-containing mixtures that a_{eff} is independent of gas composition and pressure (in the range 0.05–1.2 atm) and also agrees well with the dissociative recombination coefficient for O_2^+ . It may be concluded that under the conditions studied electron attachment to molecules and dissociative recombination with complex (O_4^+, etc) positive ions are unimportant. The main channel of recombination at elevated temperature conditions is dissociative recombination [Aleksandrov et al, 2008].

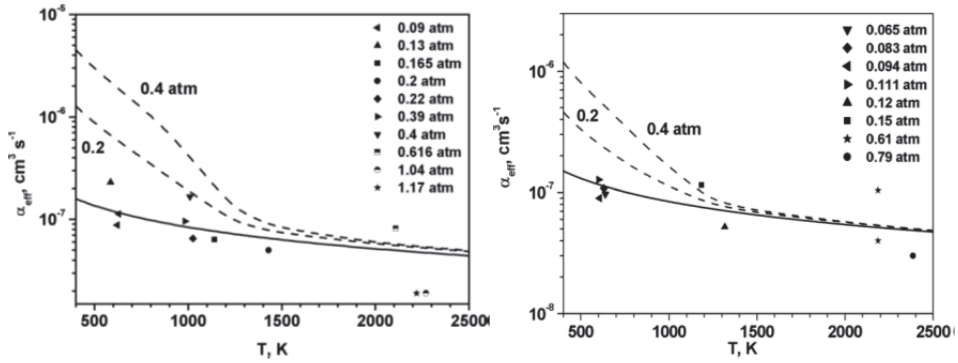


Fig. 30. The effective electron–ion recombination coefficient (symbols) as a function of temperature. The solid curve corresponds to the dissociative recombination coefficient measured in [Cunningham&Hobson, 1972] for O_2^+ and the dashed curves correspond to our calculations at various pressures in the absence of O atoms. A) Air; B) N_2:O_2:CO_2 = 86:5:9 mixture [Aleksandrov et al, 2008].

4. Boundary layer control

On the whole, plasma governs flow through two main mechanisms, either by momentum or energy transfer.
Discharge energy transfer to the flow is a rather complicated multistep process [Raizer, 1991]. Because they possess small masses and long mean free paths, the electrons gain energy from the electric field. The slow rate of energy exchange of electrons with neutral gas results in a significant deviation of the mean electron energy from the energy of translational degrees of freedom of molecules. Depending on the value of the applied electric field, the mean electron energy in the discharge can reach several electron-volts. These conditions provide active excitation of the internal degrees of freedom of molecules, as well as their dissociation and ionization by electron impact. At the same time, the energy flux into translational and fast-thermalizing rotational degrees of freedom is relatively low.

Consequently, the energy release at VT-relaxation, recombination of neutral and charged components and quenching of electronically excited molecules is the main mechanism of gas temperature increase in non- equilibrium plasma. VT relaxation and recombination are rather slow and can last tens of microseconds or longer even at atmospheric pressure, which is comparable with the typical gas dynamic times within a scale of several millimeters. Energy release into translational degrees of freedom, during excitation of electronically excited states and molecular dissociation and ionization by electron impact, is a much faster process. For instance, a molecule being excited by electron impact to a repulsive state dissociates to products with high translational energy. The time of thermalization of such "hot" atoms and radicals usually reaches units of nanoseconds. Quenching of electronically excited molecules and electron-ion and ion-ion recombination proceed almost at the same time scale and also lead to “hot” atoms and radicals formation. Such a heating mechanism can become a governing process and produce fast gas heating in the discharge region under high values of reduced electric field E/n (close to or higher than the breakdown threshold) [Popov 2001, Aleksandrov et al, 2010a,2010b].

Presently, most researchers applying plasma actuator for flow control propose to use this device to accelerate the flow in the boundary layer near the airfoil surface in the region of flow separation. They consider induced velocity to be one of the main features developed by the actuator in the discharge zone. The gas flow velocity can be changed during the interaction between the electric field and uncompensated spatial plasma charge.

The flow acceleration mechanism is connected with loss of quasi-neutrality in the plasma which conducts electric current. In the case of a small Debye radius, the existence of the electric field feeding the current is always connected with the existence of considerable uncompensated spatial charge in plasma (in the absence of the media polarization $\mathrm{div}(\varepsilon_0 E) = 4\pi\sigma$)). Gaining the momentum from the electric field, uncompensated charge causes whole gas motion [Sigmond&Lagstadt, 1993]. For instance, this pattern is typical for glow discharge.

At low ionization degree and high electron energy, the Debye radius is noticeably bigger than the typical size of the plasma region; and then, the electric field is determined only by external conditions, which leads to charged particles acceleration in the external field. The total gas acceleration is determined by the space charge of the plasma region. This charge is formed by the discharge current from the electrodes. A low-current corona discharge from the point-like electrode may be an example of such a situation.

Both gas acceleration in the boundary layer and pulse heating with further expansion may, on the whole, lead to changes of flow characteristics. It is necessary to analyze the value of gas acceleration by discharge as well as gas heating and induced flow in the discharge afterglow in order to investigate the physics of interaction between the nanosecond pulsed discharge and gas flow.

Two different mechanisms, stationary and non-stationary, lead to such interaction. In a stationary case the electrical field is limited by breakdown threshold. In the paper [Likhansky et al, 2010] the estimations based on the volumetric force equation $F = e \times n \times E$ and the Poisson equation lead to simple relation for induced velocity

$$v_g = E^*(\varepsilon_i/\rho)^{1/2}$$

where ρ is the gas density, E is an applied electric field and ε_i is the ion mobility. This equation describes the gas flow in stationary discharges using the condition that E cannot

exceed the breakdown threshold. For free space this equation predicts the maximum induced velocity up to 80 m/s, but close to the surface due to the viscous effects this maximum cannot be achieved [Likhansky et al, 2010] and actual limit was estimated ~20 m/s. Actually, the estimation proposed in [Likhansky et al, 2010] assumes the permanent presence of a spatial charge in the plasma region. In a weak electrical field under consideration this charge cannot be generated by gas ionization or emission from the electrodes [Raizer, 1991]. Thus the estimation [Likhansky et al, 2010] is an upper estimation of the induced velocity in the presence of external source of uncompensated charge in plasma region.

As a rule, the presence of high uncompensated spatial charges in gas is associated with the presence of strong electric field gradients and ionization waves [Starikovskaia et al, 2002]. A streamer discharge is an example of such a case. Uncompensated charge on the ionization wave front at the streamer is under the influence of the strong electric field of the streamer's head. This results in significant acceleration of the gas in the region of the strong field. This process lasts only fractions of nanoseconds. The calculations presented in [Opaits et al, 2005] have shown that the gas velocity in a single streamer's channel may reach units of centimeters per second. This mechanism is implemented in pulsed non-stationary discharges without bias.

AC discharges and pulsed discharges with significant bias situated in between of these two limiting cases. Presently, the possibility of gas acceleration reaching a velocity up to nearly 10 m/s has been shown with the help of positive corona [Loiseau et al, 2002; Zouzou et al, 2006; Rickard et al, 2006].

It should be noted that the nature of gas acceleration is the same in all cases. The interaction between the uncompensated plasma charge and the electric field, together with the effective momentum transfer from charged to neutral gas components, generate flux acceleration as a whole.

4.1 Laminar-turbulent transition control

In [Grundmann&Tropea, 2007] artificially excited Tollmien–Schlichting (TS) waves were cancelled using plasma actuators operated in pulsed mode. In order to achieve this a vibrating surface driven by an electromagnetic turbulator was flush mounted in a flat plate to excite the TS waves. These were amplified by an adverse pressure gradient induced by an insert on the upper wall of the test section. A control plasma actuator positioned downstream of the excitation actuator attenuates the waves by imparting an unsteady force into the boundary layer to counteract the oscillation. As a result the amplitude of the velocity fluctuations at the excitation frequency is reduced significantly depending on the distance from the wall. A parameter study was performed to identify the influence of several operation parameters of the control actuator.

The investigations have been performed in an open circuit wind tunnel with a test section of a cross section of 0.45 m by 0.45 m and a length of 2 m. An insert on the roof of the test section creates an adverse pressure gradient of 25 pa/m to promote transition on the flat plate at the relatively low velocity of 9.6 m/s measured in the smallest cross section. The boundary-layer thickness has a value of d_{99} = 5 mm at x = 590 mm yielding a Reynolds number of Re = 1100 based on the displacement thickness [Grundmann&Tropea, 2007]. Figure 31a shows the test section and Fig. 31b shows a closeup view of the two actuators and the measurement position.

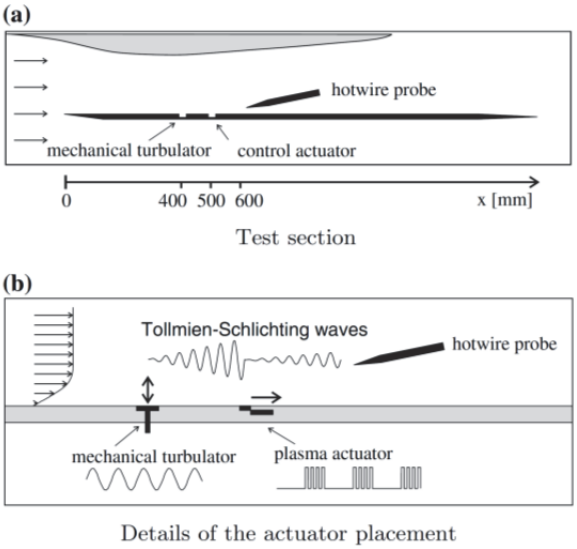


Fig. 31. Test section and detail view (a) Test section (b) Details of the actuator placement [Grundmann&Tropea, 2007].

Figure 32 gives more detailed information about the frequency content and the shape of the fluctuations with and without control. The figures on the left show the power spectra densities of the velocity fluctuations and the figures on the right show the time traces of these measurements. With the control actuator working, the amplitude (bottom of Fig. 32a) of the fundamental frequency is reduced significantly, while the modes f_2 and f_3 remain unchanged. The mode f_4 is cancelled while f_5 disappears below the background noise floor produced by the actuator.

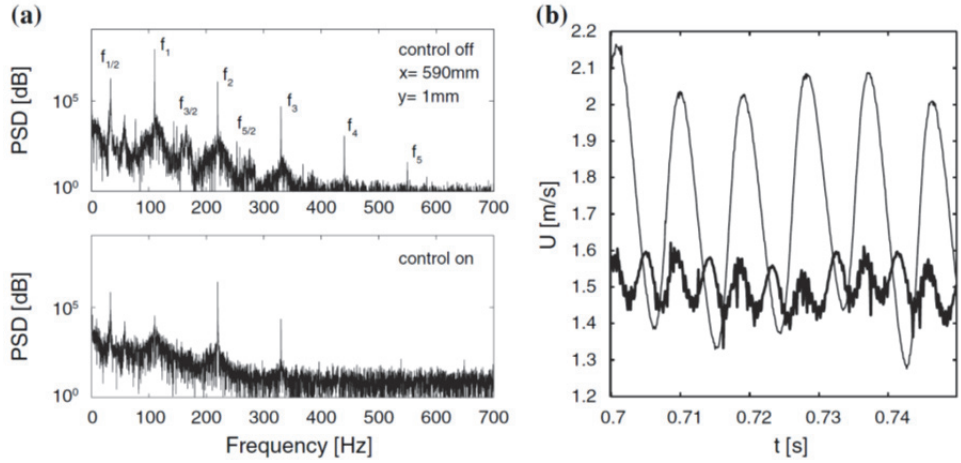


Fig. 32. Power spectra density and time traces with (thick lines) and without (thin lines) cancellation at x = 590 mm. y=1 mm [Grundmann&Tropea, 2007].

4.2 Boundary layer separation control by ionic wind

Unlike cases involving strong shock waves, a great number of papers on slow subsonic flow control point out the role of plasma effects (and ion wind in particular) in accelerating gas in the boundary layer, controlling the layer detachment and guiding the laminar-turbulent transition [Moreau, 2007].

Any surface-proximal plasma layer employed to change the flow regime can be easily generated by various techniques. For example, papers [Velkoff& Ketchman 1968; Yabe et al, 1978] and more recent publications [Leger et al, 2001a,2001b] used a direct current discharge with electrodes placed above or on the surface of the airfoil to achieve the effect. A discharge-generated ion wind can provide flow acceleration up to 3-10 m/s in the boundary layer [Moreau et al, 2005; Richard et al, 2006].

Prof. Roth and his team [Roth et al, 1998a,1998b,2000] presented another approach to generate the plasma layer near the surface to control flow. This approach is based on creating surface DBD by applying AC sinusoidal voltage. Discharge is developing in the form of thin streamers propagating along the surface above the covered low electrode [Allegraud et al, 2007].

This type of plasma actuator and its modifications have been widely investigated recently [Moreau, 2007]. Paper [Gregory et al, 2007] demonstrates the value of thrust force generated by an asymmetric actuator at the level of 0.2 mN/W. Practically the same value (0.3 mN/W) was obtained in [Abe et al, 2007]. The flow velocity generated by such an actuator may reach values up to 5 m/s according to the measurements presented in [Roth et al, 2006]. Meanwhile, paper [Forte et al, 2006] presented values of induced velocities up to 8 m/s. Such flow acceleration provides effective control of the velocity profile in the boundary layer as well as its detachment for main flow velocities reaching the value of several dozen meters per second.

For example, in paper [Do et al, 2007], the flow speed ranges from 10 m/s to 25 m/s and the corresponding *Re* numbers are from $5 \cdot 10^4$ to $5 \cdot 10^5$. In this flow regime, the separation point behind the bluff body can be moved downstream in the presence of the AC DBD. However, the separation delay effect is found to decrease as the flow speed increases. Paper [Lopera et al, 2007] described wind tunnel experiments conducted on a 47^0-sweep, scaled 1303 UAV model for flight control at low angles of attack. The actuators produced significant shifts in the lift curve, up to 25% for the most effective ramp angles of 20 and 30 degrees, in the 0-20 degree alpha range for a free-stream velocity of 15 m/s. For all ramp cases examined, the unsteady (pulsed) actuator was more effective than the steady actuator in controlling flow separation and influencing the aerodynamic lift.

In the study Post et. al. [Post et al, 2007], the effectiveness of a plasma actuator was tested on a high-speed, natural laminar flow, HSNLF(1)-0213 airfoil. The 10-kV peak-to-peak actuator is designed to simulate an aileron-up or trailing-edge flap upward deflection at M=0.1 (Re=292 K) and M=0.2 (Re=584 K). The tests are performed at various angles of attack from $\alpha = -2^0$ to 16^0. The results at M=0.2 indicate a 2% increase in C_L and up to an 8% increase in C_D.

Thus, the plasma actuators based on AC sinusoidal voltage surface dielectric barrier discharges make it possible to change the flow velocity within several meters per second (maximum induced velocity has been reported by Corke [Corke, 2011] V ~ 12 m/s) and manage the boundary layer detaching at the main flow velocities up to ~40 m/s. There are no published data on the influence of ionic wind flow acceleration for free stream velocities above 60 m/s. This result confirms the conclusion of very first paper by Mhitaryan [Mhitaryan et al, 1964] where the authors made a conclusion that the actuator affects the flow through ionic wind mechanism when induced velocity was in the order of 20-25% from the velocity of free stream.

A primary goal of the study [Thomas et al, 2009] is the improvement of actuator authority for flow control applications at higher Reynolds numbers. The study examines the effects of dielectric material and thickness, applied voltage amplitude and frequency, voltage waveform, exposed electrode geometry, covered electrode width, and multiple actuator arrays. The metric used to evaluate the performance of the actuator in each case is the measured actuator-induced thrust which is proportional to the total body force. It is demonstrated that actuators constructed with thick dielectric material of low dielectric constant produce a body force that is an order of magnitude larger than that obtained by the Kapton-based actuators used in many previous plasma flow control studies. They achieve jet velocity 5-6 m/s at the distance ~4-5 cm downstream of the actuator (Figure 33).

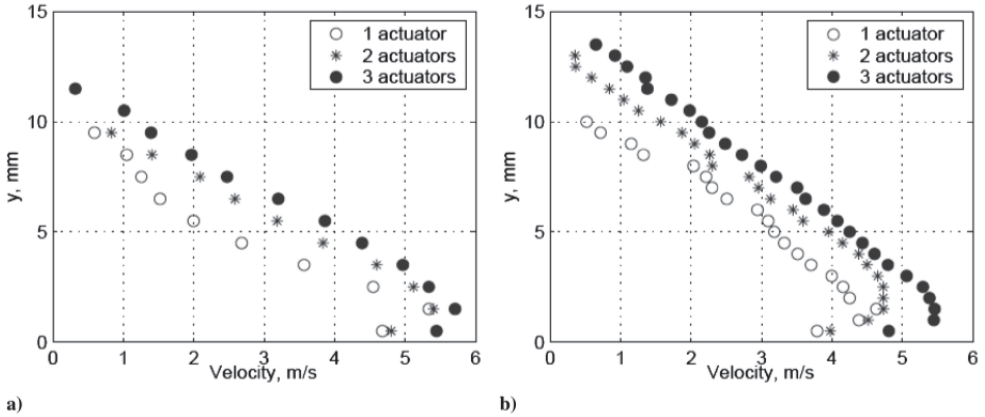


Fig. 33. Mean velocity profiles for single, dual, and triple actuator configurations: a) 3.81 cm downstream; b) 5.08 cm downstream [Thomas et al, 2009].

Combined analysis of the capacitance, light emission, size of the plasma region, force production and power consumption is presented in [Kriegseis et al, 2011]. A force-power diagram in presented in Figure 34. Such a plot led to the dimensioned coefficient of the force production efficiency

$$\eta = \frac{F}{P_a} = \frac{F/_L}{P_a/_L}$$

Measurements [Kriegseis et al, 2011] show that for thrust generation by AC plasma actuator η = 2.5×10^{-4} N/W. The same parameter calculated for Pratt & Whitney F100 Engine gives η = 1.1×10^{-3} N/W (calculated from total fuel energy). Thus even assuming no losses for electric power generation, plasma actuator is about order of magnitude less efficient than GTE. The main advantage of plasma actuators is their flexibility and fast response.
It seems that the physical restrictions employed in the mechanism of creating "an ion wind" do not allow significant improvement in performance of this technology because of physical limitations for flow acceleration in the discharge. At the same time, subsonic aerodynamics researchers are very interested in the velocity range from 100 m/s (take-off and landing velocities) to 250 m/s (cruising speed). Thus, advancing into the region of higher velocities is of great importance and urgency.

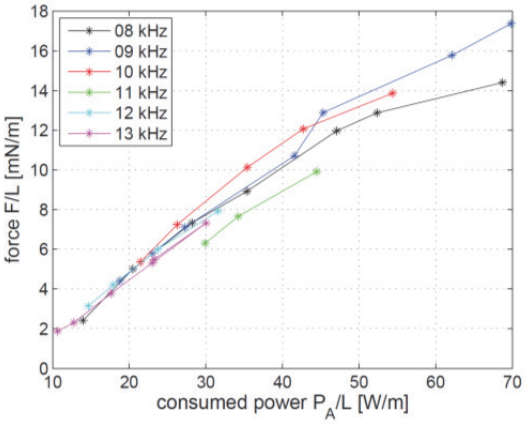


Fig. 34. Dimensioned coefficient of force production efficiency for AC plasma actuator [Kriegseis et al, 2011].

4.3 Boundary layer separation control by heat release

Paper [Opaits et al, 2005] proposed using pulsed nanosecond discharge for plasma actuator. The *E/n* value for this type of the discharge can exceed by several times the breakdown threshold. The high value of the reduced electric field seems to be an evident advantage of such a discharge. Such characteristics as relatively low energy consumption, the possibility of using such discharges within a wide range of pressures, flow velocities, and gas compositions, including high humidity, also contribute to the advantages of the approach proposed. The first experiments [Opaits et al, 2005] have shown that it is possible to firmly control the boundary layer separation using this nanosecond pulsed discharge at velocities up to 75 m/s and energy consumption lower than 1 W/cm of wingspan.

Further, the impact of pulsed sliding discharge on the flow separation has been investigated in [Roupassov et al, 2006]. The high efficiency of pulsed discharge was shown for the velocity up to 110 m/s. The main mechanism of plasma influence was concluded to be the boundary layer turbulization, rather than the gas acceleration. An optimum pulsed actuator frequency was found to maximize the actuator effect on lift and drag force and flow re-attachment, such as $f_{opt} = U_0/L$, where U_0 is the main flow velocity and L is the typical distance along the surface to the separation zone. Later, this result was confirmed by Patel et al. [Patel et al, 2007] in experiments for chord Reynolds numbers up to 10^6 and a maximum free-stream speed of 60 m/s.

Scaling effects of an aerodynamic nanosecond pulsed plasma actuator were investigated in [Sidorenko et al, 2007; Maslov et al, 2007]. Separation control experiments on a rectangular wing (dimensions 0.5×1 m²) were carried out using a dielectric barrier discharge plasma at subsonic speed for chord Reynolds numbers from 0.35 to $0.875 \cdot 10^6$. Surface pressure measurements and flow visualization show that global flow separation on the wing can be mitigated or eliminated by the plasma actuators (Figure 35). The data were obtained for a wide range of attack angles, flow speeds, plasma excitation frequencies and power. New applications of several kinds of voltage pulses for plasma excitation were discussed, including microsecond and nanosecond pulses. As in [Roupassov et al, 2006], it was found there that control efficiency strongly depends on discharge frequency (Figure 36).

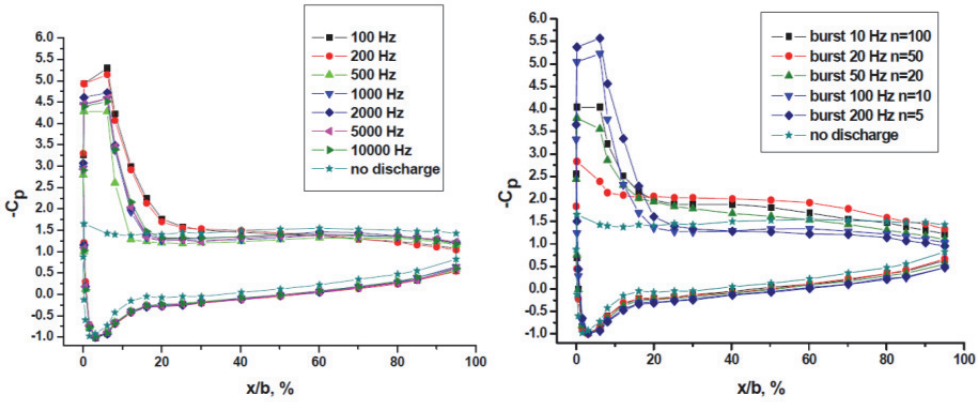


Fig. 35. C_p distribution along the model chord ($\alpha = 19^0$; $U^\infty = 19$ m/s; $V = 24$ kV; $\mathrm{Re} = 0.8\times10^6$) [Sidorenko et al, 2007]

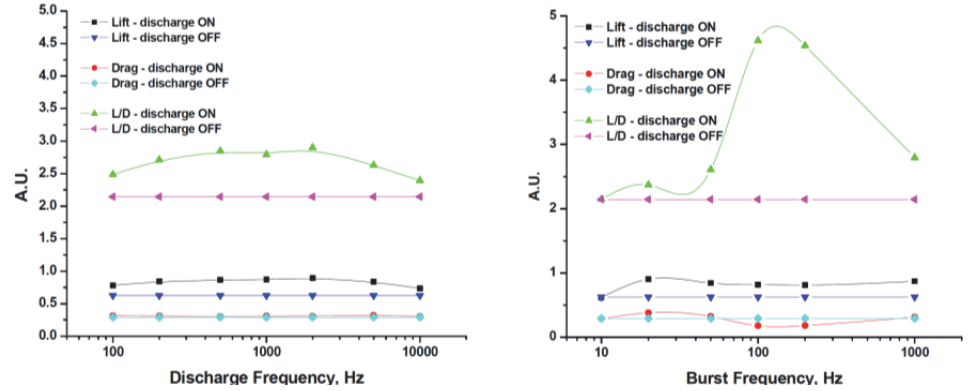


Fig. 36. Lift, Drag force and Lift-to-Drag ratio in dependence on the frequency. $\alpha = 22^0$; $U^\infty = 17.4$ m/s; a) – Periodic Mode, $P = 2.5$-250 W for $f = 100 - 10000$ Hz, respectively; b) – Burst Mode, $P = 25$ W for all regimes [Sidorenko et al, 2007].

Separation control experiments on a rectangular wing were carried out using nanosecond dielectric barrier discharge plasma at subsonic speed (M = 0.3 - 0.75) for chord Reynolds numbers between 0.5 and $2\cdot10^6$ [Roupassov et al, 2007]. This work has demonstrated the possibility to control the flow at cruising velocity with a plasma actuator. A vacuum blow-down wind tunnel has been used for the experiments. The system was modernized to perform the experiments in pulse regime. The nozzle with working chamber operates at Mach numbers from M = 0.6 to M = 0.9. Figure 37 depicts the installation.

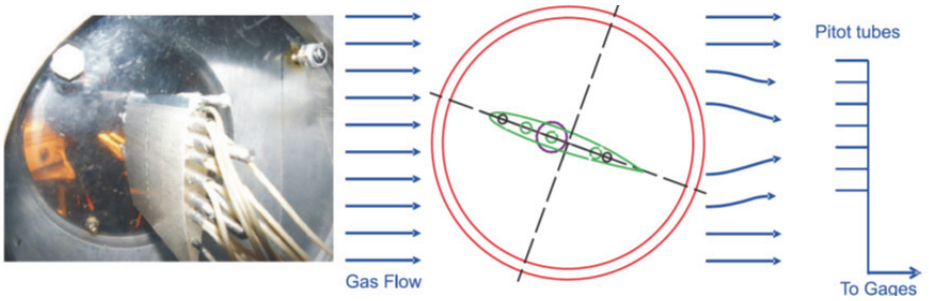


Fig. 37. Photo of the model and schematics of pressure measurements. Pressure distribution have been measured in the wake of model and on the model surface [Roupassov et al, 2007].

The discharge impact on the flow pattern near the surface has been investigated. The Mach number was equal to M = 0.65 − 0.6; 0.7 − 0.65; or 0.74 − 0.69 in different experiments. The discharge frequency in the experiments was equal to 5 kHz. High-voltage pulses have amplitude of 25 kV, pulse width was 12 ns. Discharge energy was equal to 10 mJ/pulse. The plasma impact was investigated for angles of attack between 0 and 30^0.

An unseparated flow regime with local supersonic zone and shock wave formation was observed for small angles of attack. These regimes were clearly identified by the pressure jump in the middle of the airfoil surface. This jump is associated with the shock wave location (Fig. 38). The discharge impact for angles within the range of 0 − 15^0 is negligible. For higher angles of attack, the flow separation is observed and the pattern of pressure distribution changes (Fig. 38). For angles of attack higher than the stall angle, the discharge switches the flow to the unseparated flow regime. Figure 38,a presents the pressure distribution on the upper surface of the model. The X-value corresponds to the distance from the leading edge of the model to the pressure port. The discharge was able to remove high-frequency pulsations in the wake of the model. The data from the pressure gauges for Mach number M = 0.7 are presented in Fig. 38,b to illustrate the noise reduction. Gauge N1 records the pressure at the upper surface of the model and shows the change in the attack angle. Gauges N2-4 are placed in the wake of the model. Pressure pulsations in the wake disappear when the discharge is switched on. This effect was observed at high angles of attack (starting with $\alpha = 24^0$) for Mach number M = 0.65−0.75. The mean pressure value near the model surface does not change significantly, while high-frequency pulsation amplitude decreases dramatically. Thus, the study of separation control for the model of C-141 airfoil has been carried out at transonic velocities (M = 0.65 − 0.75). Dielectric barrier discharge plasma was used for separation control. The effects of the angle of attack and flow Mach number on the efficiency of flow control were studied in experiments. Nonequilibrium plasma impact was observed for angles of attack from 18^0 to 30^0.

The discharge removes both flow separation and high-frequency pulsations in the wake. These experiments demonstrate a possibility of transonic flow separation control using low-energy pulsed nanosecond surface dielectric discharges.

Thus, nanosecond pulsed discharges have demonstrated an extremely high efficiency of operation for aerodynamic plasma actuators over a very wide velocity (M = 0.03 - 0.75) and Reynolds number (Re = 10^4 - 2×10^6) range. For further technological development, it is extremely important to understand the physics of the nanosecond plasma actuator and differences between different types of SDBD in terms of their efficiencies [Roupassov et al, 2008a,2008b,2009; Nikipelov et al, 2009; Correale et al, 2011; Rios et al, 2011].

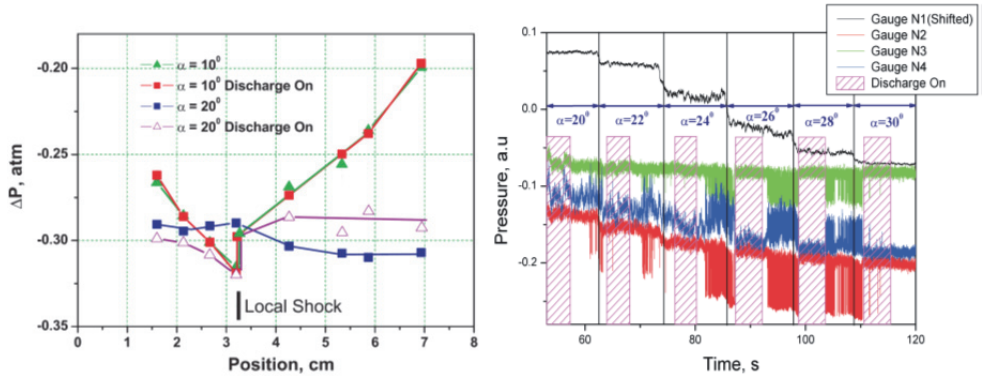


Fig. 38. a) Pressure distribution on model surface with and without discharge. Mach number M = 0.74. Total pressure P = 1 atm. b) Noise reduction in the wake of the model. Mach number M = 0.7. Total pressure P = 1 atm
[Roupassov et al, 2007].

From this point of view there are several important milestones. Paper [Roupassov et al, 2006] experimentally demonstrated that the pulsed nanosecond high-voltage discharge used for boundary layer separation control in a wide range of free stream velocity produces no gas acceleration. In [Visbal&Gaitonde, 2006] the use of a steady counter-flow DBD actuator as a boundary-layer tripping device was numerically analyzed. According to calculations, the actuator induced transition and turbulence, and generates a fuller velocity profile. This feature was exploited to delay stall of a NACA 0015 airfoil at high angle of attack using a pulsed counter-flow actuator. Thus, [Visbal&Gaitonde, 2006] demonstrated that the co-flow gas acceleration is not necessary for boundary layer control. In [Roupassov et al, 2008a] the mechanism of pulsed nanosecond high-voltage discharge influence on boundary layer separation was experimentally demonstrated. It was shown that fast nonequilibrium plasma thermalization (on the time scale of hundreds of nanoseconds) produces hot, over-pressurized gas layer in the discharge zone, followed by strong shock wave formation. It was suggested that the shock wave propagation across the boundary layer causes strong flow perturbations and provokes flow re-attachment through formation of large scale vortices in the shear layer separating free stream and separation bubble [Roupassov et al, 2008a]. Later, experimental results [Samimy et al, 2010] prove that nanosecond SDBD plasma performs as an active trip at pre-stall angles of attack and provides high amplitude perturbations that manipulate flow instabilities and generate coherent spanwise vortices at post-stall angles. These coherent structures entrain freestream momentum thereby reattaching the normally separated flow to the suction surface of the airfoil. Numerical modeling of SDBD development also shows fast formation of plasma layer and shock wave generation [Unfer&Boeuf, 2009; Starikovskii et al, 2009].
The process of nanosecond pulsed plasma layer interaction with the flow, formation of perturbations and vortices, and flow re-attachment was investigated in details in [Correale et al, 2011].
A model of NACA 63-618 airfoil with the chord of 20 cm and span of 40 cm with the actuator applied was used for experiments. Several different actuators were used, including single, double and triple ones. The flow speed was 30 m/s. Some results are shown in Figure 39.

The shock wave generated by actuators can be clearly seen, as well as large scale vortex structure as it developed 40 microseconds after the discharge [Correale et al, 2011]. It was observed that after 2-3 discharges the flow pattern changed completely. Flow reattached, separation zone shifted downstream. It was found that placing second actuator into the point to where separation was shifted by the first actuator, shifts the separation further downstream. This allows to achieve attached flow up to AoA = 32^0, using three pairs of the actuators. Summary energy consumption was less than 1 W for 40×20 cm airfoil in 30 m/s flow.

Thus typical system reaction time was 10-15 ms and was close to the time of the vortex propagation along the surface of the airfoil (Figure 39). From Figure 39 it is clear that perturbation generated by pulsed actuator initiates instability in the shear layer. This instability propagates along the shear layer; additional mixing brings additional momentum into boundary layer from the main stream and attaches the flow. It should be noted that the discharge energy plays a secondary role: two different regimes (repetitive pulse mode and burst mode) shown in the columns 1 and 2, correspondingly, demonstrate almost the same dynamics of flow attachment while the discharge energy in the second case is 10 times bigger. This means that we need high rate of energy release from discharge to translational degrees of freedom of gas. Fast transition (in time scale shorter than gas-dynamic time in plasma layer) means the efficient generation of the shock wave and efficient excitation of perturbations in the flow [Starikovskiy et al, 2009]. That is why the kinetics of energy transfer in nonequilibrium plasma is the most critical issue for pulsed SDBD actuators.

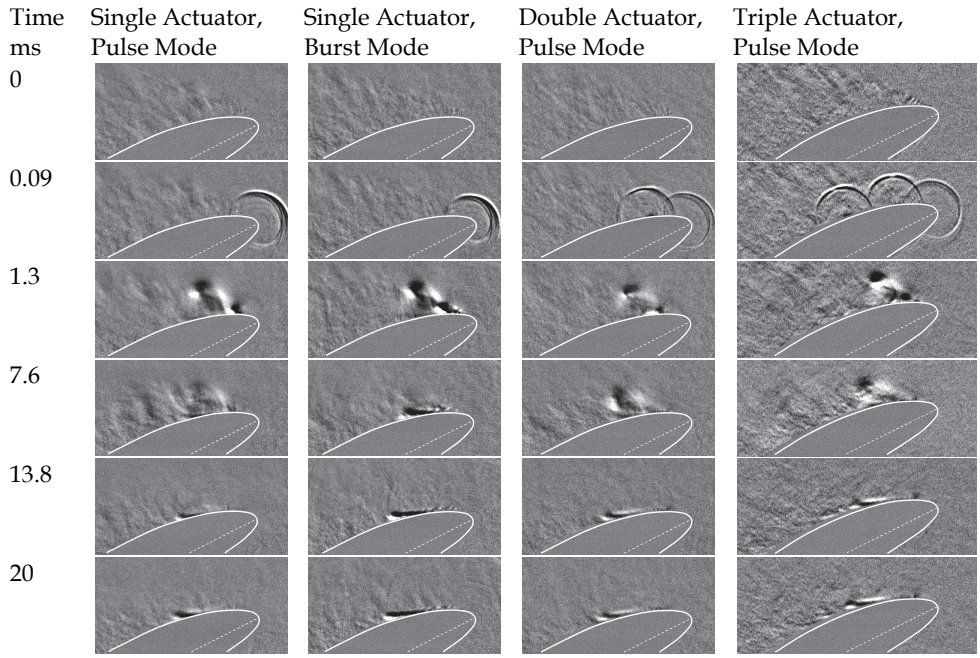


Fig. 39. Dynamics of boundary layer re-attachment. V = 30 m/s, AoA = 26^0, NACA 63-618 airfoil, chord length was 20 cm wing span was 40 cm. Discharge energy 5 mJ/pulse, discharge frequency 200 Hz in pulse and burst modes; 10 pulses per burst; $d\tau$ = 100 μs [Correale et al, 2011].

As it was indicated above, the main mechanism of pulsed nanosecond SDBD effect on the flow is an extremely fast gas heating. Energy release in the gas is sometimes considered to be $Q=U\times I\times\tau$, whereas gas heating is defined by $\Delta T = Q/C_p$. Such an estimate includes some strong assumptions. The electric field energy is supposed to be completely absorbed by gas. This is not always true in the case of strong electric fields, since part of the energy is lost in radiation processes. In the case of high-current discharges at low electric fields, some energy will be lost in the near-electrode regions. In this case, part of the energy goes to heat the electrodes. Thus, the current multiplied by voltage in the discharge gap gives only the upper estimation of energy release. Estimations of temperature changes in the discharge are still-stronger suppositions. The equation $\Delta T = Q/C_p$ is completely valid for the thermal equilibrium state when internal degrees of freedom of the gas are in equilibrium with the translational degrees of freedom. That is not the case under conditions of strongly nonequilibrium plasma of gas discharge. On the other hand, using specific heat under constant pressure C_p presumes that energy release occurs at times noticeably higher than gasdynamics times. Then, it is quite reasonable to use the supposition $P = const$.

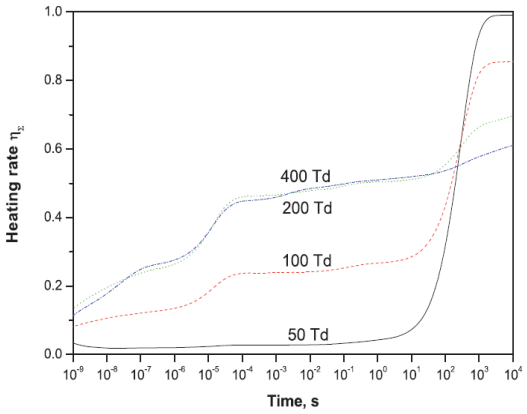


Fig. 40. Percentage of nonequilibrium energy transferred into translational degrees of freedom [Flitti&Pancheshnyi, 2009].

So, when analyzing the thermal mechanism of plasma actuator impact on the flow, it is necessary to take into account not only radiation energy loss, wall heating, etc., but also the rate of energy relaxation as compared to the typical times of plasma layer expansion.
Dynamics of plasma relaxation in the case of excitation by low and moderate electrical fields was calculated many times (see, for example, [Flitti&Pancheshnyi, 2009], Figure 40).
Mechanisms of fast gas heating under low electrical fields (E/N < 20 Td) mainly include elastic electrons scattering and rotational excitation of the molecules. Here, typical relaxation time is rather short because of fast energy exchange between rotational and translational degrees of freedom, but total energy fraction of this excitation is very small (Figure 41). According to this Figure, under moderate electrical fields (E/N = 20 - 200 Td) there is efficient excitation of vibrational and electronic degrees of freedom. VT relaxation under low temperature conditions is very slow process and curve "50 Td" in Figure 40 demonstrates that almost all the energy will be frozen for about 100 sec before real gas heating will take place. Under such conditions formation of a shock wave (strong perturbations) is impossible. Instead, weak

compression waves will appear. Under higher E/N (100-200 Td) efficient excitation of electronic degrees of freedom and molecules dissociation will take place (Figure 41.) Dissociation by e-impact takes place through repulsive states and 20-30% energy goes immediately to translational motion of fragments (for example, $e + O_2 \rightarrow e + 2O + \Delta E$). Collisional quenching of electronically excited states (in air there are nitrogen triplets – $N_2(A, B, C, a', ...)$) also lead to energy release into translational degrees of freedom:

$$e + N_2 \rightarrow e + N_2^*(A, B, C, a', ...)$$

$$N_2^*(A, B, C, a', ...) + O_2 \rightarrow N_2 + 2O + \Delta E$$

$$O(^1D) + N_2 \rightarrow O + N_2 + \Delta E$$

This mechanism was proposed for air in [Popov, 2001].
In SDBDs reduced electrical field reaches extremely high value (E/n ~ 800-1200 Td). Significant part of the electrons energy goes to gas ionization. Extension of the energy relaxation mechanism to high E/n was proposed in [Aleksandrov et al, 2010]. We have analyzed the results of two observations of nonequilibrium plasma produced by high-voltage nanosecond discharges. These results involved the measurement of the velocity of a shock wave that propagates through air heated by an impulse discharge at 20 Torr and the experimental study of a SDBD in atmospheric-pressure air. The electron power transferred into heat in air plasmas was estimated in high (~10^3 Td) electric fields. It is shown that around 50% of the discharge power can be transferred into heat for a short period of time (~ 1 μs at atmospheric pressure). This effect is much more profound than that observed at low and moderate reduced electric fields.

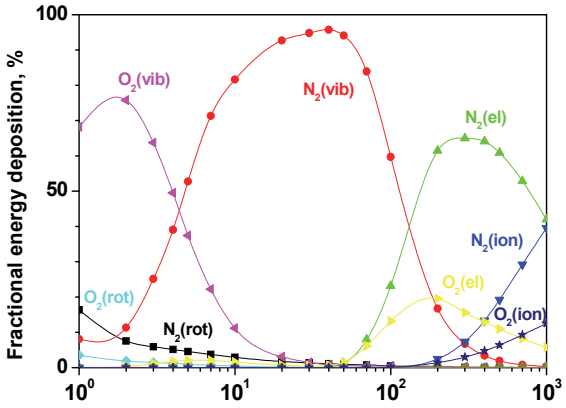


Fig. 41. Discharge energy distribution across internal degrees of freedom in air [Aleksandrov et al, 1982].

A kinetic model was suggested to simulate the fast heating of air plasmas under the conditions considered. This model extends work previously developed for describing fast heating in moderate (<10^2 Td) reduced electric fields and takes into account electron-impact excitation of high-energy states followed by their collisional quenching, as well as ion–molecule reactions and electron–ion and ion–ion recombinations. These reactions play an important role in

plasmas produced at high electric fields when most electron energy losses are due to electron-impact ionization. Based on this model, the fractional electron power transferred into heat was calculated as a function of the reduced electric field in dry and humid air at various pressures. Calculations agree well with the results of experimental analysis of SDBD at atmospheric pressure. There is also reasonable agreement between theory and measurements in the impulse high-voltage nanosecond discharge initiated in air at 20 Torr.

According to the calculation at 20 Torr, approximately equal parts (<10%) of the electron power are converted into heat: (i) through electron impact dissociation of O_2 and excitation of $N_2(A,B,C,a)$ states followed by quenching by O_2, which are suggested to describe fast gas heating in air plasmas at moderate electric fields; (ii) through electron-impact excitation of higher electronic N_2 states followed by dissociation and quenching by O_2 and (iii) through electron–ion recombination. At atmospheric pressure, the calculated total fractional electron power transferred into heat could be increased by <20% due to the three-body recombination of positive and negative ions and, to a smaller extent, due to ion–molecule reaction in the discharge afterglow. The calculations shows that, under the conditions considered, the characteristic time of gas heating lies in the range 0.3–5 ns for 1 atm and in the range 5–80 ns for a pressure of 20 Torr (Figure 42).

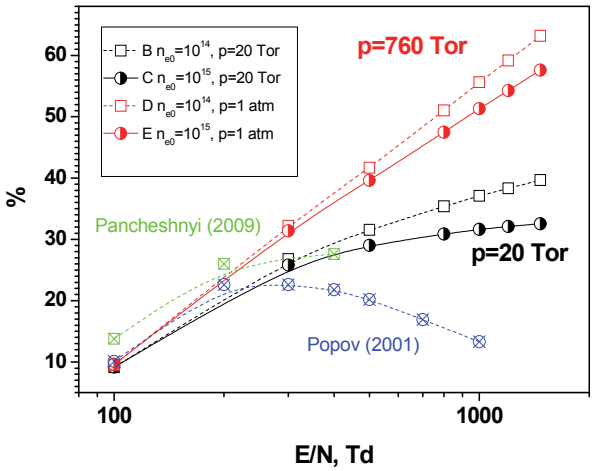


Fig. 42. The total fractional electron power transferred into heat in dry air at 20 Torr and 1 atm as a function of the reduced electric field at which the energy was deposited in a high-voltage nanosecond discharge [Aleksandrov et al, 2010]. The calculations were carried out for $n_{ef} = 10^{15}$ cm^{-3} (solid curves) and 10^{14} cm^{-3} (dash curves). Curve 1 corresponds to calculations [Flitti&Pancheshnyi, 2009] and curve 2 corresponds to the calculations assuming that 28% of the energy spent on the excitation of electronic N_2 and O_2 states is quickly transferred into gas heating [Popov, 2001].

It should be mentioned that both AC and pulsed discharges always provide a combined excitation. We can generate the flow perturbation with AC BDB (see, for example, [Visbal& Gaitonde, 2006]) and we can accelerate the gas using pulsed DBD – especially with additional bias [Opaits et al, 2010]. Using one or another discharge type allows to optimize the process and to minimize the energy consumption.

5. Conclusions

Flow control opportunities by plasma include shock wave pattern control; aerodynamic breaking; drag reduction; heat mitigation; flow vectorization, acceleration and deceleration; MHD power extraction and breaking. Boundary layer control could be subdivided into laminar-turbulent transition control; boundary layer separation control; lift and drag force control; acoustic noise control; mixing enhancement. Nonequilibrium plasma also may be very efficient in ignition and flame stabilization control; engine performance enhancement, including possibility of fast initiation of detonation waves.

This review mentions briefly the most important results obtained over the last decade in plasma assisted aerodynamics and discusses the physical mechanisms of the phenomena under consideration. There are three different physical mechanisms which control the efficiency of plasma aerodynamics: 1) gas heating; 2) electrostatic momentum transfer to the gas; 3) magneto-hydrodynamic effects, including MHD flow acceleration and on-board electricity generation using gas flow kinetic energy. It is shown that the most universal mechanism of plasma action on airflows is their local heating. This mechanism is responsible for supersonic flow and shock wave control, can play an important role in MHD flow interaction and is central to boundary layer control by pulsed nanosecond SDBD. It has been demonstrated that the pulsed nanosecond SDBD is promising for boundary layer control at take-off, landing and cruising flow speeds. The modification of boundary layer by ionic wind is important when using discharges of longer duration (for instance, with sinusoidal high-voltage power supply). However, the last achievements in this area are more moderate.

It was shown that the plasma recombination and energy release in the recombination process control the efficiency of plasma-assisted flow control. In the case of "plasma" mechanisms (electrostatic momentum transfer; magneto-hydrodynamic effects) fast plasma recombination and thermalization limits the possibilities of flow control and sometimes make their usage impossible. Vice versa, for methods based on the gas heating plasma recombination is a major source of energy and the fast heat release is the most important factor which increases the efficiency of plasma control.

Recent advances in plasma kinetics allow to build detailed kinetic models to predict the efficiency of different plasma mechanisms in different aerodynamic applications, but most of the progress in nonequilibrium plasma aerodynamics has been made experimentally. Advances in theoretical simulation of the interaction between non-equilibrium plasma and high-speed airflows have been less promising. The main reason is that we have to simulate simultaneously complicated hydrodynamic, electrodynamics and kinetic processes on wide space and time scales. In addition, there is the shortage in information about the mechanisms, rates and products of plasma processes under nonequilibrium conditions. Further study of the mechanisms responsible for plasma – gas flow interaction on various time scales would favour the progress in this area.

6. Acknowledgements

The work was partially supported by Russian Foundation for Basic Research under the project "Nonequilibrium plasma thermalization", AFOSR under the project "Fundamental Mechanisms, Predictive Modeling, and Novel Aerospace Applications of Plasma Assisted Combustion", and DOE Combustion Energy Frontiers Research Center.

7. References

Abe,T., Takizawa,Y., Sato,S., Kimura,N. "A Parametric Experimental Study for Momentum Transfer by Plasma Actuator". 45th AIAA Aerospace Sciences Meeting and Exhibit. 8 - 11 January 2007, Reno, Nevada. AIAA 2007-187

Adelgren, R., Elliott, G., Knight, D., Zheltovodov, A., and Beutner, T., "Energy Deposition in Supersonic Flows," AIAA Paper 2001-0885, 2001.

Adelgren, R., Elliott, G., Knight, D., Zheltovodov, A., and Buetner, T., "Localized Flow Control in Supersonic Flows by Pulsed Laser Energy Deposition," The Third Workshop on Magneto- and Plasma Aerodynamics for Aerospace Applications, Inst. for High Temperatures, Moscow, April 2001.

Adelgren, R., "Localized Flow Control with Energy Deposition," Ph.D. Thesis, Dept. of Mechanical and Aerospace Engineering, Rutgers Univ., Sept. 2002.

Aleksandrov N.L., and Konchakov V. Plasma Physics Reports. 1982

Aleksandrov N.L., Kirpichnikov A.A., Kindusheva S.V., Kosarev I.N., Starikovskii A.Yu. Non-Equilibrium Plasma Life Time Measurements and Flow Control. 45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 Jan 2007 Grand Sierra Resort Hotel Reno, Nevada. Paper AIAA-2007-0997

Aleksandrov N.L., Kindusheva S.V., Kirpichnikov A.A., Kosarev I.N., Starikovskaia S.M. and Starikovskii A.Yu. Plasma decay in N_2, CO_2 and H_2O excited by high-voltage nanosecond discharge. J. Phys. D: Appl. Phys. 2007 40 4493

Aleksandrov N.L., Kindusheva S.V., Kosarev I.N. and Starikovskii A.Yu. Plasma decay in air and N_2 : O_2 : CO_2 mixtures at elevated gas temperatures. J. Phys. D: Appl. Phys. 2008. 41 No 21. 215207.

Aleksandrov N., S.Kindusheva, I.Kosarev, A.Starikovskii, Plasma Decay in Air and N_2:O_2:CO_2 Mixtures at Elevated Gas Temperatures. AIAA-2009-1048. 47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition, Orlando, Florida, Jan. 5-8, 2009

Aleksandrov N.L., Kindusheva S.V., Nudnova M.M. and Starikovskiy A.Yu. Mechanism of ultra-fast heating in a nonequilibrium weakly-ionized air discharge plasma in high electric fields. J. Phys. D: Appl. Phys. 43 (2010) 255201 (19pp)

Aleksandrov N., Kindusheva S., Nudnova M., Starikovskiy A. Rate of Plasma Thermalization of Pulsed Nanosecond Surface Dielectric Barrier Discharge. 48th AIAA Aerospace Sciences Meeting. Orlando, Florida. Jan 2010. Paper AIAA-2010-0465

Aleksandrov N.L., Kindusheva S.V., Kirpichnikov A.A., Kosarev I.N., Starikovskaia S.M. and Starikovskii A. Yu. Plasma decay in air excited by high-voltage nanosecond discharge. 10th Intern.Workshop on Magneto-Plasma Aerodynamics. March 22-24, 2011. Moscow, Russia

Allegraud,K., Guaitella,O. and Rousseau,A. "Spatio-temporal breakdown in surface DBDs: evidence of collective effect". 2007. J. Phys. D: Appl. Phys. 40 7698-7706

Anokhin E.M., Starikovskaia S.M., Starikovskii A.Yu. Energy transfer in hypersonic plasma flow and flow structure control by low temperature nonequilibrium plasma. 42nd AIAA Aerospace Sciences Meeting and Exhibit 5 - 8 January 2004, Reno, Nevada AIAA 2004-674.

Azarova O., Knight D., Kolesnichenko Y. Characterization of Flowfield Types Initiated by Interaction of Microwave Filament with Supersonic Body 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida AIAA 2011-1026

Bletzinger,P., Ganguly,B.N., VanWie,D., and Garscadden,A. "Plasmas in high speed aerodynamics". J. Phys. D: Appl. Phys. 38 (2005) R33-R57

Bobashev S. V., Vasil'eva R. V., Erofeev A. V., Lapushkina T. A., Ponyaev, D. M. Van Wie. "Relaxation of the shock-wave configuration in a diffuser after termination of the action of magnetic and electric fields", *Technical Physics Letters*, V.32, N2, 2006, p. 106-109.

Collins C.B. Collisional-dissociative recombination of electrons with molecular ions. Phys. Rev. 140 (1965) A1850-575

Corke TC, Enloe CL, and Wilkinson SP. 2010 Plasma actuators for flow control. Annual Review of Fluid Mechanics 42: 505-529, 2010. 11

Corke T.C. Dielectric Barrier Discharge Plasma Actuators. Lecture Series Notes for Von Karman Institute Lectures. 2011.

Correale G., Popov I.B., Rakitin A.E., Starikovskii A.Yu., Hulshoff S.J., Veldhuis L.L.M. Flow Separation Control on Airfoil with Pulsed Nanosecond Discharge Actuator. 49th AIAA Aerospace Sciences Meeting. Orlando, Florida. Jan 2011. Paper AIAA-2011-1079

Cunningham and Hobson. (1972)

Do,H., Kim,W., Mungal,M.G., Cappelli,M.A. "Bluff Body Flow Separation Control using Surface Dielectric Barrier Discharges". 45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 January 2007, Reno, Nevada AIAA 2007-939

Erdem E., Yang L., Kontis K. Drag Reduction Studies by Steady Energy Deposition at Mach 5. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. 4 - 7 January 2011, Orlando, Florida. AIAA 2011-1027.

Erofeev, Lapushkina T., Poniaev S., and Bobashev S. "Supersonic Body Streamline at Different Configuration Gas Discharge", AIAA-2010-1382, 48th AIAA Aerospace Sciences Meeting and Exhibit and 12th Weakly Ionized Gas Workshop, Orlando, Florida, Jan.4-7, 2010.

Flitti O. and Pancheshnyi S., Eur. Phys. J. Appl. Phys. 45, 21001 (2009)

Florescu-Mitchell A.I. and Mitchell J.B.A. Dissociative recombination. Physics Reports. 430. (2006) 277-374

Forte,M., Jolibois,J., Moreau,E., Touchard,G. and Cazalens,M. "Optimization of a dielectric barrier discharge actuator by stationary and instationary measurements of the induced flow velocity, application to airflow control".2006. AIAA Meeting (San Francisco, USA, June 2006) paper 2006-2863

Gregory,J.W., Enloe,C.L., Font,G.I., and McLaughlin,T.E. "Force Production Mechanisms of a Dielectric-Barrier Discharge Plasma Actuator". 45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 January 2007, Reno, Nevada. AIAA 2007-185

Griffiths W. J.Aeron.Sci. 1956. V.23. N 1. P.19.

Grundmann S and Tropea C. 2007 Experimental transition delay using glow-discharge plasma actuators. Experimental Fluids 41: 653-657, 2007.

Gurijanov EP and Harsha PT. 1996 AJAX: new directions in hypersonic technology. AIAA-1996-4606.

Kalra CS, Zaidi SH, Miles RB, and Machret SO. 2010 Shockwave-turbulent boundary layer interaction control using magnetically-driven surface discharges. Experiments in Fluids DOI 10.1007/s00325-010-0898-9, 2010.

Khorunzhenko V.I., Roupassov D.V., Starikovskii A.Yu. Hypersonic Flow and Shock Wave Structure Control by Low Temperature Nonequilibrium Plasma of Gas Discharge. 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibition, 2002. Indianapolis, AIAA 2002-2497.

Khorunzhenko V.I., Roupassov D.V., Starikovskaia S.M., Starikovskii A.Yu. Hypersonic Shock Wave — Low Temperature Nonequilibrium Plasma Interaction. 39th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, 2003. AIAA Paper AIAA 2003-5048.

Klimov,A.I., Koblov,A.N., Mishin,G.I., Serov,Yu.L., and Yavor,I.P. "Shock wave propagation in a glow discharge" 1982 Sov. Tech. Phys. Lett. 8

Knight D. 2008 Survey of aerodynamic drag reduction at high speed by energy deposition. Journal of Propulsion and Power 24: 1153-1167, 2008. 12

Kolesnichenko,Yu.F., Brovkin,V.G., Khmara,D.V., Lashkov,V.A., Mashek,I.Ch. and Ryvkin,M.I. "Fine Structure of MW Discharge: Evolution Scenario". 2004 4th Int. Workshop on Thermochemical and Plasma Processes in Aerodynamics (July 2004)

Kossyi I.A., A.Yu.Kostinsky, A.A.Matveyev and V.P.Silakov. Kinetic scheme of the non-equilibrium discharge in nitrogen–oxygen mixtures. Plasma Source Sci. Technol., 1 (1992) 207–220

Kriegseis J., Moller B., Grundmann S., and Tropea C. Light Emission, Discharge Capacitance and Thrust Production of DBD Plasma Actuators 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida AIAA 2011-155

T.A.Lapushkina, A.V.Erofeev, S.A.Ponyaev, S.V.Bobashev, "Supersonic flow of a nonequilibrium gas-discharge plasma around a body", *Technical Physics*, V.54, No 6, 2009, P. 840-848.

T.A.Lapushkina, A.V.Erofeev, S.A.Poniaev, S.V.Bobashev. Air supersonic flow control by energy deposition and mhd action near body front part. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida AIAA 2011-1025.

Lashin A.M., Starikovskiy A.Yu. Shock and detonation waves interaction with entropy layers. Technical Physics Technical Physics. Springer. 1995. V 65. N.8. C.11-24.

Leger,L., Moreau,E. and Touchard,G. "Control of low velocity airflow along a flat plate with a DC electrical discharge". 2001.Proc. IEEE-IAS World Conf. on Industrial Applications of Electrical Energy (Chicago, USA, 30 September - 4 October)

Leger,L., Moreau,E., Artana,G. and Touchard,G. "Influence of a DC corona discharge on the airflow along an inclined flat plate". 2001.J. Electrostat. 50-51. 300-6

Likhanskii, A.V., Shneider M.N., Opaits D.F., Miles R.B., Macheret S.O. Limitations of the DBD effects on the external flow. 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2010, Orlando, Florida AIAA 2010-470.

Loiseau,J.F., Batina,J., Noel,F., and Peyrous,R. "Hydrodynamical simulation of the electric wind generated by successive streamers in a point-to-plane reactor". 2002. J. Phys.D: Appl. Phys. 35 1020-31

Lopera,J., and Ng,T.T., Corke,T.C. "Aerodynamic Control of 1303 UAV Using Windward Surface Plasma Actuators on a Separation Ramp". 45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 January 2007, Reno, Nevada, AIAA 2007-636

Macheret SO, Shneider MN, and Miles RB. 2001a Potential performance of supersonic MHD power generators. AIAA-2001-0795.

Macheret,S.O., Ionikh,Y.Z., Chernysheva,N.V., Yalin,A.P., Martinelli,L., and Miles,R.B. "Shock Wave Propagation and Dispersion in Glow Discharge Plasmas," Physics of Fluids, Vol. 13, No. 9, September 2001, pp. 2693-2705.

Mark H. J.Aeron.Shi. 1957. V.24. N 4. P. 304.

Maslov A.A., Sidorenko A.A., Zanin B.Yu., Postnikov B.V., Budovsky A.D., Starikovskii A.Yu., Roupassov D.V., Zavialov I.N., Malmuth N.D. Pulsed Discharge Actuators for Rectangular Wing Separation Control 45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 Jan 2007 Grand Sierra Resort Hotel Reno, Nevada. Paper AIAA-2007-0941

Meyer,R., Palm,P., Ploenjes,E., Rich,J.W. and Adamovich,I.V. "Nonequilibrium Radio Frequency Discharge Plasma Effect on a Conical Shock Wave: M=2.5 Flow", AIAA Journal, vol. 41, No. 5, 2003, pp. 465-469

Mhitaryan A.M., Labinov S.D., Fridland V.Ya. Electro-hydrodynamic method of boundary layer control // in paper collection "Some problems of aerodynamics and electro-hydrodynamics", vol.1, Kiev's Institute of Civil Aviation Engineers, Kiev, 1964.

Mhitaryan A.M., Maksimov V.S., Flidland V.Ya., Labinov S.D. Technology of boundary layer investigation with a new type work panel. IFZh. Volume 4, 1961.

Miles RB, Macheret SO, Martinelli L, Murray R, Shneider M, Ionikh Y.Z, Kline J, and Fox J. 2001 Plasma control of shock waves in aerodynamics and sonic boom mitigation. AIAA-2001-3062.

Miles R.B., Plasma flow control, fundamentals, modeling, and applications. Lecture Series Notes for Von Karman Institute Lectures. 2011.

Moreau,E. "Airflow control by non-thermal plasma actuators". J. Phys. D: Appl. Phys. 40 (2007) 605-636

Nemchinov I.V., Artemiev V.I., Bergelson V.I. New gasdynamic structures development by gas density change in thing longitudinal channels ahead of shock wave front. Mathematical Modeling. 1989. V. 1. N 8. PP. 1–11.

Nikipelov A., Nudnova M., Roupassov D., Starikovskiy A. Acoustic Noise and Flow Separation Control by Plasma Actuator. AIAA-2009-695. 47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition, Orlando, Florida, Jan. 5-8, 2009

Opaits,D.F., Roupassov,D.V., Starikovskaia,S.M., Starikovskii,A.Yu. "Shock Wave Interaction With Nonequilibrium Plasma of Gas Discharge". 42nd AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 2004.

Opaits D.F., Roupassov D.V., Starikovskaia S.M., Starikovskii A.Yu., I.N.Zavialov, S.G.Saddoughi Plasma Control of Boundary Layer Using Low-Temperature Non-equilibrium Plasma of Gas Discharge. 43-rd AIAA Aerospace Sciences Meeting and Exhibit, 2005. Reno, Nevada, USA, paper AIAA 2005-1180, 2005.

Patel,M.P., Ng,T.T., Vasudevan,S., Corke,T.C., Post,M.L., McLaughlin,T.E., Suchomel,C.F. "Scaling Effects of an Aerodynamic Plasma Actuator". 45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 January 2007, Reno, Nevada AIAA 2007-635

Popov N.A. Plasma Physics Reports. 27, 866 (2001).

Post,M.L., Greenwade,S.L., and Yan,M.H., Corke,T.C., Patel,M.P. "Effects of an Aerodynamic Plasma Actuator on an HSNLF Airfoil". 45th AIAA Aerospace Sciences Meeting and Exhibit. 8 - 11 January 2007, Reno, Nevada. AIAA 2007-638

Raizer,Yu.P. "Gas Discharge Physics". Berlin, Heidelberg, New York. Springer-Verlag. 1991.

Rickard,M., Dunn-Rankin,D., Weinberg,F., and Carleton,F. "Maximizing ion-driven gas flows". 2006. J. Electrostat. 64 368-76

Rios M., Nudnova M., Kindusheva S., Aleksahdrov N., Starikovskiy A. Fast Nonequilibrium Plasma Thermalization in N_2-O_2 Mixtures for Different Pulse Polarities. 49th AIAA Aerospace Sciences Meeting. Orlando, Florida. Jan 2011. Paper AIAA-2011-1275.

Roth J.R. Aerodynamic flow acceleration using paraelectric and peristaltic electrohydrodynamic effects of a one atmosphere uniform glow discharge plasma. Phys. Plasmas 10: 2117–26, 2003.

Roth,J.R., Sherman,D.M. and Wilkinson,S.P. "Boundary layer flow control with a one atmosphere uniform glow discharge surface plasma" 1998.AIAA Meeting (Reno, USA, January 1998) paper N98-0328

Roth,J.R. "Electrohydrodynamically induced airflow in a one atmosphere uniform glow discharge surface plasma". 1998.25th IEEE Int. Conf. Plasma Science (Raleigh, USA)

Roth,J.R., Sherman,D.M. and Wilkinson,S.P. "Electrohydrodynamic flow control with a glow discharge surface plasma". 2000. AIAA J. 38 1172-9

Roth,J.R., and Dai,X. "Optimization of the Aerodynamic Plasma Actuator as an Electrohydrodynamic (EHD) Electrical Device". 44th AIAA Aerospace Sciences Meeting and Exhibit 9 - 12 January 2006, Reno, Nevada

Roupassov,D.V., Zavyalov,I.N., Starikovskii,A.Yu., Saddoughi,S.G. "Boundary Layer Separation Plasma Control Using Low-Temperature Non-Equilibrium Plasma of Gas Discharge". 44th AIAA Aerospace Sciences Meeting and Exhibit, 9 - 12 January 2006, Reno, Nevada. AIAA 2006-373

Roupassov,D.V., Zavyalov,I.N., Starikovskii,A.Yu., Saddoughi,S.G. "Boundary Layer Separation Control by Nanosecond Plasma Actuators". AIAA 2007-4530. Joint Prorulsion Conference, 2007, Miami, Florida.

Roupassov,D., Nudnova,M., Nikipelov,A., Starikovskii,A. "Sliding DBD for Airflow Control: Structure and Dynamics". 46th AIAA Aerospace Sciences Meeting and Exhibit 7 - 10 Jan 2008 Grand Sierra Resort Hotel Reno, Nevada. Paper AIAA-2008-1367.

Roupassov D.V., Starikovskii A.Y., Development of Nanosecond Surface Discharge in "Actuator" Geometry. IEEE Transactions on Plasma Science Volume 36, Issue 4, Part 1, Aug. 2008 Page(s):1312 – 1313

Roupassov D., Nikipelov A., Nudnova M., Starikovskii A., Flow Separation Control by Plasma Actuator with Nanosecond Pulsed-Periodic Discharge. AIAA Journal (2009) 0001-1452 vol.47 no.1 (168-185)

Samimy M, Kim J-H, Kastner J, Adamovich I, and Utkin Y. 2007 Active control of high-speed, and high-Reynolds number jets using plasma actuators. Journal of Fluid Mechanics 578: 305-330, 2007.

Shang,J.S. "Plasma injection for hypersonic blunt body drag reduction". 2002.AIAA J. 40(6), 1178-1186.

Shneider M.N., Macheret S.O., and Miles R.B. Analysis of Magnetohydrodynamic Control of Scramjet Inlets. AIAA Journal. Vol. 42, No. 11, November 2004

Sidorenko,A.A., Zanin,B.Yu., Postnikov,B.V., Budovsky,A.D., Starikovskii,A.Yu., Roupassov,D.V., Zavialov,I.N., Malmuth,N.D., Smereczniak,P., Silkey,J.S. "Pulsed Discharge Actuators for Rectangular Wing Separation Control". 45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 January 2007, Reno, Nevada. AIAA 2007-941

Sigmond,R.S., and Lagstadt,I.H. "Mass and species transport in corona discharges". 1993. High Temp. Chem. Proces. 2 221-9 71 417-36

Starikovskaia,S.M., Anikin,N.B., Pancheshnyi,S.V., Starikovskii,A.Yu. "Time resolved emission spectroscopy and its applications to study of pulsed nanosecond high-voltage discharge". 2002. Proceedings of SPIE, V. 4460, pp. 63-73.

Starikovskii A.Yu. *Plasma supported combustion*, Proceedings of the Combustion Institute, 30, 2405-2417 (2005).

Starikovskii A., Anikin N., Kosarev I., Mintoussov E., Nudnova M., Rakitin A., Roupassov D., Starikovskaia S., Zhukov V., Nanosecond-Pulsed Discharges for Plasma-Assisted Combustion and Aerodynamics. Journal of Propulsion and Power 2008. 0748-4658 vol.24 no.6 (1182-1197)

Starikovskii A.Yu., Nikipelov A.A., Nudnova M.M. and Roupassov D.V. SDBD plasma actuator with nanosecond pulse-periodic discharge Plasma Sources Sci. Technol. 18 No 3 (August 2009) 034015 (17pp)

Thomas F.O., Corke T.C., Iqbal M., Kozlov A., and Schatzman D. Optimization of Dielectric Barrier Discharge Plasma Actuators for Active Aerodynamic Flow Control AIAA Journal Vol. 47, No. 9, September 2009

Unfer T. and Boeuf J.P. Modeling of a nanosecond discharge actuator", J. Phys. D: Appl. Phys. 42 194017 (2009)

Van Wie D. Paper presented at the RTO AVT Lecture Series on —Critical Technologies for Hypersonic Vehicle Development, at the von Kármán Institute, Rhode-St-Genèse, Belgium, 10- 14 May, 2004, and published in RTO-EN-AVT-116.

Van Wie D. M., Risha D. J., and Suchomel C. F., "Research Issues Resulting from an Assessment of Technologies for Future Hypersonic Aerospace Systems, AIAA-2004-1257, January 2004.

Velkoff,H. and Ketchman,J. "Effect of an electrostatic field on boundary layer transition" 1968.AIAA J. 16 1381-3

Visbal M.R., Gaitonde D.V. Control of Vortical Flows Using Simulated Plasma Actuators. 44th AIAA Aerospace Sciences Meeting and Exhibit 9 - 12 January 2006, Reno, Nevada AIAA 2006-505.

Yabe, A., Mori, Y., Hijikata, K. "EHD study of the corona wind between wire and plate electrodes",1978, AIAA J., Vol. 16 pp.340-5.

Zhukov V.P., Kindusheva S.V., Kirpichnikov A.A., Kosarev I.N., Starikovskaia S.M., Starikovskii A.Yu. Plasma Production for MHD Power Generation by Nanosecond Discharge. 44th AIAA Aerospace Sciences Meeting and Exhibit 9 - 12 Jan 2006. Reno Hilton. Reno, Nevada. Paper AIAA-2006-1370

Zouzou,N., Moreau,E., and Touchard,G. "Precipitation electrostatique dans une configuration pointe-plaque". 2006. J.Electrostat. 64 537-42

4

Numerical Investigation of Plasma Flows Inside Segmented Constrictor Type Arc-Heater

Kyu-Hong Kim
School of Mechanical and Aerospace Engineering, Seoul National University
Korea

1. Introduction

Arc-jet wind tunnels produce the flow conditions as an entry vehicle entering the atmosphere of a planet. They are widely used in experimental investigations of heat shield materials and thermal protection systems (TPSs). For a detailed analysis of experimental data, it is very important to know the conditions of the plasma flow that reaches a specimen of heat shield material or the TPS; e.g., the total pressure, total temperature and turbulence intensity of the plasma flow. However, it is hard to measure the properties of plasma flow because of its high temperature. In particular, it is almost impossible to measure the turbulence intensity inside an arc heater. On the other hand, if accurate physical models are adopted, computational fluid dynamics (CFD) can provide very useful information on plasma flow, and play an important role on the analysis of experimental data.

Several types of arc heaters have been developed according to the required range of total pressure and total enthalpy; these include the Huels type, constricted type, segmented constrictor type, inductively-coupled plasma (ICP), MPD (magnetoplasmadynamic) and so on. Among them, the segmented type and the Huels-type arc heaters have been widely used in aerospace applications since they can provide a large amount of plasma flow at a high total pressure condition. Compared to other types of arc heaters, the Huels arc heater produces a severe arc fluctuation phenomenon in time and space whereas a segmented arc heater produces stable arc flow throughout a long constrictor. At present, CFD analysis is more useful for the analysis of segmented type arc heaters since it is not easy to calculate a Huels-type arc heater accurately, even though state-of-the-art CFD technologies have been applied.

There have been several numerical investigations of the flow inside segmented arc heaters; these studies mainly focused on the improvement of physical models including radiation and the turbulence model. In the 1970s, Nicolet et al. (1975), based on the work of Watson and Pegot (1967), developed the ARCFLO numerical code. The ARCFLO code adopted a two-band radiation model and an algebraic turbulence model. However, since ARCFLO uses the space marching technique, it requires knowledge of downstream conditions in advance. Thus, the applicability of the ARCFLO code was quite restrictive in the actual design of arc heaters. Therefore, Kim et al. (2000) developed a time-marching code, ARCFLO2, to overcome this shortcoming. ARCFLO2 retained the two-band radiation model and the algebraic turbulence model. Sakai and Olejniczak (2001, 2003) improved the numerical models of ARCFLO and ARCFLO2 by adopting a new three-

band radiation model that was quite consistent with detailed line-by-line calculation; this code was named ARCFLO3. However, ARCFLO2 and ARCFLO3 could not always provide acceptable accuracy for a wide range of arc heaters since it could not predict turbulence physics accurately. Recently, Lee et al. (2007) improved ARCFLO3 and named their revised code ARCFLO4. The ARCFLO4 code adopted a two-equation turbulence model instead of an algebraic turbulence model while retaining the three-band radiation model of ARCFLO3. ARCFLO4 adopted k-ε (Jones & Launder, 1972), k-ω (Wilcox, 1998), and kω-SST model (Menter, 1994) which are able to express convective physics of turbulence and do not explicitly require a mixing length. Subsequently, ARCFLO4 was improved to handle an air–argon mixture as a working gas, and its numerical accuracy was sufficiently enhanced to be used as an arc heater design tool (Lee et al., 2008). Recently, ARCFLO4 was improved again to solve electronic field and magnetohydrodynamic (MHD) equations.

This chapter is focused on two issues. First, the accuracy level of recent CFD technology for arc-heated plasma flow is introduced and validated. As previously mentioned, arc heaters are widely used for high enthalpy flow experiments, and the values of total pressure, total enthalpy, and turbulence intensity at the surface of the specimen are important for a detailed analysis of experimental data. For example, in ablation experiments, the turbulent intensity of plasma flow may accelerate an ablation several-fold compared to laminar plasma flow. In order to validate the ARCFLO4 code, I first describe the results of the CFD analysis by comparison with real experimental data. Second, the potential use of the computational code as a design tool for arc heater is discussed. The empirical design process is possible if an established database exists based on the extensive experience of designers. However, if this is not the case, it is hard to determine the required total enthalpy, the amount of the heat flux that should be removed by cooling process, and so on. On the other hand, the required database can be obtained easily by CFD analysis if an accurate calculation of arc-heated flow is possible. Thus, in this chapter, I consider the performance variation of an arc heater according to changes in design variables in order to assess the use of ARCFLO4 as a design tool for arc heaters.

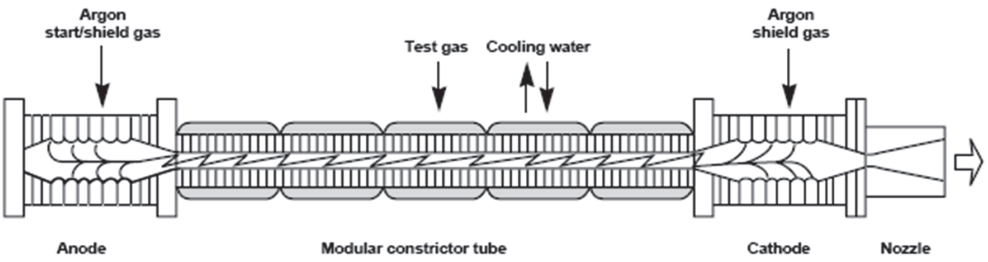


Fig. 1. Schematic Drawing of a Segmented Arc Heater

The contents of the remainder of this chapter are as follows. The governing equation and the physical modeling are explained in Section 2. In Section 3, validations and a flow analysis of the arc heater flow are described. Finally, Section 4 describes a parametric study process to examine the possibility of using a CFD code as an arc heater design tool.

2. Governing equations and physical modeling

2.1 Governing equations

The governing equations are hyperbolic 2-D axisymmetric Navier-Stokes equations that include Joule heating by arc, radiation, the swirling effect, and turbulence phenomena. The governing equations are as follows:

$$\frac{\partial \mathbf{Q}}{\partial t}+\frac{\partial \mathbf{E}}{\partial x}+\frac{\partial \mathbf{F}}{\partial y}+\mathbf{H}=\frac{\partial \mathbf{E_v}}{\partial x}+\frac{\partial \mathbf{F_v}}{\partial y}+\mathbf{H_v}+\mathbf{I} \tag{1}$$

where

$$\mathbf{Q}=\begin{bmatrix}\rho\\ \rho_{Ar}\\ \rho u\\ \rho v\\ \rho w\\ \rho E\end{bmatrix},\ \mathbf{E}=\begin{bmatrix}\rho u\\ \rho_{Ar}u\\ \rho uu+p\\ \rho uv\\ \rho uw\\ \rho uH\end{bmatrix},\ \mathbf{F}=\begin{bmatrix}\rho v\\ \rho_{Ar}v\\ \rho uv\\ \rho vv+p\\ \rho vw\\ \rho vH\end{bmatrix},\ \mathbf{H}=\frac{1}{y}\begin{bmatrix}\rho v\\ \rho_{Ar}v\\ \rho uv\\ \rho vv-\rho ww\\ 2\rho vw\\ \rho vH\end{bmatrix},$$

$$\mathbf{E_v}=\begin{bmatrix}0\\ \rho D_{Ar}\dfrac{\partial c_{Ar}}{\partial x}\\ \tau_{xx}\\ \tau_{xy}\\ \tau_{x\theta}\\ u\tau_{xx}+v\tau_{xy}+w\tau_{x\theta}-q_{c,x}-q_{R,x}\end{bmatrix},\ \mathbf{F_v}=\begin{bmatrix}0\\ \rho D_{Ar}\dfrac{\partial c_{Ar}}{\partial y}\\ \tau_{xy}\\ \tau_{yy}\\ \tau_{y\theta}\\ u\tau_{xy}+v\tau_{yy}+w\tau_{y\theta}-q_{c,y}-q_{R,y}\end{bmatrix},$$

$$\mathbf{H_v}=\frac{1}{y}\begin{bmatrix}0\\ \rho D_{Ar}\dfrac{\partial c_{Ar}}{\partial y}\\ \tau_{xy}\\ \tau_{yy}-\tau_{\theta\theta}\\ 2\tau_{y\theta}\\ u\tau_{xy}+v\tau_{yy}+w\tau_{y\theta}-q_{c,y}-q_{R,y}\end{bmatrix},\ \mathbf{I}=\begin{bmatrix}0\\ 0\\ 0\\ 0\\ 0\\ -j\cdot E\end{bmatrix}. \tag{2}$$

In practical applications, argon and air are injected separately in a segmented arc heater (Terrazas-Salinas & Cornelison, 1999). Thus, the continuity equation for argon gas is also solved. The thermodynamic and transport properties of the mixture, including Ar and Ar+, are calculated in a thermal equilibrium state. I is the Joule heating term and q_R is the radiant flux. As shown in Eq. 2, swirl velocity w is considered in the momentum equation and the energy equation to reflect the effects of swirl injection. Numerically, the governing equations are then discretized using the finite volume method. The AUSMPW+ scheme

(Kim et al., 2001) is used for the spatial discretization of the inviscid flux. A central difference scheme is used for the calculation of the viscous flux. For time integration, LU-SGS (Jameson & Yoon, 1987) method is used and the spatial accuracy is improved by MLP(Kim & Kim, 2005).

2.2 Physical modeling

2.2.1 Equilibrium composition, thermodynamic properties and transport properties

In this chapter, the working gas is assumed to be in a chemical equilibrium state. Chemical equilibrium is described by minimization of the Gibbs free energy formulation. The equilibrium composition is obtained with 11-species (N_2, O_2, Ar, N, O, NO, Ar+, N+, O+, NO+ and e-). Thermodynamic data of each species are from the NASA Glenn thermodynamic database (Mcbride et al., 2002) which is known to be valid for temperatures up to 20,000 K. Transport properties are computed using the approximate formulations of Gupta et al. (1990) and Yos (1963). The collision cross sections ($\pi\Omega ij(1,1)$, $\pi\Omega ij(2,2)$) including the air species are obtained from Park (2001) and Kim et al. (2006). However, for the ionized species, the effective Coulomb cross section of Gupta et al. (1990) and Yos (1963) is used. To solve the continuity equation for the argon gas, the effective binary diffusion coefficient is required. Thus,

$$D_{im} = \frac{1 - n_i}{\sum_{j=1, j\neq i}^{N_S} n_j / D_{ij}} \tag{3}$$

where D_{ij} is an ordinary binary diffusion coefficient and can be computed as follows:

$$D_{ij} = \frac{kT}{p\Delta_{ij}^{(1)}}, \tag{4}$$

with

$$\Delta_{ij}^{(1)} = \frac{8}{3}\left[\frac{2M_iM_j}{\pi RT\left(M_i + M_j\right)}\right]^{1/2} \pi\Omega_{ij}^{(1,1)} \tag{5}$$

In this computation, the continuity equation for argon gas does not distinguish between neutral argon and ionized argon. Thus, the diffusion coefficient for argon gas is finally calculated as follows (the details are introduced in Lee et al. (2008)):

$$D_{Ar} = \frac{D_{Arm}c_{Ar} + D_{Ar^+m}c_{Ar^+}}{c_{Ar} + c_{Ar^+}} \tag{7}$$

2.2.2 Turbulence model

Turbulence is known to play an important role in the flow physics of an arc heater. In numerical calculations, algebraic turbulence models have been used wherein the turbulent fluctuating quantities are correlated to the mean flow quantities using algebraic relations. However, such calculations do not reflect the convection effect of turbulence since the local

rate of turbulence production is equal to that of turbulence dissipation. In addition, the algebraic model needs the turbulence mixing length in order to provide an accurate numerical result. To overcome this shortcoming, more advanced turbulence models that possess less empirical turbulence coefficients are introduced for arc-heated flow (Lee et al. 2007). In this chapter, I introduce three two-equation turbulence models: the k–ε model of Jones and Launder (1972), the k–ω model of Wilcox (1998), and the k–ω SST model of Menter (1994).

2.2.2.1 *k*-ε two-equation model (Jones & Launder, 1972)

The k-ε model is the most widely known and extensively used two-equation eddy viscosity model. The main references for this model are described by Jones and Launder (1972). The model gives reasonably good results for free-shear-layer flows with relatively small pressure gradients. The Reynolds stresses are modeled as follows:

$$\tau_{tij} = 2\mu_t\left(s_{ij} - s_{nn}\delta_{ij}/3\right) - 2\rho k\delta_{ij}/3\,, \tag{6}$$

where μ_t is the eddy viscosity, s_{ij} is the mean-velocity strain-rate tensor, ρ is the fluid density, k is the turbulent kinetic energy, and δ_{ij} is the Kronecker delta. The eddy viscosity is defined as a function of the turbulent kinetic energy, k, and the turbulent dissipation rate, ε.

$$\mu_t = c_\mu f_\mu \rho k^2/\varepsilon\,. \tag{7}$$

The model coefficient, c_μ, is determined by equilibrium analysis at high Reynolds numbers, and the damping function, f_μ, is modeled in terms of a turbulence Reynolds number, $Re_t=\rho k^2/\varepsilon\mu$. The turbulence transport equations for the *k*-ε model are defined as follows:
Turbulence energy transport equation:

$$\frac{\partial\rho k}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho u_j\frac{\partial k}{\partial x_j} - \left(\mu + \frac{\mu_\tau}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right) = \tau_{tij}s_{ij} - \rho\varepsilon\,. \tag{8}$$

Energy dissipation transport equation:

$$\frac{\partial\rho\varepsilon}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho u_j\varepsilon - \left(\mu + \frac{\mu_\tau}{\sigma_\varepsilon}\right)\frac{\partial\varepsilon}{\partial x_j}\right) = c_{\varepsilon 1}\frac{\varepsilon}{k}\tau_{tij}s_{ij} - c_{\varepsilon 2}\rho\frac{\varepsilon^2}{k}\,. \tag{9}$$

The model constants are defined as

$$c_\mu = 0.09\,,\ c_{\varepsilon 1} = 1.45\,,\ c_{\varepsilon 2} = 1.92\,,\ \sigma_k = 1.0\,,\ \sigma_\varepsilon = 1.3\,,\ Pr_t = 0.9\,. \tag{10}$$

2.2.2.2 *k*-ω two-equation model (WILCOX, 1998)

The *k*-ω model is a well known and widely tested two-equation eddy viscosity model. Convective transport equations are solved for the turbulent kinetic energy and its specific dissipation rate, k and ω, respectively. The *k*-ω model has proven to be superior in numerical stability to the *k*-ε model primarily in the viscous sub-layer near the wall. In the logarithmic region, the model gives good agreement with experimental data for mild adverse pressure gradient flows. The Reynolds stresses are modeled in terms of the eddy viscosity as follows:

$$\tau_{tij} = 2\mu_t\left(s_{ij} - s_{nn}\delta_{ij}/3\right) - 2\rho k\delta_{ij}/3\,. \tag{11}$$

The eddy viscosity is defined as a function of the turbulent kinetic energy, *k*, and the specific rate of dissipation, *ω*, as follows:

$$\mu_t = \rho k/\omega\,. \tag{12}$$

The two transport model equations are defined as follows:
Turbulence energy transport equation:

$$\frac{\partial\rho k}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho u_j k - \left(\mu + \sigma^*\mu_\tau\right)\frac{\partial k}{\partial x_j}\right) = \tau_{tij}s_{ij} - \beta^*\rho\omega k\,, \tag{13}$$

Specific dissipation rate equation:

$$\frac{\partial\rho\omega}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho u_j\omega - \left(\mu + \sigma\mu_\tau\right)\frac{\partial\omega}{\partial x_j}\right) = \alpha\frac{\omega}{k}\tau_{tij}s_{ij} - \beta\rho\omega^2\,. \tag{14}$$

As described by Wilcox (1998), the model constants are defined as

$$\alpha = \frac{5}{9},\ \beta = \frac{3}{40},\ \beta^* = \frac{9}{100},\ \sigma = 0.5,\ \sigma^* = 0.5,\ \mathrm{Pr}_t = 0.9\,. \tag{15}$$

2.2.2.3 *k*-ω SST two-equation model (Menter, 1994)

The *k*-ω SST shear-stress-transport model combines several desirable elements of existing two-equation models. The two major features of this model are a zonal weighting of model coefficients and a limitation on the growth of the eddy viscosity in rapidly strained flows. The zonal modeling uses Wilcox's *k*-ω model near solid walls and the standard k-ε model near boundary layer edges and in free-shear layers. This switching is achieved with a blending function of the model coefficients. The eddy viscosity is defined as the following function of the turbulent kinetic energy, *k*, and the specific dissipation rate or turbulent frequency, ω:

$$\mu_t = \frac{\rho k/\omega}{\max\left[1, \Omega F_2/a_1\omega\right]},\quad a_1 = 0.31\,. \tag{16}$$

In turbulent boundary layers, the maximum value of the eddy viscosity is limited by forcing the turbulent shear stress to be bounded by the turbulent kinetic energy time a_1. This effect is achieved with an auxiliary function F_2 and the absolute value of the vorticity, Ω. The auxiliary function, F_2, is defined as a function of wall distance *y* as follows:

$$F_2 = \tanh\left\{\left(\max\left[2\frac{\sqrt{k}}{0.09\omega y}, \frac{500\mu}{\rho\omega y^2}\right]\right)^2\right\}. \tag{17}$$

The transport equation for the turbulent kinetic energy is

$$\frac{\partial\rho k}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho u_j k - \left(\mu + \sigma_k\mu_\tau\right)\frac{\partial k}{\partial x_j}\right) = \tau_{tij}s_{ij} - \beta^*\rho\omega k \tag{18}$$

and the transport equation for the specific dissipation of turbulence is

$$\frac{\partial \rho\omega}{\partial t}+\frac{\partial}{\partial x_j}\left(\rho u_j\omega-\left(\mu+\sigma\mu_t\right)\frac{\partial \omega}{\partial x_j}\right)=p_\omega-\beta\rho\omega^2+2(1-F_1)\frac{\rho\sigma_{\omega 2}}{\omega}\frac{\partial k\partial\omega}{\partial x_j\partial x_j}, \tag{19}$$

where the last term of Eq. 19 represents the cross-diffusion term that appears in the transformed *k*-ω equation from the original *k*-ε equation. The production term of ω is sometimes approximated as proportional to the absolute value of vorticity, as follows:

$$p_\omega\equiv 2\gamma\rho\left(2s_{ij}-\omega s_{nn}\delta_{ij}/3\right)s_{ij}\approx\gamma\rho\Omega^2. \tag{20}$$

The function F_1 takes the value of one on no-slip surfaces and near one over a large portion of the boundary layer, and goes to zero at the boundary layer edge. This auxiliary blending function, F_1, is defined as

$$F_1=\tanh\left\{\left(\min\left[\max\left[\frac{\sqrt{k}}{0.09\omega y},\frac{500\mu}{\rho\omega y^2}\right],\frac{4\rho\sigma_{\omega 2}k}{CD_{k\omega}y^2}\right]\right)^4\right\}, \tag{21}$$

where

$$CD_{k\omega}=\max\left[2\rho\sigma_{\omega 2}\frac{1}{\omega}\frac{\partial k}{\partial x_j}\frac{\partial\omega}{\partial x_j},10^{-20}\right], \tag{22}$$

where $CD_{k\omega}$ stands for cross-diffusion in the *k*-ω model. The constants of the SST model are

$$a_1=0.31,\ \beta^*=0.09,\ \kappa=0.41. \tag{23}$$

The model coefficients β, γ, σ_k and σ_ω denoted with the symbol ϕ are defined by blending the coefficients of the original k-ω model, denoted as ϕ_1, with those of the transformed k-ε model, denoted as ϕ_2.

$$\phi=F_1\phi_1+(1-F_1)\phi_2,\ \text{where}\ \phi=\{\sigma_k,\sigma_\omega,\beta,\gamma\} \tag{24}$$

with the coefficients of the original models defined as
Inner model coefficients:

$$\sigma_{k1}=0.85,\ \sigma_{\omega 1}=0.5,\ \beta_1=0.075,\ \gamma_1=\beta_1/\beta^*-\sigma_{\omega 1}\kappa^2/\sqrt{\beta^*}=0.553. \tag{25}$$

Outer model coefficients:

$$\sigma_{k2}=1.0,\ \sigma_{\omega 2}=0.856,\ \beta_2=0.0828,\ \gamma_2=\beta_2/\beta^*-\sigma_{\omega 2}\kappa^2/\sqrt{\beta^*}=0.440. \tag{26}$$

2.2.3 Joule heating model

In general, Maxwell's equations should be solved in order to calculate Joule heating. However, if the current distribution of the arc heater is known, the Joule heating can be calculated using Ohm's law(the current distribution inside a constrictor is constant since the

constrictor wall is insulated electrically). With the assumption that the voltage gradient is independent of radius, Joule heating can be calculated as follows. Ohm's law for a cylindrical column is given by

$$j(x,y) = \sigma(x,y) \cdot E(x) \,, \tag{27}$$

where

$$E(x) = \frac{j(x,y)}{\sigma(x,y)} = \frac{\int_0^R 2\pi y j(x,y) dy}{\int_0^R 2\pi y \sigma(x,y) dy} = \frac{I}{\int_0^R 2\pi y \sigma(x,y) dy} \tag{28}$$

and

$$I = \int_0^R 2\pi y j(x,y) dy = const \tag{29}$$

from Kirchhoff's law of conservation of current. Finally, Joule heating can be calculated as

$$j(x,y) \cdot E(x) = \frac{I^2 \sigma(x,y)}{\left[\int_0^R 2\pi y \sigma(x,y) dy\right]^2} \,. \tag{30}$$

Figure 2 shows the typical current distribution in an arc heater along the axial direction.

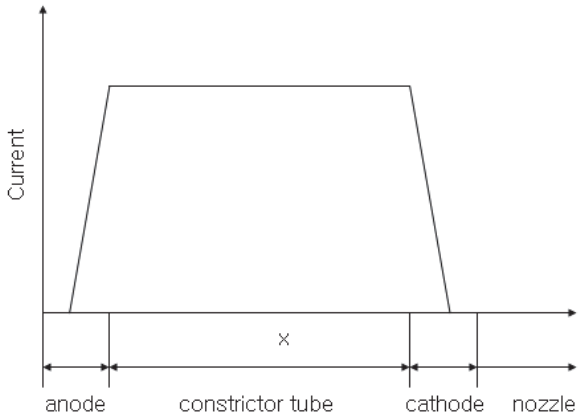


Fig. 2. Current Distribution in the Arc Heater

2.2.4 Radiation model

For high temperature flows, radiation becomes an important heat transfer mode in addition to thermal convection. The radiative transfer equation is as follows:

$$-\frac{1}{\rho \kappa_\nu} \frac{dI_\nu}{ds} = I_\nu - B_\nu \,, \tag{31}$$

where I_ν is the radiative intensity traveling along a ray denoted by, s. The absorption coefficient κ_ν is a function of frequency, temperature, and pressure; these are obtained from experiment. The Planck function B_ν is expressed as follows:

$$B_\nu = \frac{2h}{c^2} \frac{\nu^3}{e^{h\nu/kT} - 1} \tag{32}$$

The radiant flux per unit frequency expressed in cylindrical coordinates can be calculated using Eq. 33, when the radiative intensity at a given point is calculated for all directions:

$$q_\nu(r) = \int_\Omega I_\nu(r) \cos\theta \, d\Omega \tag{33}$$

where θ is the angle between the ray and the outward normal to the cylindrical surface and Ω is the solid angle. The total radiative flux at any radius r can be expressed as,

$$q_R(r) = \int_0^\infty q_\nu(r) d\nu \, . \tag{34}$$

Under the band model assumption the total radiative flux can be written as,

$$q_R(r) = \sum_{l=1}^{m} q_l(r) \, , \tag{35}$$

where m is the total number of bands and $q_l(r)$ is the radiant flux contribution from the l^{th} band to the total flux which is given by

$$q_l(r) = \int_{\Delta\nu_l} q_\nu(r) d\nu \, . \tag{36}$$

where $\Delta\nu_l$ is the band-width for the l^{th} band. For accurate calculation of the radiant flux, a line-by-line calculation (Whiting et al., 1996) is the most desirable approach. However, such a line-by-line calculation requires a huge computational cost. Thus, Sakai and Olejniczak (2003) developed a new three band model that can compute radiative transport equations 400 times faster than the line-by-line calculation without losing accuracy. In this chapter, all calculations are performed by using this three-band radiation model. Details of the three band model are introduced in Sakai and Olejniczak (2003).

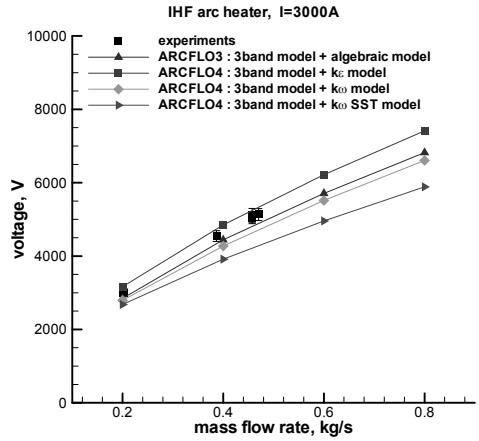
3. Validation of the CFD code

3.1 Validation

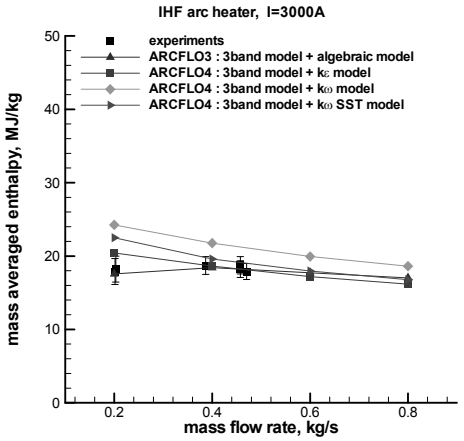
3.1.1 NASA Interaction Heating Facility (IHF) arc heater

The IHF is operated with a 60 MW constricted arc heater that can operates at a pressure range of 1 to 9 atm and an enthalpy level of 7 to 47MJ/kg. In an experiment, Hightower et al. (2002) measured the performance of the IHF arc heater and Sakai and Olejniczak (2003) numerically calculated the flow inside the IHF arc heater. In this section, I compare the calculated results from ARCFLO4 with experimental data and from the data calculated by the ARCFLO3 code which is developed by Sakai and Olejniczak (2003). The calculated voltage, mass averaged enthalpy, chamber pressure, and efficiency are compared in terms of mass flow rate. Results are illustrated for the cases of I=3000A and 6000A. Figure 3 shows

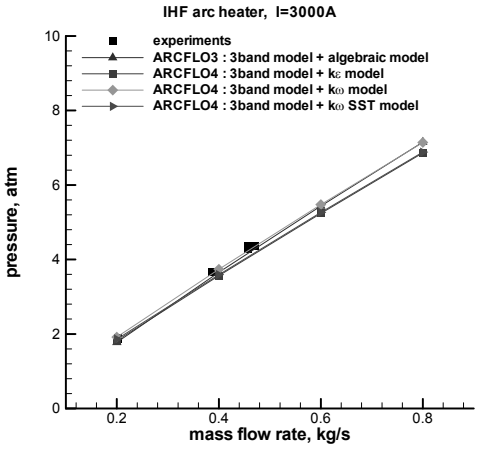
calculated and experimental data for the case of I=3000A. The numerical results using three two-equation turbulence models are illustrated. As shown in the figure, the k-ε turbulence model predicted the operational data most accurately. It appears that k-ω turbulence model overestimated the mass-averaged enthalpy and efficiency. The k-ω SST turbulence model predicted mass-averaged enthalpy and chamber pressure fairly well, whereas the calculated voltage and efficiency deviated somewhat from the experimental data. That is, the combination of the k-ε turbulence model and the three-band radiation model seems to be appropriate for arc heater flow calculation.



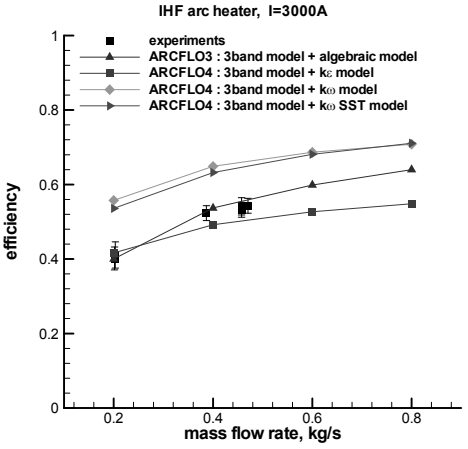
(a) Voltage



(b) Mass-Averaged Enthalpy



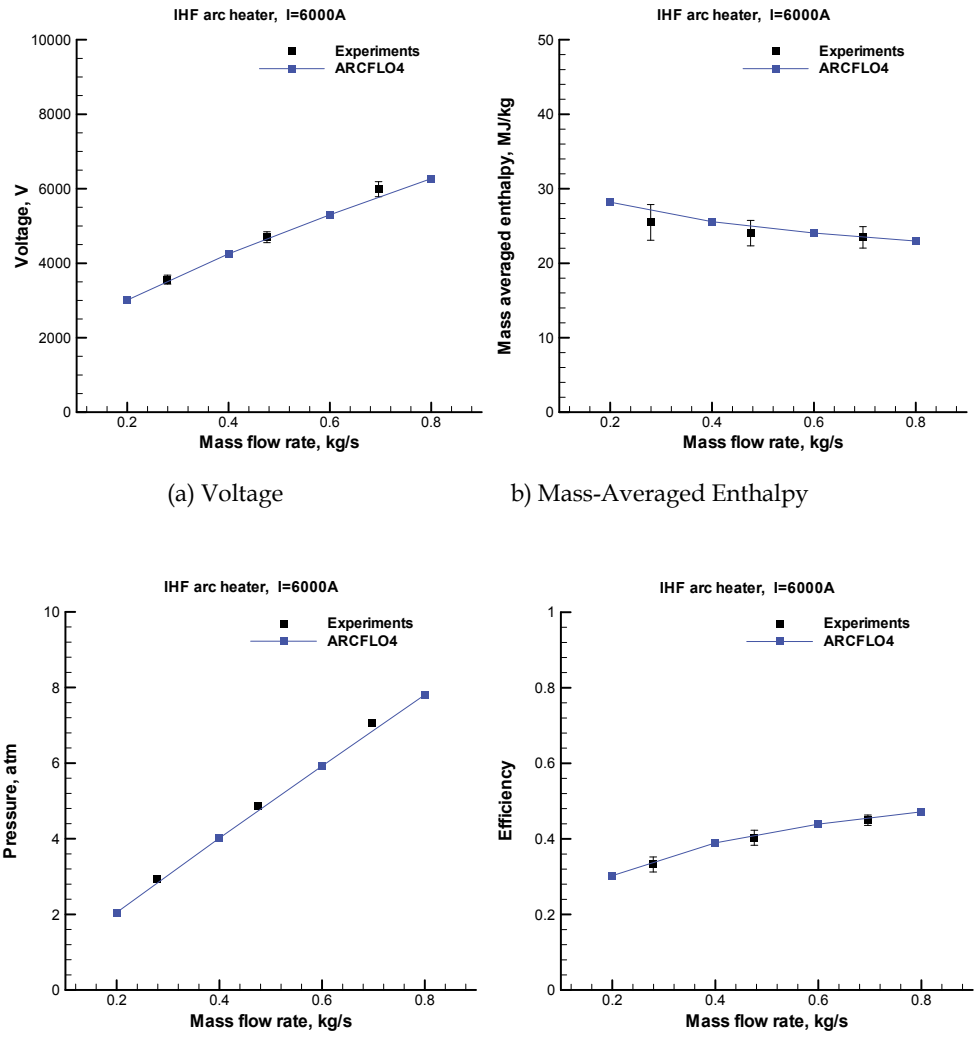
(c) Pressure



(d) Efficiency

Fig. 3. Comparison between Calculation and Experiment (Lee et al., 2007)

Figure 4 shows the results for I = 6000A. As shown in the figure, the overall results follow the same tendency as the case of I = 3000A. Again, the computational results of ARCFLO4 predicted the flow inside the arc heater accurately. The combination of the two-equation turbulence model and the three-band radiation model provide accurate performance parameters for the arc heater. Since the numerical solution using two-equation turbulence model does not require user experience for turbulent flows, it is expected that the numerical approach of ARCFLO4 can be applied to a wide range of arc heaters.



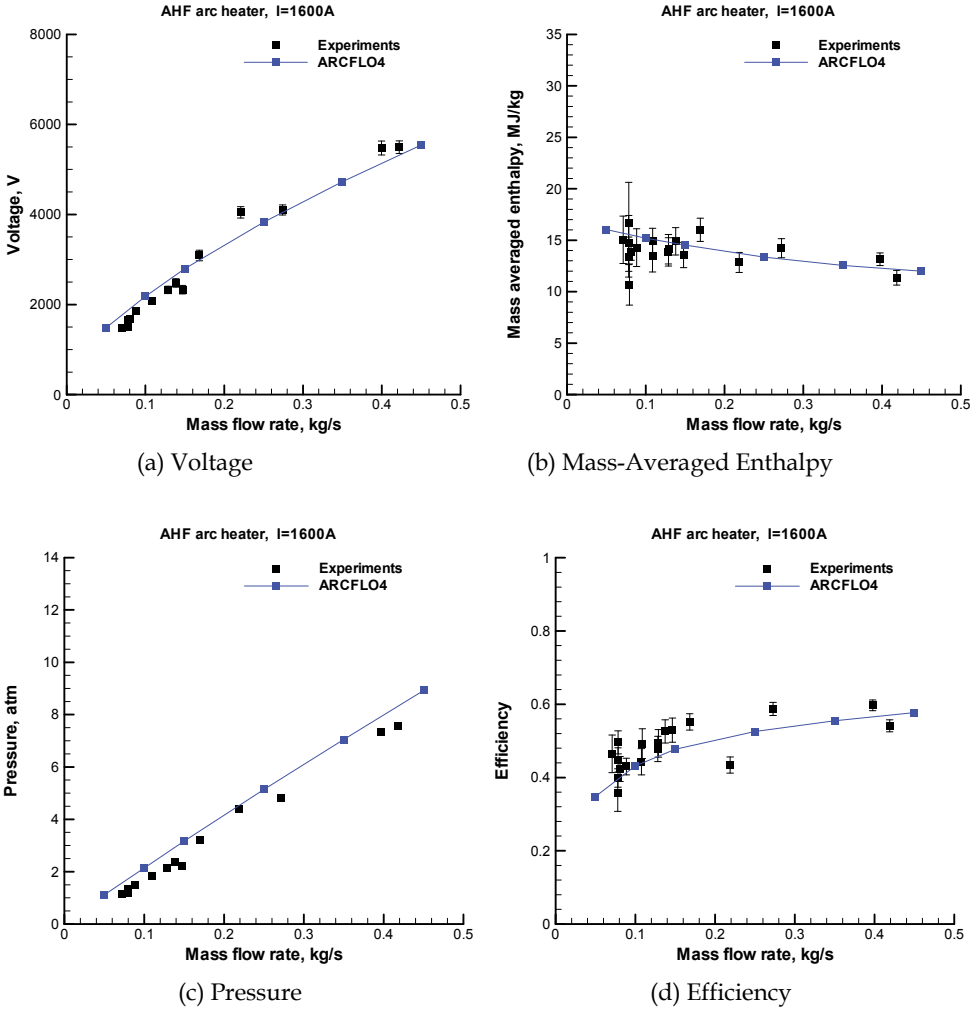
(a) Voltage b) Mass-Averaged Enthalpy

(c) Pressure (d) Efficiency

Fig. 4. Comparison between Calculation and Experiment

3.1.2 NASA Aerodynamic Heating Facility (AHF) arc heater

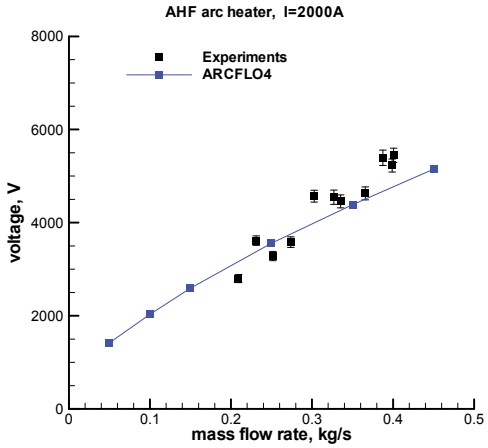
The AHF is operated with a 20MW constricted arc heater that has a pressure range of 1 to 9atm and an enthalpy level of 1 to 33MJ/kg. Hightower et al. (2002) measured the performance of the AHF experimentally and Sakai and Olejniczak (2003) computed the corresponding flows using ARCFLO3. In this section, a comparison between computation and experiment is carried out in a manner similar way to previous IHF analyses. The calculated voltage, mass-averaged enthalpy, chamber pressure, and efficiency are compared using experimental data for the cases of I = 1600 and 2000 A. Figure 5 shows results for I = 1600 A. Similar to the IHF cases, the calculated results by ARCFLO4 are in good agreement with the experimental data.



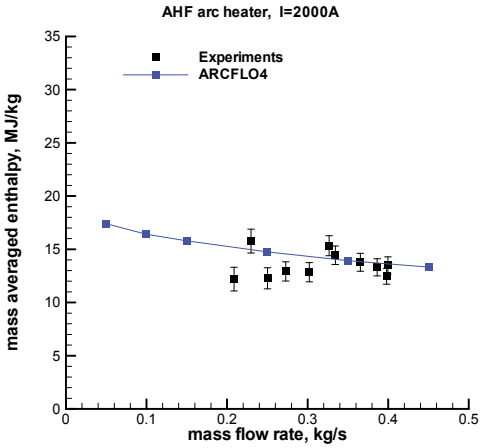
(a) Voltage (b) Mass-Averaged Enthalpy

(c) Pressure (d) Efficiency

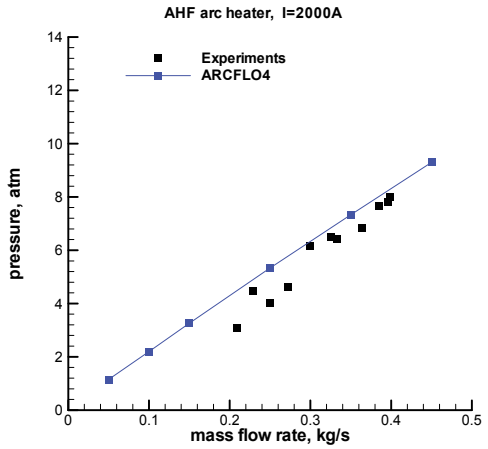
Fig. 5. Comparison between Calculation and Experiment



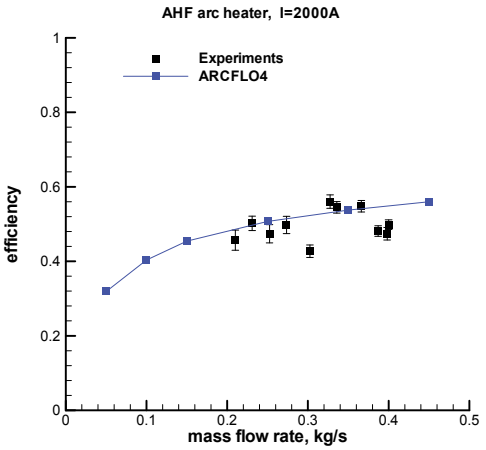
(a) Voltage



(b) Mass-Averaged Enthalpy



(c) Pressure



(d) Efficiency

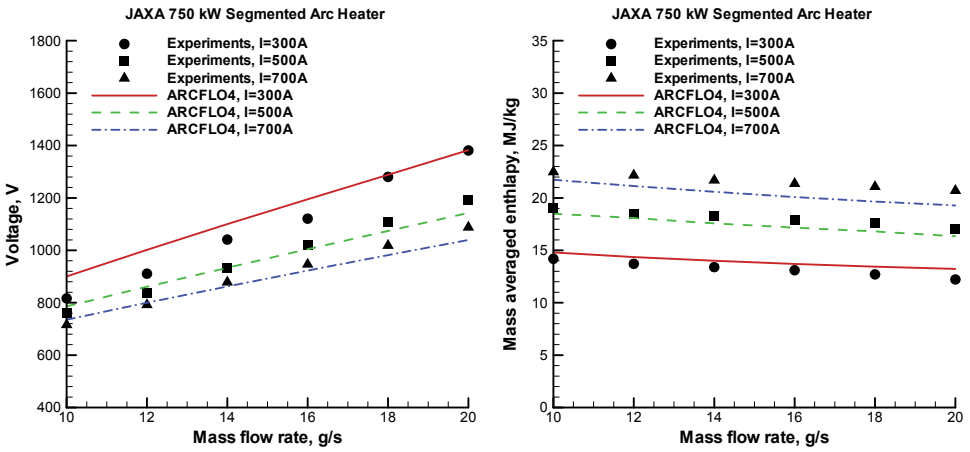
Fig. 6. Comparison between Calculation and Experiment

Figure 6 shows the results for I = 2000 A. As shown in the figure, the overall results show a tendency similar to the case of I = 1600 A and are in good agreement with the experimental results. Considering the results described in Sections 3.1.1 and 3.1.2, we can say that the ARCFLO4 code predicted the arc heater flow accurately for high electric power cases.

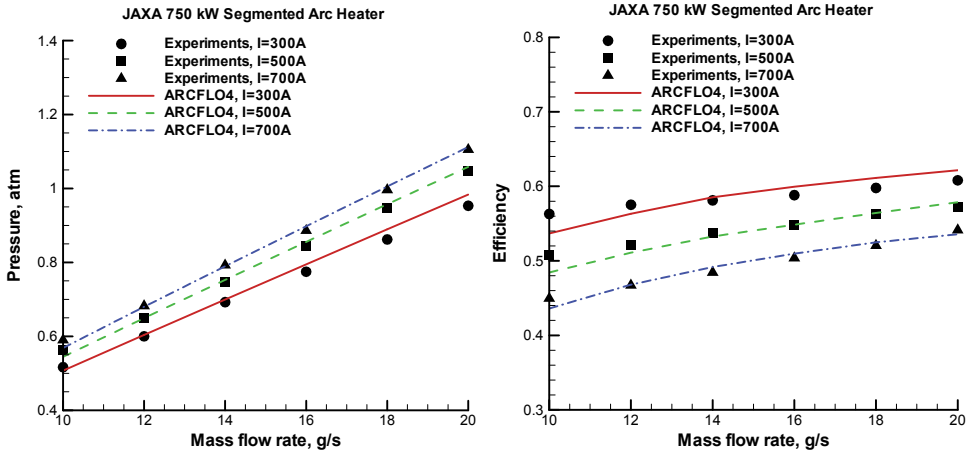
3.1.3 JAXA 750KW arc heater

The Japan Aerospace Exploration Agency (JAXA) has serviced a 750 kW segmented arc heater since the 1990s, and its operational data are available through the references of

Matsuzaki et al. (2002) and Sakai et al. (2007). The JAXA 750 kW segmented arc heater operates at a current between 300 and 700 A and a mass flow rate between 10 and 20 g/s. The constrictor length and diameter are 39 cm and 2.54 cm, respectively. The diameter of the nozzle throat is 2.5 cm. The diameter and the radius of the electrode is 7.6 cm and 1.9 cm, respectively. In this section, a numerical flow calculation of the JAXA 750 kW arc heater is introduced as a low electric power case. The voltage between electrodes, the mass-averaged enthalpy at the nozzle throat, the pressure in the cathode chamber, and the arc heater efficiency are calculated and compared to the experimental data.



(a) Voltage (b) Mass-Averaged Enthalpy



(c) Pressure (d) Efficiency

Fig. 7. Comparison between Calculation and Experiment (Lee & Kim, 2010)

Figure 7 shows a comparison of the operational data plotted in terms of mass flow rates. As shown in the figure, the computed operational data are in good agreement with the experimental data. Thus, it is confirmed that the ARCFLO4 simulation of low electric power segmented arc heater flows is valid.

3.1.4 150KW arc heater

A 150 kW arc heater in Korea was analyzed in order to validate ARCFLO4 for a lower electric power regime. This arc heater is basically a Hules-type heater. However, to stabilize the arc, the constrictor is located at the center of the heater. The details of the configurations are shown in Fig. 8, and the test cases for present analysis are given in Table 1.

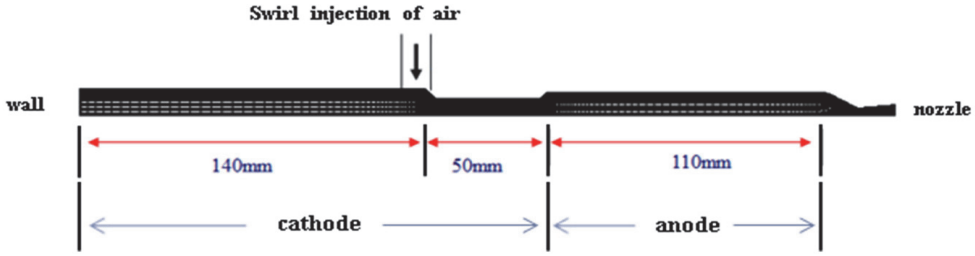


Fig. 8. Computational Grid

	Current(Amphere)	Mass flow rate(g/s)
CASE1	363	11.78
CASE2	393.3	10.11
CASE3	383	9.08
CASE4	374.4	7.53

Table 1. Test Cases

Generally, radiant heat flux is mainly generated at the constrictor and has almost zero value at the cathode and anode for the case of a long constrictor. Therefore, the ARCFLO4 code calculates the radiant flux using the assumption of long cylindrical coordinates. However, this 150 kW arc heater has a relatively short constrictor length, so the assumption is not valid. Considering the short length of the constrictor, the calculation of radiant flux was slightly corrected using a configuration factor, as shown in Fig. 10. The details of the correction are available in Han et al., 2011.

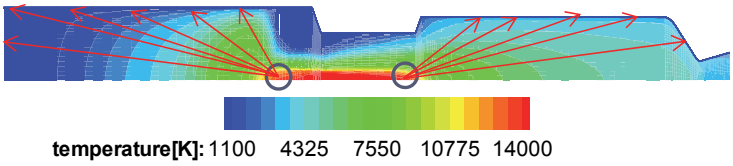


Fig. 9. Correction of Radianit Heat Flux Using Configuration Factor (Han et al., 2011)

Table 2 shows a comparison of the ARCFLO4 numerical results and the experimental results. The table shows that the calculated voltage and pressure are in very good agreement with the experimental data. That is, ARCFLO4 showed good accuracy again for the flow inside the low electric power arc heater.

	Pressure(atm)			Voltage(volt)		
	Cal.	Exp.	Error	Cal.	Exp.	Error
Case1	6.25	6.33	1.3%	385	392	2.8%
Case2	5.62	5.60	0.35%	345	344	0.3%
Case3	5.05	4.93	2.4%	328	320	2.5%
Case4	4.41	4.56	3.1%	325	335	2.8%

Table 2. Comparisons between Calculations and Experiments (Han et al., 2011)

Considering the results described in Sections 3.1.1 to Sec. 3.1.4, the ARCFLO4 code predicted the flow inside the arc heater accurately for a wide range of electric power (150 kW to 60 MW). It is also confirmed that the turbulence model used in ARCFLO4 reflected the convection physics of turbulence properly near the wall region.

4. CFD code as a design tool of the arc heater

The NASA Ames Research Center developed a segmented arc heater in the 1960s. Currently, NASA Ames has three segmented arc heater facilities: the 20 MW Aerodynamic Heating Facility, the 20 MW Panel Test Facility, and the 60 MW Interactive Heating Facility (Terrazas-Salinas and Cornelison, 1999). In the 1990s, Europe and Japan began to develop segmented arc heaters. In Europe, a 6 MW segmented arc heater was developed and operated with an L3K arc heated facility of the German Aerospace Center (Smith et al., 1996). Recently, 70 MW segmented arc heater was added to the SCIROCCO arc heated facility of the Italian Aerospace Research Center (Russo, 1993). Japan has serviced the 750 kW segmented arc heater since the 1990s. Despite these arc heater development experiences, a design process has been accomplished by only a few research centers and companies. In the development stage, there was probably considerable trial and error since the flow phenomena inside segmented arc heaters had not been characterized. Also, the higher cost would have been spent during the development of the segmented arc heater. In an effort to reduce the difficulties and cost during arc heater development, Lee et al. (2007, 2008) recently developed the ARCFLO4 computational code to study the flow physics in segmented arc heaters. As described in Section.3, the code accurately simulated existing arc heaters under various operating conditions. It predicted well the operational data of the AHF, IHF (Lee et al., 2007, 2008) and JAXA 750 kW arc heater (Lee & Kim, 2010). Since ARCFLO4 can accurately predict operational data and the wall heat energy loss, development costs can be reduced without previous design experience.

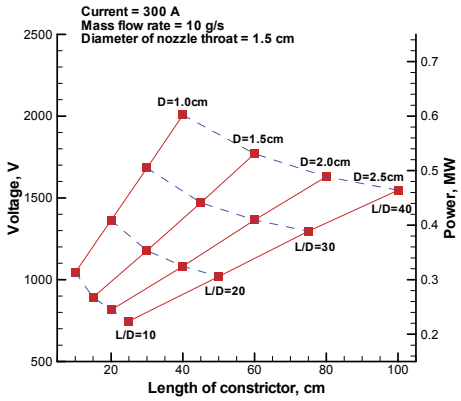
In this section, the effects of configuration and input operational conditions on the performance of an arc heater are investigated in order to provide fundamental data for the design of segmented arc heaters. A parametric study is performed to determine the main design variables that strongly affect arc heater performance. First, performance changes in terms of constrictor length, constrictor diameter, and nozzle throat diameter are investigated. Then, performance changes due different input currents and mass flow rates are examined.

4.1 Parametric study

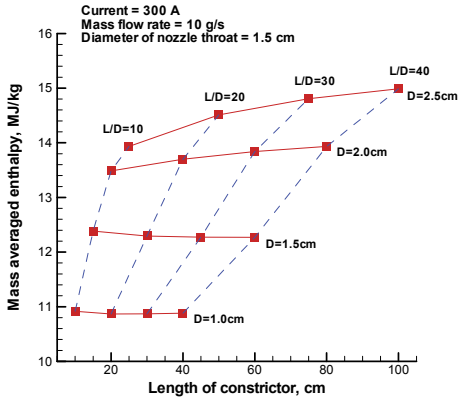
The relationship between performance and main design parameters, such as configuration and input operational conditions is investigated. The 750 kW JAXA segmented arc heater is chosen as a baseline model. To study the effect of configuration on arc heater flows, a constrictor length, a constrictor diameter, and a nozzle throat diameter are changed. Then, the input current and mass flow rate are changed to determine the effect of input operational conditions on arc heater flows.

4.1.1 Length of the constrictor

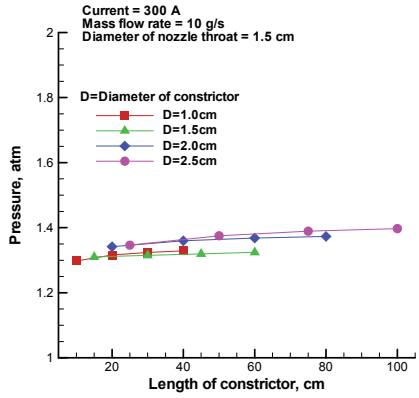
Generally, the arc length inside a segmented arc heater is similar to the constrictor length. Thus, the constrictor length is one of the key factors that affects arc heater flows. In this section, a parametric study according to the various constrictor lengths is described. The constrictor length varies from 10 to 100 cm with other parameters are fixed for comparison. In order to maintain an input electric power lower than 1 MW, a current of 300 A and a mass flow rate of 10 g/s were selected. The nozzle throat diameter is 1.5 cm. Figure 10 shows



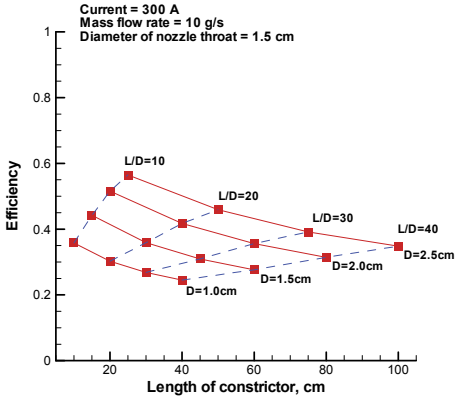
(a) Voltage & Power



(b) Mass-Averaged Enthalpy



(c) Pressure



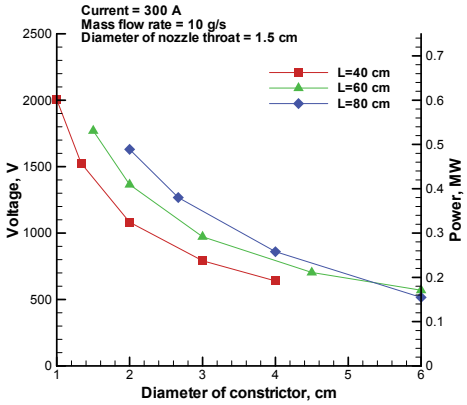
(d) Efficiency

Fig. 10. Operational Data (Lee & Kim, 2010)

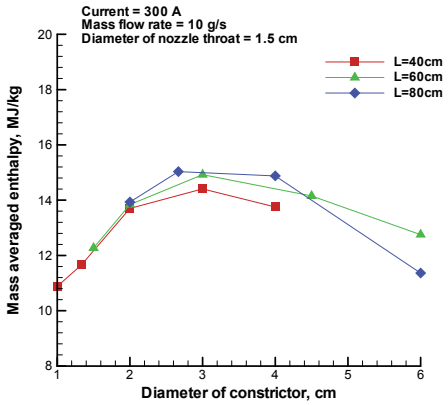
operational data in terms of a constrictor length at specific constrictor diameters. As shown in Fig. 10a, the voltage and the electric power are increased proportionally to the constrictor length. On the other hand, as shown in Figs. 10b and 10c, the effects of constrictor length on the mass-averaged enthalpy and the cathode chamber pressure are relatively small. It is shown that the efficiency decreases as the constrictor length increases. In general, the efficiency is strongly related to the amount of heat energy loss at the arc heater wall. The heat energy loss per unit length increases and the electric power input per unit length decreases, by increasing the constrictor length. Therefore, the longer the constrictor length, the lower the total efficiency becomes.

4.1.2 Diameter of the constrictor

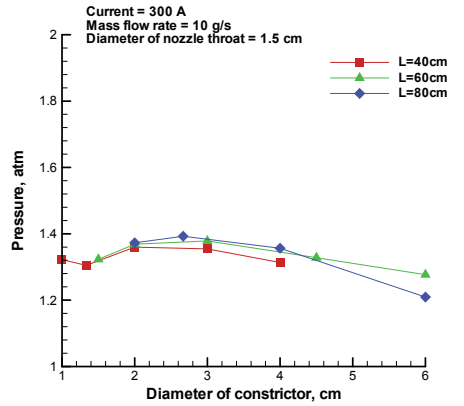
The effects of the constrictor diameters are also investigated. The constrictor diameters vary from 1.0 to 6.0 cm, while other configurations are fixed. The nozzle throat diameter is 1.5 cm. The current and mass flow rate are also fixed at 300 A and 10 g/s, respectively. Figure 11



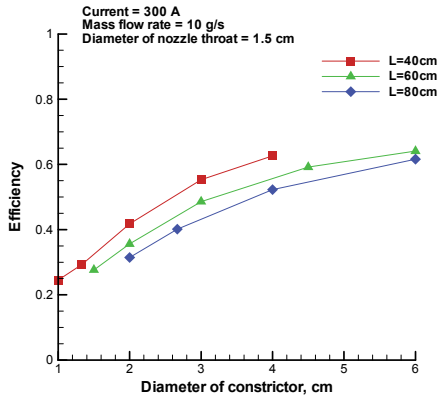
(a) Voltage & Power



(b) Mass-Averaged Enthalpy



(c) Pressure



(d) Efficiency

Fig. 11. Operational Data (Lee & Kim, 2010)

shows operational data in terms of constrictor diameter. As shown in the figure, the voltage, mass-averaged enthalpy, and efficiency are strongly affected by the constrictor diameter. As shown in Fig. 11a, the voltage and the electric power increase as the constrictor diameter decreases. For the mass-averaged enthalpy, the effect of the constrictor diameter is greater than that of the constrictor length, as shown in Figs. 10b and 11b. In Fig. 11c, we note that the cathode chamber pressure is weakly affected by the constrictor diameter. Finally, Fig. 11d shows that the efficiency decreases as the constrictor diameter decreases.

To understand the change in efficiency, we consider the heat energy loss on the arc heater wall as illustrated in Fig. 12. In the figure, as the constrictor diameter decreases, both the conductive and radiant energy losses increase, and thus the efficiency decreases. Generally, if a constrictor diameter decreases, the quantity of injecting working gas per unit area increases. Thus, the axial speed of the working gas increases, and thus a viscous dissipation phenomenon due to turbulence is strongly generated near the wall. Therefore, the heat energy loss by thermal conduction increases as the constrictor diameter decreases. Moreover, the distance from the core to the wall is small; thus, only a small amount of radiation is absorbed by the surrounding gas on its way to the wall.

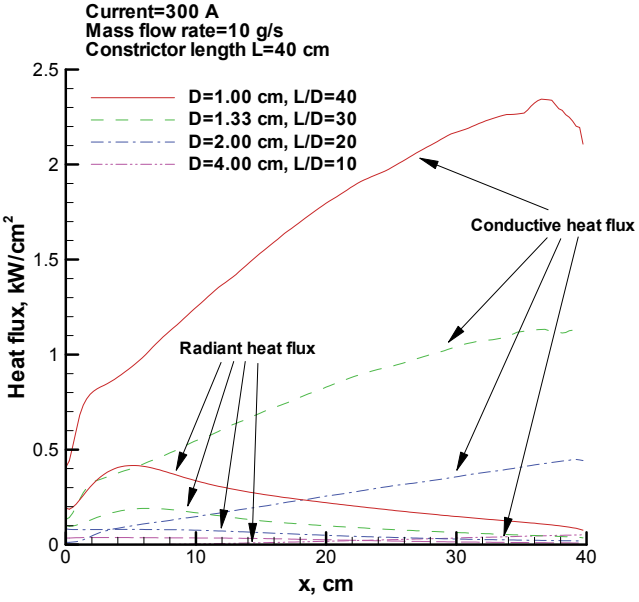


Fig. 12. Heat Flux (Lee & Kim, 2010)

The effect of the ratio of constrictor length to constrictor diameter, L/D, on the stability of an arc discharge is investigated. Figure 13 shows the temperature distribution in the radial direction. In the figure, we can define a region where the temperature is greater than 9,000 K and the current density is high, as an arc column. It is shown that the thickness of the arc column is large at the upstream region of the constrictor where L/D is greater than 30. Also,

the arc column broadens as L/D increases. If an arc column broadens, there is not enough room for the arc column to fluctuate and the stability of an arc discharge improves. Generally, it is known that L/D should be greater than 30 to stabilize an arc discharge (Sakai et al., 2007).

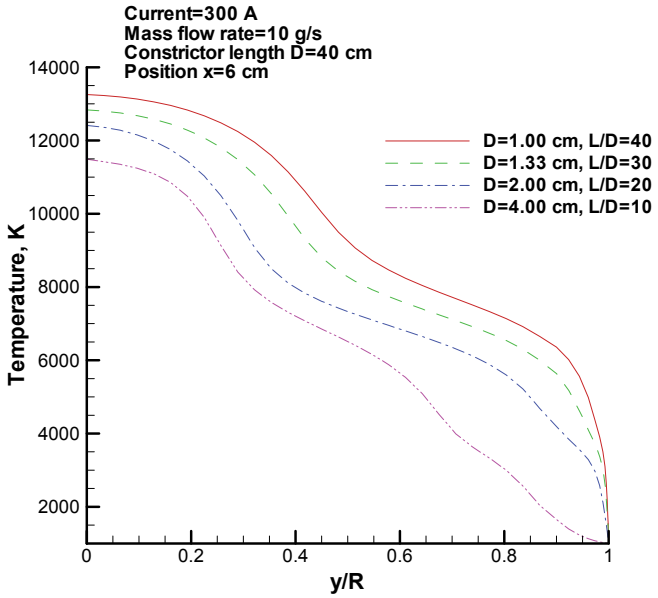


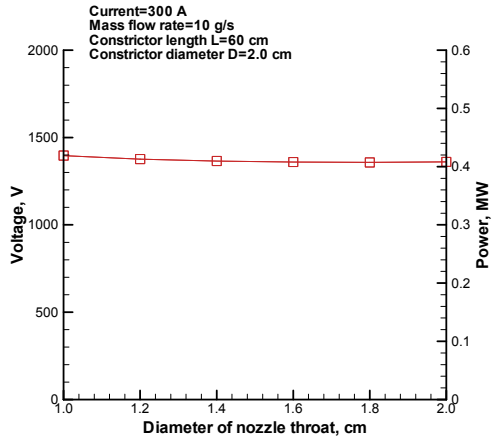
Fig. 13. Temperature (Lee & Kim, 2010)

4.1.3 Diameter of nozzle throat

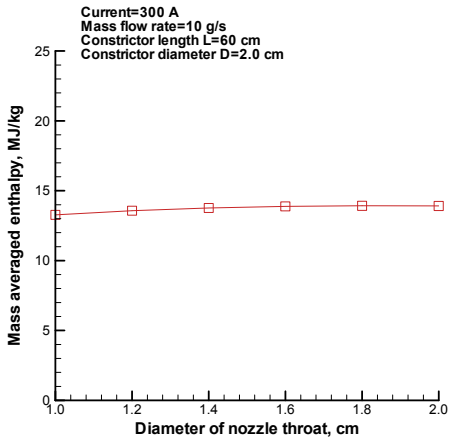
To investigate the effect of nozzle throat diameter on the arc heater flow, the nozzle throat diameter is chosen to vary from 1.0 to 2.0cm, while other parameters are fixed. The length and the diameter of the constrictor are 60.0 cm and 2.0 cm, respectively. Figure 14 shows operational data in terms of the nozzle throat diameter. As shown in the figure, the nozzle throat diameter does not affect operational data, such as electric voltage, mass averaged enthalpy, and efficiency. However, the chamber pressure is strongly affected by the nozzle throat diameter since the pressure is inversely proportional to nozzle area for a fixed mass flow rate. The pressure decreases as the nozzle throat diameter increase.

4.1.4 Input current

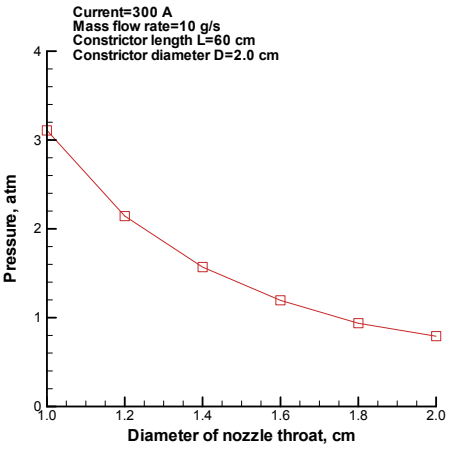
When designing a segmented arc heater, a range of input currents must be determined as well as arc heater configurations. In this section, the effects of the input current on arc heater flow are investigated. The input current is defined to vary from 100 to 900 A. The length and the diameter of the constrictor are 60.0 cm and 2.0 cm, respectively. The diameter of the nozzle throat is 1.5 cm.



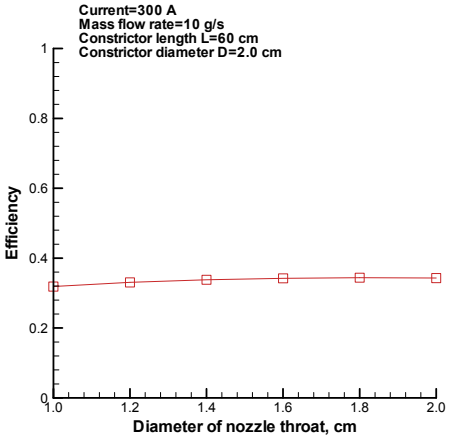
(a) Voltage & Power



(b) Mass-Averaged Enthalpy



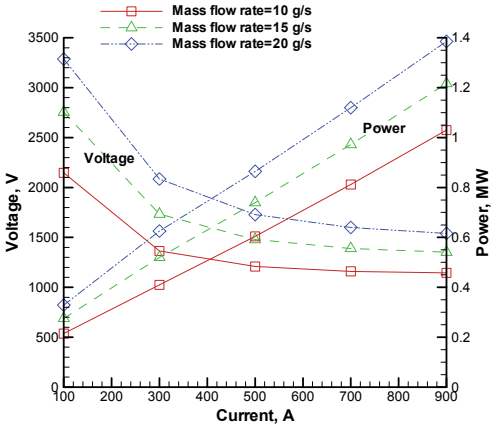
(c) Pressure



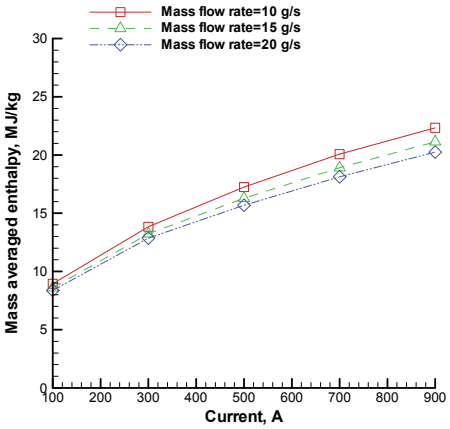
(d) Efficiency

Fig. 14. Operational Data (Lee & Kim, 2010)

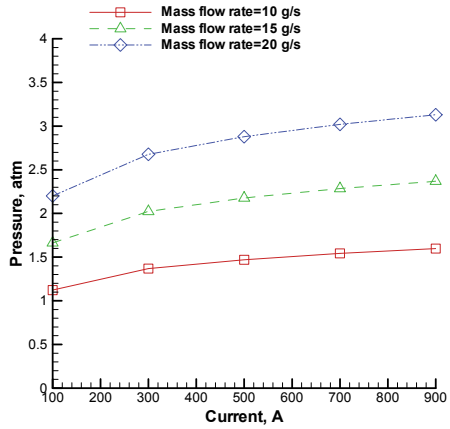
Figure 15 shows operational data in terms of input current at the following mass flow rates: 10, 15, and 20 g/s. Figure 15a shows that the electric power is almost proportional to the input current, while the voltage decreases as the input current increases. The reason is that constrictor length dominantly determines the voltage value. Accordingly, the mass-averaged enthalpy and pressure increase under the condition of constant mass flow rate, as shown in Figs. 15b and c. Efficiency decreases as the input current increases.



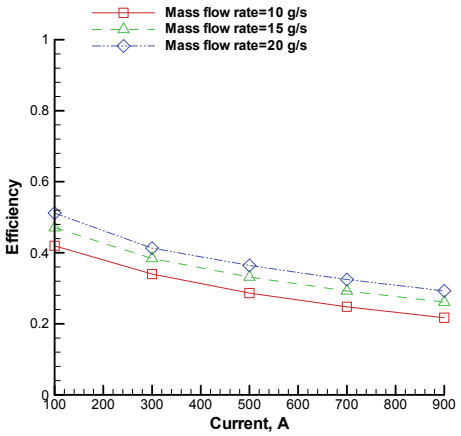
(a) Voltage & Power



(b) Mass-Averaged Enthalpy



(c) Pressure



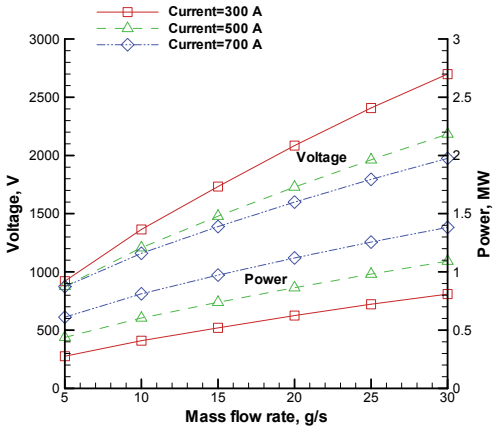
(d) Efficiency

Fig. 15. Operational Data (Lee & Kim, 2010)

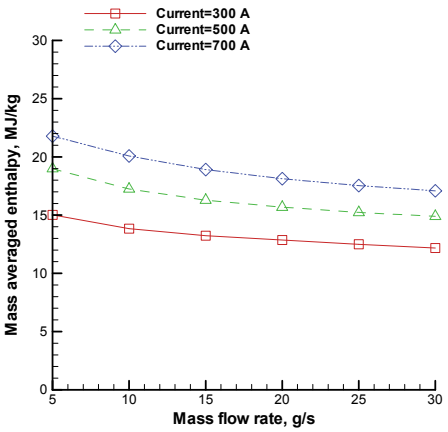
Efficiency is strongly related to temperature distribution. As the input current increases, the core temperature increases and the arc column broadens. Generally, if the current increases, the temperature increases due to high Joule heating. On the other hand, strong radiation prohibits the core temperature from increasing. Instead, it makes the temperature distribution to be flat at the core region and arc column broader, which leads to enhanced radiation throughout the wall. Also, the temperature gradient near the wall increases, which increases the heat energy loss by thermal conduction. As a consequence, efficiency decreases due to high heat energy loss caused by radiation and thermal conduction.

4.1.5 Mass flow rate

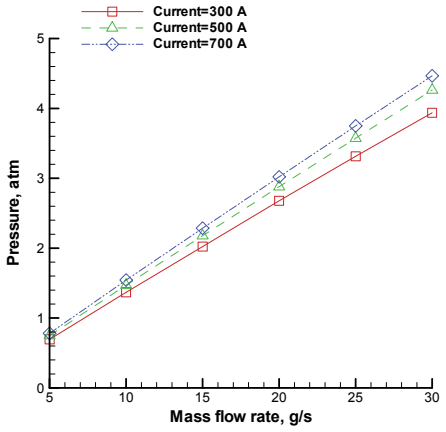
A parametric study according to a mass flow rate is performed. The mass flow rate changes from 5 to 30 g/s. The length and diameter of the constrictor are 60.0 cm and 2.0 cm, respectively. The diameter of the nozzle throat is 1.5 cm. Figure 16 shows operational data in terms of the mass flow rate for three input currents: 300, 500, and 700 A. Figure 16a shows that the voltage and the electric power increase as the mass flow rate increases. This is



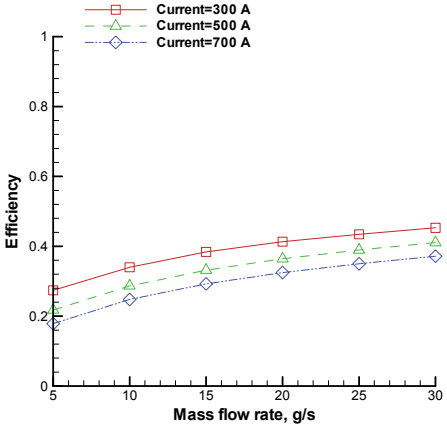
(a) Voltage & Power



(b) Mass-Averaged Enthalpy



(c) Pressure



(d) Efficiency

Fig. 16. Operational Data (Lee & Kim, 2010)

because that the ionization rate decreases due to the high mass flow rate inside the arc column. As shown in Fig. 16b, the mass-averaged enthalpy decreases as the mass flow rate increases. On the other hand, in Fig. 16c, the cathode chamber pressure increases as the mass flow rate increases. In addition, efficiency increases as the mass flow rate increases. Figure 17 shows the temperature distribution along the radial direction at the middle cross section, which is 30 cm from the constrictor starting point. As shown in the figure, the core temperature decreases and the arc column becomes narrower as the mass flow rate increases. This temperature distribution makes the voltage higher and reduces the energy loss due to radiation, as shown in Fig. 18. On the other hand, viscous dissipation by turbulence occurs noticeably near the wall due to the high mass flow rate and high axial speed of the working gas. Hence, both the temperature gradient near the wall and the heat energy loss due to thermal conduction increase, as shown in the Fig. 18. Consequently, the more the mass flow rate increases, the higher the total energy loss becomes. However, as a result, the efficiency increases slowly along with the increase in mass flow rate because the total heat energy loss increases more slowly than the total electric power input.The mass flow rate also has an influence on the stability of arc discharge. Figure 17 shows that the arc column becomes narrower with increased mass flow rate. That is, the stability of the arc discharge becomes worse as the mass flow rate increases. In a real design process, the maximum value of the mass flow rate should be determined considering the stability of the arc discharge for a given arc heater configuration.

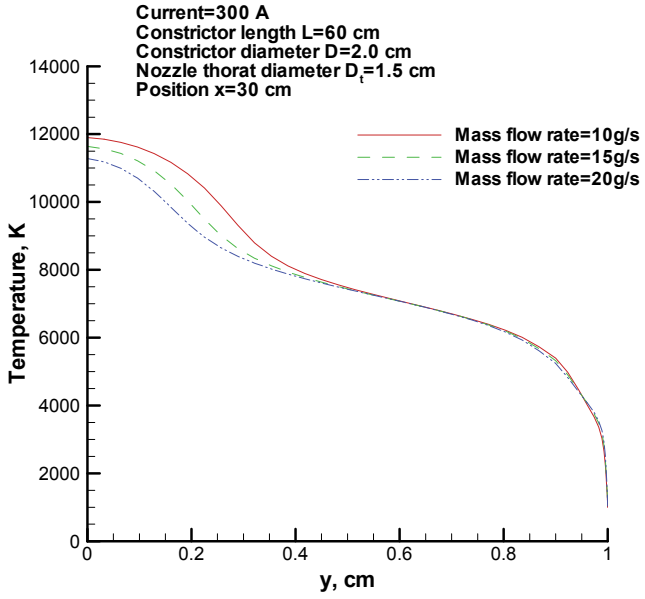


Fig. 17. Temperature (Lee & Kim, 2010)

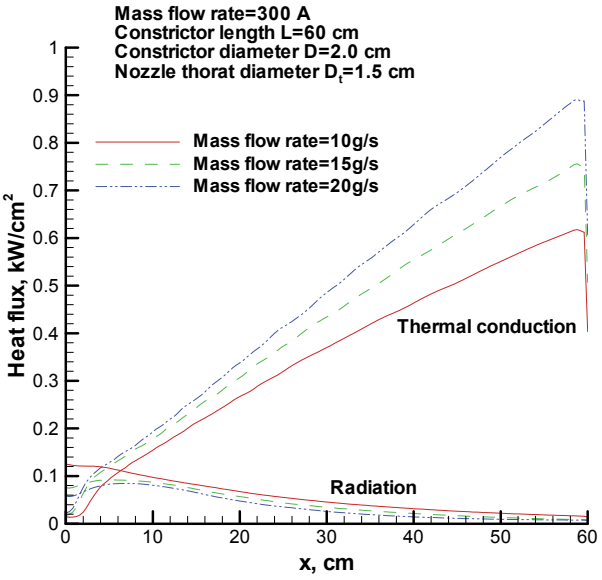


Fig. 18. Heat Flux (Lee & Kim, 2010)

4.2 Design of a segmented arc heater

In Section 3, the numerical code, ARCFLO4, is validated using the real operating data of the arc heater. In Section 4.1, a parametric study is performed for the various design parameters of the arc heater. Since the accuracy of the numerical results is rigorously proven, the database obtained by the parametric study is quite reliable. Therefore, it is expected that the database can be used in arc heater design processes. If target parameters such as total pressure and total enthalpy are given, the configuration and operational conditions such as the size of the constrictor and the nozzle throat, and the range of the input current and the mass flow rate, can be directly determined through the database based on the parametric study. Moreover, in the design of a cooling system, the database can be effectively used since the CFD code predicts the wall heat flux value quite well, i.e., if the heat flux value on the arc heater is known, the size of the cooling system and pipe configuration can be determined directly.

5. Conclusion

The accuracy level of current CFD analysis on arc heater flows is introduced. It is shown that current state-of-the art CFD technologies can predict the plasma flow inside the arc heater well. Both the high input power cases (60MW, 20MW) and the low input power cases (750kW, 150kW) are validated successfully using the ARCFLO4 computational code. The k-ε turbulence model combined with the 3-band radiation model provides good solutions for arc heater flows. Moreover, the possibility of the present computational code as a design tool for arc heater is introduced. A parametric study is performed to investigate the relation between arc heater performance and the design parameters. In the case of constrictor length,

as the constrictor length increases, the voltage and electrical power increase while the efficiency decreases. It is also shown that the voltage, the mass-averaged enthalpy, and the efficiency are strongly affected by the constrictor diameter. The mass-averaged enthalpy seems to be affected more by the constrictor diameter than by the constrictor length. From the view point of arc stability, as the L/D ratio increases, the arc column broadens, which means that the stability of the arc improves. Based on a parametric study of the nozzle throat diameter, it is determined that the nozzle throat diameter strongly affects the pressure. The effects of input operational conditions such as input current and mass flow rate are also discussed. It appears that the electric power increases as the input current and the mass flow rate increase. Moreover, arc stability becomes worse as the mass flow rate increases or the input current decreases. It appears that if the configuration of the arc heater is known, the minimum value of input current and maximum value of the mass flow rate can be determined using the numerical parametric study results. Therefore, it is expected that the ARCFLO4 code could play an important role in the design process of arc heater.

6. Acknowledgement

The authors would like to specially thank Chul Park for a technical guidance at the Korea Advanced Institute of Science and Technology. The authors would like to also thank Takehara Sakai for an offer of his three-band radiation code at the Nagoya University.

7. References

Gupta, R. N., Yos, J. M., Thompson, R. A., & Lee, K. P. (1990). A Review of Reaction Rates and Thermodynamic and Transport Properties for an 11-Species Air Model for Chemical and Thermal Nonequilibrium Calculations to 30000 K, NASA RP-1232, August 1990.

Han, S. H., Byeon, J. Y., Lee, J. I, & Kim, K. H. (2011). Numerical analysis of a 150kW enhanced Huels type arc heater, 49th AIAA aerospace science meeting, Orlando, January 2011.

Hightower, T. M., Balboni, J. A., MacDonald, C. L., Anderson, K. F., & Martinez, E. R. (2002). Enthalpy by Energy Balance for Aerodynamic Heating Facility at NASA Ames Research Center Arc Jet Complex, 48th International Instrumentation Symposium, the Instrumentation, Systems, and Automation Society, Research Triangle Park, NC, May 2002.

Jameson, A. & Yoon, S. (1987). Lower-Upper Implicit Schemes with Multiple Grids for the Euler Equations, AIAA Journal, Vol. 25, No. 7, 1987, 929-935.

Jones, W. P., & Launder, B. E. (1972). The Prediction of Laminarization with a Two-Equation Model of Turbulence, International Journal of Heat and Mass Transfer, vol. 15, No. 2, 1972, 301–314.

Kim, K. H. & Kim, C. (2005). Accurate, efficient and monotonic numerical methods for multi-dimensional compressible flows Part II: Multi-dimensional limiting process, Journal of Computational Physics, Vol. 208, No. 2, September 2005.

Kim, J. G., Oh, J. K., & Park, C. (2006). A High Temperature Elastic Collision Model for DSMC Based on Collision Integrals, AIAA paper 2006-3803, June 2006.

Kim, K. H., Rho, O. H., & Park, C. (2000). Navier-Stokes Computation of Flows in Arc Heaters, Journal of Thermophysics and Heat Transfer, Vol. 14, No. 2, 2000, 250-258.

Kim, K. H., Kim, C., & Rho, O. H. (2001). Methods for the Accurate Computations of Hypersonic Flows: I. AUSMPW+ Scheme, Journal of Computational Physics, Vol. 174, No.1, November 2001, 38-80.

Lee, J. I., Kim, C., & Kim, K. H. (2007). Accurate Computations of Arc-heater Flows Using Two-equation Turbulence Models, Journal of Thermophysics and Heat Transfer, Vol. 21. No. 1, 2007, 67-76.

Lee, J. I., Han, S. H., Kim, C., & Kim, K. H. (2008). Analysis of Segmented Arc-heater Flows with High Argon Concentration, Journal of Thermophysics and Heat Transfer, Vol. 22, No. 2, 2008, 187-200.

Lee, J. I. & Kim, K. H. (2010). Numerical Parameter Study of Low Electric Power Segmented Arc Heaters, AIAA 2010-230, January 2010.

Matsuzaki, T., Ishida, K., Watanabe, Y., Miho, K., Itagaki, H., & Yoshinaka, T. (2002). Construction and Characteristics of the 750 kW Arc Heated Wind Tunnel, Rept. TM-760, October 2002.

McBride, B. J., Zehe, M. J., & Gordon, S. (2002). NASA Glenn Coefficients for Calculating Thermodynamic Properties of Individual Species, NASA TP 2002-211556, September 2002.

Menter, F. R. (1994). Two-Equation Eddy Viscosity Turbulence Models for Engineering Applications, AIAA Journal, Vol. 32 No. 8, Nov. 1994, 1598-1605.

Nicolet, W. E., Shepard, C. E., C. E., Clark, K. J. Balakrishnan, A., Kesselring, J. P., Suchsland, K. E., & Reese, J. J. (1975). Analytical and Design Study for a High-Pressure, High-Enthalpy Constricted Arc Heater, AEDC-TR-75-47, July 1975.

Park, C. (2001). Chemical-Kinetic Parameters of Hyperbolic Earth Entry, Journal of Thermophysics and Heat Transfer, Vol. 15, No. 1, 2001, 76-90.

Russo, G. (1993). The Scirocco Wind Tunnel Project - Progress report 1993, AIAA Paper 1993-5117, November 1993.

Sakai, T. & Olejniczak, J. (2001). Navier-Stokes Computations for Arcjet Flows, AIAA Paper 2001-3014, June 2001.

Sakai, T. & Olejniczak, J. (2003). Improvement in a Navier-Stokes Code for Arc Heater Flows, AIAA Paper 2003-3782, June 2003.

Sakai, T. (2007). Computational Simulation of High-Enthalpy Arc Heater Flows, Journal of Thermophysics and Heat Transfer, Vol. 21, No. 1, 2007, 77-85.

Smith, R., Wagner, D. A., & Cunningham, J. W. (1996). Experiments with a Dual Electrode Plasma Arc Facility at the Deutsche Forschungsanstalt fur Luft-und Raumfahrt E.V. (DLR), AIAA Paper 96-2211, June 1996.

Terrazas-Salinas, I. & Cornelison, C. (1999). Test Planning Guide for ASF Facilities, A029-9701-XM3 Rev. B, March 1999.

Watson, V. R. & Pegot, E. B. (1967). Numerical Calculations for the Characteristics of a Gas Flowing Axially Through a Constrictor Arc, NASA TN D-4042, June 1967.

Wilcox, D. C. (1998). Turbulence Modeling for CFD, 2nd ed., DCW Industries, La Canada, CA, 1998, 119–122.

Whiting, E. E., Park, C., Liu, Y., Arnold, J. O., & Paterson, J. A. (1996). NEQAIR96, Non-equilibrium and Equilibrium Radiative Transport and Spectra Program: User Manual, NASA Reference Publication 1389, December 1996.

Yos, J. M. (1963). Transport Properties of Nitrogen, Hydrogen, Oxygen, and Air to 30,000 K, RAD-TM-63-7, March 1963.

5

Physico - Chemical Modelling in Nonequilibrium Hypersonic Flow Around Blunt Bodies

Ghislain Tchuen[1] and Yves Burtschell[2]
[1]*Institut Universitaire de Technologie F-V, LISIE, Université de Dschang, Bandjoun - Cameroun*
[2]*Ecole Polytechnique Universitaire de Marseille, Université de Provence, DME, 5 rue Enrico Fermi Technopole de Chateau Gombert, Marseille, France France*

1. Introduction

The development of new space transportation vehicle requires better knowledge of hypersonic flow around blunt bodies and an accurate prediction of thermal protection system for extremely high temperatures. The complex domain of this hypersonic research program concerns the fully understanding and the control of reentry flowfield. The vehicle flying with high velocity through the upper layers of the atmosphere with low density. A very strong bow shock wave around vehicle is generated and converted the high kinetic energy into internal energy, thus increasing the temperature of the gas. Therefore, shock layer is the site of intensive physico-chemical nonequilibrium processes such as vibrational excitation, dissociation, electronic excitation, even the ionization and radiation phenomena. Under this typical hypersonic condition, air must be considered as a plasma around the vehicle which perturbes traditionally the communication between the vehicle and ground control station because the plasma absorbs radio waves. The computation of such flowfield is a challenging task.

The successful conception of such high technology would not have been possible without some knowledge of these thermochemical nonequilibrium phenomena and how they affect the performance of the vehicle. Some of these informations can either be obtained from experimental facilities such as wind tunnel and ballistic range, or large scale fight experiments, and/or numerical simulations. Moreover, small scale laboratory experiments are severely limited by impossible exact simulation of thermo-chemical nonequilibrium flow around a full scale hypersonic vehicle, and flight experiments are too costly to allow their widespread usage. Therefore, much of these aerothermodynamics informations needed to design future hypersonic vehicle will have to come from numerical predictions (the least expensive approach) which is a reasonable alternative after sufficient validations.

The numerical simulation of hypersonic flow in thermochemical nonequilibrium past a blunt body presents considerable difficulties for accurate solutions in the stagnation region. The computational results depend on the choice of the thermochemical model and the strategy of resolution. Generally, efforts provided to solve these types of flows have been based on the full coupling between Navier-Stokes equations and the thermochemical phenomena. Many

researchers have developed different thermal and chemical models for the description of hypersonic flowfield with the same experimental configurations. Therefore, it is important to determine an adequate model for accurate description of hypersonic flowfield.

Some of the largest uncertainties in the modeling of reacting hypersonic flow are the chemical reaction rates and the coupling between thermochemical phenomena. The uncertainties about the thermochemical processes render the calculations doubtful. Whereas methods for analyzing the aerodynamics in equilibrium flow have achieved a level of maturity, uncertainties remain in their nonequilibrium counterparts due to the incomplete modeling of chemical processes. Consequently, a good knowledge of the chemical modeling is required. For example, several chemical kinetics model give only the forward reaction rates. Many options are available for the calculation of the backward reaction rate with the equilibrium constant (Park[1], Gupta[2], Gibb[3]) and lead to different results. The method of computation of the backward reaction rate affects flowfield structure, shock shapes, and vehicle surface properties. It is necessary therefore, to make a judicious choice of an adequate model through a comparative study.

In the literature, a multitude of models for chemical kinetics of air exist. These models are built on different simplifying assumptions, and have all advantages and disadvantages depending on the problem simulated. The objective of the present study is to investigate results obtained with four different models of chemical kinetic. Solutions from models proposed by Gardiner[4], Moss[5], Dunn and Kang[6], and Park[7] are compared. Particular attention has been devoted to the way in which the backward reactions have been obtained. Gupta[2] high temperature least-squares equilibrium constant curves fits are also included. The influence of the formulations of Hansen[8], and Park[1] for the coupling between a molecule's vibrational state and its dissociation rate are compared. Several studies were presented in the past on the dissociation of nitrogen or of oxygen separately[9; 10]. The extension of these works to the complexity of overall reactions of air remains questionable. The present chapter attempt to identify the model with an acceptable confidence for a wide range of Mach number. The gas was chemically composed either by seven species ($O, N, NO, O_2, N_2, NO^+, e^-$) with 24 step chemical reactions or by 17 reactions involving five species (O, N, NO, O_2, N_2) and, or by nitrogen dissociated partially (N_2, N). One approach to validate the thermochemical model in CFD codes is to compare the shock standoff distance and the stagnation heating point along a sphere, with the experimental data.

Moreover, it has been shown that the description of the flow with a one temperature model leads to a substantial overestimation of the rate of equilibration when compared with the existing experimental data [1]. Much work on nonequilibrium flow are based on a model with two [11] or three [12] temperatures. For two temperature model, vibrational and electronic mode of molecules are described by a single temperature. This assumption is made to simplify the calculation. For the model at three temperature, it is assumed that a single temperature control the translational-rotational modes, a second temperature for vibrational mode for all molecules, and the third temperature for electronic-free electron modes. This was a resonable assumption if the vibrational-vibrational coupling between the various molecular species is very strong. It is well known that the vibrational and electronic temperatures play important role in a high temperature gas because they improve the definition and evaluation of the physical properties of nonequilibrium hypersonic flow. In the present chapter, four-temperature model is used with two vibrational temperatures and the numerical results obtained for RAM C flight have well been compared with experimental data.

The results present in this chapter were obtained using an improved version of the time marching Navier-Stokes code CARBUR, originally developed at IUSTI Marseille. The code has been extensively tested in the past[13–16], and it's used here for the solutions of the stagnation-region flowfield. The scheme is based on a multiblock finite volume technique. The convective numerical flux is calculated by upwind technology with Riemann's solvers algorithms. The second-order central differences are used to discretize the viscous fluxes. An accurate second order algorithm in space and time is obtained by employing the MUSCL approach in conjunction with the Minmod limiter and the time predictor-corrector schemes. The source terms are treated implicitly to relax the stiffness. The steady state is obtained after convergence of the unsteady formulation of the discretized equations. We have included the recent definition and improvement in physical modelling. Special attention will be given to treatment of chemical phenomena that take place during reentry phase, in order to complete some description and modelisation of thermochemical nonequilibrium flow around atmospheric reentry vehicle.

2. Nomenclature

$A_{i,j}$	area of the cell (i,j)
A_r	constant for evaluating foward reaction rate coefficient $K_{f,r}$
$C^s_{v,q}$	specific heat at constant volume for species s for energy mode q, (where $q \equiv$ Translation, rotation, vibration, electronic)
$C^s_{p,q}$	specific heat at constant pressure for species s for energy mode q
D_s	diffusion coefficient of species s
$\rho e, e$	total energy per unit volume, mass
$\rho e_e, e_{e_s}$	electron-electronic energy per unit volume, mass of species s
$\rho e_{v_m}, e_{v_m}$	vibrational energy per unit volume, mass of molecules m
$H_{i,j}$	axisymmetric source term
$K_{eq,r}$	equilibrium constant for reaction r
$K_{f,r}, K_{b,r}$	forward and backward reaction-rate coefficient for reaction r
L_e	Lewis constant number, $(= \rho C_p D/\lambda)$
M_s	molecular weight of species s
NM	total number of molecules
N_k	outward normal vector on each side of the cell
NR	total number of reactions
NS	total number of species
NSV	total number of molecules in vibrational nonequilibruim
p_s	pressure of species s
Q_{T-e}	translation-electronic energy transfer rate
Q_{T-v_m}	translation-vibration energy transfer rate
Q_{v-v_m}	vibration-vibration energy transfer rate
$r_{i,j}$	radius of the cell-center position
T	Translational-rotational temperature
T_a	geometrically averaged temperature $(= T^q T_v^{1-q}$ or $= T^q T_e^{1-q})$
T_e	electron-electronic excitation temperature
T_{v_m}	vibrational temperature of molecule m
t	time

U	vector of conserved quantities
u, v	velocity in x and y directions
Y_s	mass fraction of species s $(= \frac{\rho_i}{\rho})$
Greek symbols	
λ_{tr}	translational thermal conductivity of mixture
$\lambda_{v,m}$	vibrational thermal conductivity of molecule m
λ_{el}	electronic thermal conductivity of mixture
ω_s	mass production rate for species s
$\alpha_{s,r}$	stoichiometric coefficient for reactant s in reaction r
$\beta_{s,r}$	stochiometric coefficient for product s in reaction r
γ	ratio of specific heat $(\gamma = C_p/C_v)$
$\Omega_{i,j}$	source term
θ_r	characteristic temperature of reaction r
$\theta_{v,m}$	characteristic temperature of vibration
τ_m	average vibrational relaxation time of molecule m
$\tau_{m,s}^{VT}$	vibrational relaxation time for collision pair $m - s$
ρ	total density $(= \Sigma_s \rho s)$
ρ_s	density of species s
Subscripts	
eq	equilibrium
m	molecule
s	species
w	wall
∞	freestream

3. Analysis

The governing equations that describe the weakly ionized, thermo-chemical nonequilibruim flow have been developed by Lee [17]. In this work, the following assumptions are introduced: 1) The flow regime is continuum. 2) The energy level of each mode are populated following boltzmann distribution with a characteristic temperature. 3) The rotational mode energy of molecules is fully equilibrated with the translational mode of heavy particles, and therefore translational and rotational temperatures of molecules are equal. 4) The harmonic oscillator model is employed for the vibrational energy. 5) The gas in the shock layer does not emit nor absorb radiation. 6) When ionization is taken into account, absence of the conduction current is assumed and an induced electric field is built up by charge separation [18], the magnitude of this field is predicted to be: $E_i \cong -\frac{1}{N_e e}\frac{\partial p_e}{\partial x_i}$.

The full laminar Navier-Stokes equations for two-dimensional conservation equations are written as:

The mass conservation equation for each species, s,

$$\frac{\partial \rho_s}{\partial t} + \frac{\partial \rho_s u_j}{\partial x_j} + \frac{\partial \rho_s V_s^j}{\partial x_j} = \omega_s \tag{1}$$

The momentum conservation equation in x and y directions,

$$\frac{\partial \rho u_i}{\partial t} + \frac{\partial (\rho u_i u_j + p\delta_{ij})}{\partial x_j} + \frac{\partial \tau_{ij}}{\partial x_j} = \sum_s \epsilon N_s Z_s E_i \tag{2}$$

The total energy equation,

$$\frac{\partial \rho e}{\partial t} + \frac{\partial ((\rho e + p)u_j)}{\partial x_j} + \frac{\partial u_i \tau_{ij} + q_{t_j} + q_{v_m j} + q_{e_j} + \sum_s^{ns} (\rho_s h_s V_s^j)}{\partial x_j} = \sum_s \epsilon N_s Z_s E_i u_i \tag{3}$$

The conservation equation of vibrational energy for each nonequilibrium molecule,

$$\frac{\partial \rho e_{v_m}}{\partial t} + \frac{\partial (\rho e_{v_m} u_j)}{\partial x_j} + \frac{\partial (q_{v_j} + \rho_m e_{v_m} V_m^j)}{\partial x_j} = Q_{T-v_m} + Q_{v_m - v_r} + Q_{v_m - e} \tag{4}$$

When the electronic relaxation is accounted, the electron-electronic energy conservation equation

$$\frac{\partial \rho e_e}{\partial t} + \frac{\partial ((\rho e_e + p_e)u_j)}{\partial x_j} + \frac{\partial (q_{e_j} + \sum_s \rho_s e_{e_s} V_s^j)}{\partial x_j} = u_j \frac{\partial p_e}{\partial x_j} - Q_{v_m - e} + Q_{T-e} + Q_{el} \tag{5}$$

In these equations, the electric field due to the presence of electrons in flow is expressed as:

$$\overrightarrow{E} \simeq -\frac{1}{N_e \epsilon} \overrightarrow{\nabla} p_e \tag{6}$$

The mixture is assumed to be electrically neutral ($\sum_s \epsilon N_s Z_s E_i \simeq 0$) as a consequence of the chemical kinetic mechanism; for each ion produced/consumed in the flow an electron is also produced/consumed. The local charge neutrality is also assumed. Thus, the number of electrons is equal to number of ions at each point:

$$\rho_e = \hat{M}_e \sum_{s=ions} \frac{\rho_s}{\hat{M}_s} \tag{7}$$

The state equation of the gas allows to close the system of equations (1-5). The total pressure is given as sum of partial pressures of each species regarded as perfect gas.

$$p = \sum_{s=1}^{NS} p_s = \sum_{s \neq e} \rho_s \mathcal{R}_s T + \rho_e \mathcal{R}_e T_e \tag{8}$$

The total energy of the mixture per unit volume

$$\rho e = \sum_{s \neq e} \rho_s C_{v,tr}^s T + \frac{1}{2} \sum_s \rho_s u_s^2 + \sum_{m=1}^{NM} \rho_m e_{v_m} + \rho e_e + \sum_{s=1}^{NS} \rho_s h_s^0 \tag{9}$$

is splitted between the translational-rotational, kinetic, vibrational, electron-electronic contributions, and the latent chemical energy of the species. T and T_e are deduced through the equation (5) and (9) with an iterative method. The vibrational temperature of the diatomic

species m is determined by inverting the expression for the energy contained in a harmonic oscillator at temperature T_{v_m}:

$$e_{v_m} = \frac{\mathcal{R}}{\hat{M}_m} \frac{\theta_{v,m}}{e^{\theta_{v,m}/T_{v_m}} - 1} \tag{10}$$

The speed of sound plays a major role in flux-split algorithm. It's evaluation in the case of one translational temperature is no longer applicable in the case of multiple translation temperature. The correction of the speed of sound due to electronic contribution and the presence of electron translational temperature has been included[19]:

$$a^2 = \gamma \left(\frac{p}{\rho}\right) + (\gamma - 1)\left(\frac{T}{T_e} - 1\right)\frac{p_e}{\rho} \tag{11}$$

where classical speed of sound is obtained when $T = T_e$.

3.1 Transport coefficients

The transport coefficient modelling can have a considerable quantitative influence on practically relevant quantities such as skin friction and heat flux at the vehicle walls. Accurate measurements of transport coefficients at the high temperatures of interest for hypersonic applications are very difficult to realize and there is accordingly severe dearth of reliable experimental data for these thermophysical properties. Several useful simplifications have been used but their relevance and reliability, other than for quick estimates of order of magnitude, become rather questionable under increasingly advanced demands on accuracy. Some of the formulas in use become less reliable at high temperatures where ionization becomes important. In this study, the viscous stresses τ_{ij} are defined with the hypothesis of Stokes. The dynamic viscosity is given by Blottner[20] interpolation law. The thermal conductivity of each species is derived from Eucken's[3] relation. The Wilke's semi-empirical mixing rule[21] is used to calculate total viscosity and conductivity of the gas. For simplicity, the mass diffusion fluxes for neutral species are given by Fick's law with a single diffusion coefficient [19]. The diffusion of ions is modeled with ambipolar diffusion coefficient $D_{ion}^{ambi} = 2D_s$. To improve this formula, we used $D_{ion}^{ambi} = D_{ion}(1 + T_e/T)$ as recommended in[22]. The effective diffusion coefficient of the electrons (D_e) is proportional to the ambipolar diffusion coefficient of the ions [19].

The total heat flux is assumed to be given by the Fourier's law as:

$$Q = \sum_s q_s = -\sum_s \left(\lambda_{tr,s}\nabla T + \lambda_{v,s}\nabla T_{v,s} + \lambda_{el,s}\nabla T_e + \rho D_s h_s \nabla Y_s\right) \tag{12}$$

which is the resultant of the flux of conduction, vibration, electronics and the diffusion of the total energy. After using an extension form of Masson and Monchick assumptions[23], and the relation given by Athye[24] to connect thermal conductivity of vibration with the diffusion coefficient, a more convenient form of the total heat flux is obtained as:

$$\overrightarrow{Q} = -\frac{\lambda_{tr}^{'}}{C_{pf}^{'}}\left[\nabla h + (L_e^{'} - 1)\left(\sum_{s=1}^{NS} h_s \nabla Y_s + \sum_{s=1}^{NSV} Y_s \nabla e_{v,s} + \nabla e_e\right)\right] \tag{13}$$

Where the Lewis number $L_e^{'} = \frac{\rho D C_{pf}^{'}}{\lambda_{tr}^{'}}$ represents the ratio of the parts of the heat flux due to the energy transport by the diffusion gas mixture components and by heat conduction which depends on the translational temperature, and $\lambda_{tr}^{'} = \lambda_{tr} + \sum_{I=eq} \lambda_{v,I}$ (I: molecules in equilibrium); and the similar expression for $C_{pf}^{'} = C_{p_{tr}} + \sum_{I=eq} Y_I C_{v_{vib,I}}$.

3.2 Energy exchange model

The energy exchange between translation and vibrational mode Q_{T-v} is described according to Landau-Teller theory[1]

$$Q_{T-v_m} = \rho_m \frac{e_{v_m}(T) - e_{v_m}(T_{v_m})}{\tau_m} \tag{14}$$

and τ_m is the relaxation time expressed as in reference[19].
The vibrational energy transfer between the different molecules is modelled by Candler[25]:

$$Q_{v-v_m} = \sum_{s \neq m} P_{sm} Z_{sm} \frac{\hat{M}_s}{\mathcal{N}} [e_{vs}(T_{vsm}) - e_{vs}] \tag{15}$$

where T_{vsm} is the same vibrational temperature obtained after the collision of the two molecules, Z_{sm} is the $s-m$ collision number per unit volume which is determined from kinetic theory[3], P_{m-s} and P_{s-m} are the two probabilities originating from the work of Taylor et al.[26]. These probabilities have been presented recently in an exponential form by Park and Lee[7].
When electron-electronic energy is taken into account, the expression of the energy exchanged during electron-heavy particles collisions is derived from Lee[17]

$$Q_{T-e} = 3\mathcal{R}\rho_e(T - T_e)\sqrt{\frac{8\mathcal{R}T_e}{\pi M_e}} \sum_{r \neq e} \frac{\rho_r N}{\hat{M}_r^2} \sigma_{er} \tag{16}$$

where σ_{er} are the collision cross sections for interaction electron-other particle. The value is assumed to be constant and equal to $10^{-20} m^2$.
Q_{e-v_m} is the energy source term from vibration-electron coupling. It is assumed that only $N_2 - e$ coupling is strong[17], and its expression is assumed to be of the Landau-Teller form :

$$Q_{e-v_m} = \rho_m \frac{\hat{M}_m}{\hat{M}_e} \frac{e_{v_m}(T_e) - e_{v_m}(T_{v_m})}{\tau_{em}}; \quad m = N_2 \tag{17}$$

where the relaxation time τ_{em} is a function of electron-electronic temperature and electron pressure as presented in[25]. The term Q_{el} accounts for the rate of electron energy loss when a free electron strikes a neutral particle and ionizes it, with a loss in electron translational energy.

3.3 Chemical processes

The accurate characterization of the shock layer in hypersonic flow requires the good knowledge of the species mass, produced by chemical reactions which take place according to a suitable chemical kinetic model. An instantaneous accurate prediction of the mass fraction and the heating rate require a correct simulation of the chemical behavior of the flow field. The

mechanism by which the considered chemical reactions appear is little known. Several ways lead to intermediate products or highly unstable excited states, which are available virtually for each reaction. It is difficult to produce theoretical formulations that involve all species, and to conduct experimental work for the kinetic data that shows the evolution of chemical reactions. Solutions from Dunn and Kang[6], the Gardiner[4], Moss[5] and Park[7] reaction rate sets are compared. All these models are different essentially in the data of the forward and backward coefficients of reactions rate. In Dunn and Kang, the ratio of rate is not necessarily equal to the equilibrium constant as required by Eq.19. A modified Dunn and Kang model, was created when the backward rate is computed either with exact or curve fit equilibrium constants. Three different models of chemical reactions are simulated in this study.
- The first model applied to 5 species (O, N, NO, O_2, N_2) with 17 elementary reactions are grouped into 15 reactions of dissociation and 2 exchange reactions.
• Dissociation

$$\mathrm{O_2 + M \rightleftharpoons O + O + M}$$
$$\mathrm{N_2 + M \rightleftharpoons N + N + M}$$
$$\mathrm{NO + M \rightleftharpoons N + O + M}$$

• Exchange or Zel'dovich reaction

$$\mathrm{O + N_2 \rightleftharpoons NO + N}$$
$$\mathrm{NO + O \rightleftharpoons O_2 + N}$$

M represents the collision partner which is one of the 5 species of the mixture.
- The second model is applied to 7 air species ($O, N, NO, O_2, N_2, NO^+, e^-$) with 24 elementary reactions. In this case, the reaction of ionization is added to the model above

$$\mathrm{N + O \rightleftharpoons} NO^+ + e^-$$

- The last model is applied to the dissociation of nitrogen gas (N_2, N).

The increase or decrease of the species concentration due to chemical reactions is given by source terms ω_s as follows:

$$\omega_s = \hat{M}_s \sum_{r=1}^{NR} (\beta_{s,r} - \alpha_{s,r}). \left[K_{f,r} \prod_{s=1}^{NS} (\rho_s / \hat{M}_s)^{\alpha_{s,r}} - K_{b,r} \prod_{s=1}^{NS} (\rho_i / \hat{M}_s)^{\beta_{s,r}} \right]$$

where the associated forward and reverse reaction rate coefficients are assumed to satisfy the generalized Arrhenius law:

$$K_{f,r} = A_r T^{\eta_r} e^{-\theta_r / T} \tag{18}$$

Backward rates $K_{b,r}$ are defined either with curve fit constants in the Arrhenius expression, or with the equilibrium constant $K_{eq,r}$:

$$K_{eq,r} = \frac{K_{f,r}}{K_{b,r}} = \frac{\prod_{i=1}^{NS} (\rho_i / \hat{M}_i)^{\alpha_{i,r}}}{\prod_{i=1}^{NS} (\rho_i / \hat{M}_i)^{\beta_{i,r}}} \tag{19}$$

$K_{eq,r}$ can be determined analytically with the minimization of Gibbs free energy of each species under thermodynamics assumptions [3]. Moreover, for thermo-chemical nonequilibrium,

the theoretical formulation becomes poorly defined due to the complexity of the physical phenomenon. However, the fitting of experimental data would be more reliable than the theoretical formulation, because every simplification introduce in the thermodynamics model affects results.
According to Park, the equilibrium constant $K_{eq,r}$ is given by the following exponential polynomial

$$K_{eq,r} = exp(A_{1,r} + A_{2,r}z + A_{3,r}z^2 + A_{4,r}z^3 + A_{5,r}z^4), \quad z = 10000/T \tag{20}$$

This Park curve fits were performed from data points only at 2 000, 4 000, 6 000, 8 000 and 10 000K. Gupta[2] has examined the curve fits for equilibrium constants to temperatures above 30 000K. He demonstrates that the five parameters used in both the Park reaction rate set and those derived from the Dunn and Kang[2] rates are incorrect at temperatures above 10 000K. Gupta least squares curve fitted, with a six parameters function for equilibrium constant is of the form

$$K_{eq,r} = exp(a_{1,r}z^5 + a_{2,r}z^4 + a_{3,r}z^3 + a_{4,r}z^2 + a_{5,r}z + a_{6,r}) \tag{21}$$

The forward reactions rate of the dissociation of the nitrogen molecule are represented in Fig.1 and it is observed that the Dunn and Kang kinetic is largely seperated from the others. The causes of uncertainties on the determination of the reactions rate coefficients are multiple, and is still rather poorly known. The present work examines various options for calculating the backward rates for several chemical kinetic models. The vibration - dissociation coupling is also considered in this chapter.

3.4 Vibration-dissociation coupling

The vibration - dissociation coupling is very important behind a strong shock in thermo-chemical nonequilibriun flow. Directly behind the shock, the translational temperature reaches a maximum value while the vibrational temperature takes a time to be excited before reaching its equilibrium value. Therefore, a model of dissociation depending only on the translational temperature will tend to overestimate total dissociation. The most important and poorly understood issue is how to model the coupling between a molecule's vibrational state and its dissociation rate. Many analyses of this coupling have been made in the past, either with more or less realistic and sophisticated physico-chemical models, or with semi-empirical methods easily usable in hypersonic computation codes[1]. There are perhaps 10 such models available in literature, however only the Park and Hansen models are used in this study.
Chemical reaction rates are affected by the extent to which the internal modes of atoms and molecules are excited. The coupling factor translating the influence of the vibration on dissociation is given by the ratio:

$$Z(T, T_v) = \frac{K_f(T, T_v)}{K_f(T)} \tag{22}$$

There are several methods for including thermal non-equilibrium effects in the chemical kinetic of air. Some have a semi-empirical origin, based on experimental results and

are often extrapolated well beyond their validity domains (without physical justification). Whereas those with analytical origin, arise from particle collision theory. Park[7] suggested a geometrically average rate controlling temperature T_a for dissociation reactions. T_a is defined as :

$$T_a = T^q T_{v_m}^{1-q} \qquad and \qquad K_{f,r}(T, T_{v_m}) = A_r T_a^{\eta_r} e^{-\theta_r / T_a} \tag{23}$$

with $q = 0.5$ in the original formulation. Sharma et al.[27] suggested that the value of q varied from 0.6 to 0.7 might be more realistic for high enthalpy flow and consequently, a value of 0.7 is used in the present study. Sharma's modification of the model is based on the analysis of experimental data for the reaction velocity observed in shock tubes. Hassen used collisional cross section theory to compute dissociation reaction rates up to 40,000K for flow in thermal nonequilibrium. He suggested an alternative formulation for the rate controlling temperature, similar to the Park approach

$$q = \alpha - \beta \left(\frac{T_v}{T} \right) \tag{24}$$

with α and β values reported in table 1.
The various values of the coupling factor for O_2 and N_2 at T=20,000K are represented in Fig.2. Losev's[28] coupling is also included. As evident in Fig. 2, below 15,000K the magnitude of the coupling is very sensitive to the coupling model used.

4. Numerical methods

4.1 Numerical procedure

The governing equations (1-5) are discretized in space by using a finite-volume approach with a central formulation over structured mesh. The inviscid fluxes at cell interfaces are computed using AUFS[13] method. The second-order spatial accuracy is obtained by employing the monotonic upstream schemes for conservation laws (MUSCL-Hancock)[29] with a TVD extension type scheme approach. The limiter function Mimmod is used for the inviscid fluxes. The viscous terms are classically discretized by second-order central difference approximation.
The explicit formulation which gives the variation of $U_{i,j}$ during time Δt on each cell (i, j) can be written in two dimensional axisymmetric coordinate as:

$$\frac{\Delta U_{i,j}}{\Delta t} + \frac{1}{A_{i,j} r_{i,j}^{\alpha}} \sum_{k=1}^{4} F_k N_k = \alpha H_{i,j} + \Omega_{i,j} \tag{25}$$

where $\alpha = 1$ for an axisymmetry coordinate system, and $\alpha = 0$ for planar two-dimensions. The source terms Ω is treated implicitly to relax the stiffness. A time predictor-corrector algorithm is used to obtain second-order time accuracy.
The structure of this code is multi-blocks and uses parallel processing machine architecture for significant enhancement of efficiency in treating of complex flow configuration. A detailed description of this numerical method may be found in [19].

4.2 Boundary conditions

The freestream conditions is hypersonic. The outflow is supersonic and the zero gradient exit condition is appropriate. Along the stagnation line, the flowfield is symmetric. The wall temperature (T_w) is fixed and zero normal pressure gradient is imposed. The no-slip and

no-temperature jump conditions are used. At the wall, the flow is generally assumed to be in thermal equilibrium. Because of the shielding effect, we have assumed either $T = T_{v_s} = T_e$ when vibrational and electronic energy content is greater than the electron translational energy content or $\partial T_{v_s}/\partial n = \partial T_e/\partial n = 0$ otherwise. The wall is supposed to be chemically noncatalytic. The upstream flow conditions, the number of nodes and the minimum grid spacing in x and y directions considered in this study are reported on Table 2. The distribution of points in the x and y directions of the computational domain is stretched exponentially in the region where there are strong shock interactions, while being clustered more near the surface to ensure y^+ for the first grid point less than 1.2. The time step is computed locally in each cell from a given CFL number. In each case, the free-stream air/nitrogen gas is composed of 79 % N_2 and 21 % O_2 or 92.7 % N_2 and 7.3 % N.

4.3 Condition of numerical simulation

The computational results are validated with experimental observations reported by Lobb[30], Rose et al.[31] and Hornung[32] for the shock standoff distance and surface heat transfer distributions. Results of the current numerical simulations are validated with the shock tunnel data presented in Table 2.

The first case study involves Lobb's experiment[30] devoted to the measure of shock detachment at 5.28 km/s. The second computation is also validated with the shock tube measurements of stagnation point heat flux for blunt bodies at 5.74 km/s from the work of Rose et al.[31]. The RAM-CII flight is a program conducted by the NASA Langley Research Center to assess the effects of black-out during re-entry. Computations were performed for RAM-CII conditions at Mach 25.9 for comparison with available numerical simulations. The case $M_\infty = 23.9$ is also included. This case has a relatively high density conditions of the flow field which leads to a high chemical reactivity. For the last case that involved hypersonic nitrogen flow over cylinders, the focus was to reconstruct the interferogram experiments performed by Hornung[32]. The results have been found to compare favourably with experimental results obtained with a free-piston driven shock tube wind tunnel which was used to obtained interferograms of the flow of pure nitrogen over blunt-nosed bodies. At the end, an extension is made for study the influence of the choice of chemical model for a hypersonic flow around a double-wedge.

5. Results and discussions

5.1 flow past blunt bodies

A grid resolution study was performed in the former work[14] to determine the number of points required for different cases. The current numerical algorithm was used to achieve grid independent solutions and for a flow in thermochemical nonequilibrium. The same CFD code and mesh points are used to examine the influence of all the current chemical kinetic models selected on flowfield structures, shock shapes, and vehicle surface properties. The results obtained are presented and compared with experimental data and numerical results.

5.1.1 Lobb's experiment

The predicted temperatures distribution along the stagnation line of T and $T_{v_{N_2}}$ for five chemical models are shown in Fig.3. The computations are done with Park curve fit equilibrium constants in conjunction with Park CVD coupling. The maximum value of T

and $T_{v_{N_2}}$ behind the shock wave is obtained with the slow kinetic models. It is clearly seem that the vibrational relaxation time for molecular nitrogen in the post shock region is greatest with the slow kinetic models and lowest with the fast kinetic models, caused by higher density in the shock layer which enhances the collisional energy transfer mechanisms. It is observed that all temperatures tend towards a thermodynamic equilibrium near the stagnation point. Note that the features of the flow physics is evident by the existence of a significant thermal nonequilibrium behing the shock wave and, the vibrational temperature remains below the associated translational temperature whatever the choice of chemical kinetic model. The original Dunn and Kang model predicts the lowest temperatures which become as high as the Gardiner prediction when the backward rates are calculated with the Park equilibrium constant.

The predicted pressure distributions along the stagnation line are plotted in Fig.4. It is noted that all chemical models converge to the same peak value of pressure at the stagnation point, and that the difference between the Hansen and Park CVD coupling cannot be discriminated. These distributions are used to identify the shock location. One can see that the shock standoff distance tends to increase when the kinetic velocity is decreased, approaching the frozen flow. The shock standoff distance is the smallest for the case of fast kinetic models, suggesting the higher degree of dissociation compared to other kinetics. The peak temperature, however, is highest for Gardiner model and lowest for Dunn and Kang model, as expected. A comparison with the location given by Lobb[30] is reported in Table 3. The values reported for the shock standoff distance correspond to the point of the maximum pressure gradient associated with the thickness of the mesh. With respect to the shock standoff distance observed in the Lobb experiment (0.087 R_s with a +/- 5.75% experimental uncertainty), the calculated standoff value predicted using Park's (93) data is 4.36% smaller, while, the value predicted with the Gardiner data is 7.46% greater. The result produced by the Park(93) model is closest to the experimental value.

The evolutions of the calculated wall heat flux are distinct and presented in Fig.5 as a function of the angle θ is affected. The fast kinetic models produces similar profiles. The same observation is made with the slow kinetic models. At the stagnation point, the current chemical models produce stagnation heating values that range from 23 to 38 MW/m^2. The calculation of the backward reactions with the analytic equilibrium constant tends to increase the wall heat flux with Park's (93) model while Gardiner's model tends to decrease it about 13 MW/m^2 at the stagnation point. The numerical results obtained by other authors are about $25MW/m^2$[33; 34]. This case demonstrates that Park(93) model with Park curve fit constant has a better prediction of shock stand-off distance and wall heat flux value.

5.1.2 Rose and Stankevics's experiment

Calculations for the second case focus on the Mach 18 air experimental case of Rose. Computed temperature distributions obtained with the Park(93) model along the stagnation line are shown in Fig.6. The backward reaction rates are evaluated either with analytic or Park and/or Gupta equilibrium constants. In the Park(93) model, the manner in which the backward reaction rates are calculated do not affect profiles. The nonequilibrium behavior of the vibrational energy mode can be seen from the temperature profiles near the shock region. It can be seen from the variation of the vibrational temperatures in the vicinity of the shock wave that the vibrational relaxation time for O_2 is the lowest. This figure clearly shows that the influences of the CVD coupling of Park and Hansen are not visible in the Park(93) model. The

evolution of temperatures obtained with the Gardiner model presented in Fig.7 are affected by the method of computation of the backward reactions. Note that the same observation applies to the Mach 15.3 case where the temperatures obtained with the slow kinetic models behind the shock wave are high.
It is important to determine the correct dissociation rate in hypersonic flow, because the degree of dissociation significantly affects the heat transfer, when the surface is catalytic with recombination of atoms (as is the case for most thermal protection materials). Figure 8 shows the heat transfer Q_w along the body surface for four chemical kinetic models. As expected, the higher values of Q_w are obtained with slower kinetics. The stagnation point heat transfer are reported in Table 4 and compared with experimental data[31]. Computed stagnation heat flux is over predicted with Gardiner, Moss and modified Dunn and Kang models. In the attempt to improve Dun and Kang model, a modified Dunn and Kang was obtained with the using of equilibrium constants in the calculation of the backward reactions. The model of Park (93) gives better agreement with the experimental measurement.
When weakly ionization and electron-electronic energy are taken into account in conjunction with chemical model of Park(93), the predicted temperature distribution along the stagnation line is shown in Fig.9. The electronic relaxation and the effects of the modified speed of sound are examined here. The electron-elctronic temperature is lower with the correction. This translates that the distribution of the total energy on various internal modes has been a more accurately accounted form in results. The flux sppliting procedure and the physics of the flow take into account this correction which allows to approach correctly the local speed of sound in the presence of the electron translational temperature.

5.1.3 RAM-CII experiment

There are no experimental measurements of heat flux in this case, the interest of this study is to submit the chemical models to very high temperatures and to compare numerical results. The graphs presented here are obtained with the Gupta equilibrium constant whose field of validity is wider.
In Fig.10 the translational temperature is around 20 000K and the extrapolation of certain formulations becomes erroneous. For example at this temperature, it is necessary to account for ionization although the chemistry of Moss and Gardiner are conceived only for neutral air gas. There is a significant decrease in the temperature predicted by the modified Dunn and Kang model. These differences indicate that the Dunn and Kang rate set is not consistent with Gupta's equilibrium constants at the higher temperatures associated with Mach 25.9. This conclusion is easily verified for each reaction with a direct comparison of the equilibrium constants to the ratio of rates. Inconsistency with Gupta's equilibrium constants does not assure that, Dunn and Kang rates will not produce correct species distributions. This inconsistency, however, raises doubt as to recommending this rate set as the correct data for these flows.
The surface heat transfer Q_w is shown in Fig.11 for four chemical kinetics models. The difference of peak between calculations with Park CVD coupling and Hansen at the stagnation point is about 2,5%. The maximum value of the heat flux is located at the position $\theta = 0^o$. The computed stagnation point heat flux results are compared with other numerical results in Table 5. The large dispersion of the results obtained puts into doubt the accuracy of certain reactions rates, or indicates that some other unknown thermo-chemical process occurs in these flows. The numerical results reported in Table 5 are obtained with a non-catalytic wall. The

numerical methods and the selected chemical model are different. Walpot[36] and Soubrie et al.[37] have used a model with 11-species while a model with 7-species has been considered by others [14; 38]. The flux value for Candler and MacCormack given in Table 5 is estimated from the Stanton number. It is noted that the values obtained with the slow kinetic models are higher. This was encountered in all the cases that were studied.
An extension is made to the case $M_\infty = 23.9$ with a four temperature model and weakly ionized gas considered under Park(93) chemical kinetic. Figure 12 shows the predicted temperature distribution obtained along the stagnation line at Mach 23.9. The comparative evolution of the various temperatures justifies the choice of the model at four temperatures. The difference between $T_{v_{O_2}}$ and $T_{v_{N_2}}$ becomes larger for a strong Mach number in a flow at relatively low density condition. The use of a model at one temperature of vibration can induce miscalculations in the kinetic of the reactions, which will influence the calculation of the thermodynamic parameters. Figure 13 shows the comparison of the computed and the measured peak electron number density along the body for Mach 23.9 and 25.9. The comparison shows very good agreement, with a scatter from the experimental values. Most of the ionization occurs at stagnation region. The electron number density is maximal in the stagnation point of the obstacle and decreases along the body.

5.1.4 Hornung's experiment

Calculations involving hypervelocity nitrogen flow over cylinders are performed under conditions reported in table 2. The results obtained demonstrate the sensitivity of flowfield to the effect of finite rate chemistry. The free-stream gas is composed by nitrogen dissociated partially ($Y_{N_2} = 0.927$, $Y_N = 0.073$). The wall is assumed to be adiabatic and non-catalytic. This case implementes the experiment performed by Hornung[32] for 1 inch diameter cylinders with the dissociation of nitrogen at 5.59 km/s. The calculations use three different sets of dissociation rates of N_2.
The calculations using the corresponding free-stream conditions have been made for the body and figures 14- 16 show interferograms and computed fringe patterns derived from the simulation results. The exact position of the shock in both the experiments and the simulations is uncertain : in the experiments because of the shock curvature in the transverse direction and in the simulations because the shock is smeared over a few cell widths. To compare the experiments and the simulations more closely it is necessary to compare fringe shifts at different locations in the flow. Figure 14 compares the computed interferograms obtained with Dunn and Kang and the Modified Dunn and Kang. In the modified model, the backward rates are calculed with the Gupta curve fit equilibrium constant. The two solutions are quite different. The distance from the detachment of the shock is sensitive to the chemical model selected. Although the Dunn and Kang model is very poor, the modified version improves the thickness of the shock layer.
The fringes for the case with Park(93) and Gardiner with analytic equilibrium constant are plotted in Figs. 15 and 16. The frames are quantitatively well predicted. The experimental and computational interferograms have the same general features, but differ little in the details. The best comparison is obtained with the Park (93) model.

5.2 Interaction of type IVr around a double-wedge in hypersonic flow

Shock/shock interactions are common in many aerodynamic configurations such as those encountered during hypersonic flight, around launchers or in propulsion systems. Such shock

	α	β
N_2	0.90	0.30
O_2	0.85	0.25
NO	0.80	0.30

Table 1. CVD coefficients for Hassen's model

Experiments	Lobb[30]	Rose[31]	Ram C[13]	RamC[13]	Hornung[32]	Double-Wedge
Gas	Air	Air	Air	Air	N_2,N	Air
$R_s,(mm)$	6.35	6.6	152.4	152.4	0.5 inch	500
M_∞	15.3	18	25.9	23.9	6.14	9
$U_\infty,(m/s)$	5263	5742.49	7658.6	7636.4	5590	3130.137
$P_\infty,(P_a)$	664	432.2	4.764	19.85	2910	10^5
$T_\infty,(K)$	293.0	252	216	254	1833	300
$T_{wall},(K)$	1000	555.5	1500	1500	adiabatic	adiabatic
R_{e_∞}	14605	12840	6280	19,500	6000	10^5
IM nodes	50	60	70	60	50	400
JM nodes	50	60	70	70	50	500
$\Delta x_{min}(m)$	4.1936x10^{-6}	1.423x10^{-6}	2.002x10^{-5}	2.1171x10^{-5}	4.214x10^{-5}	4,1122x10^{-6}
$\Delta y_{min}(m)$	7.2055x10^{-5}	4.6432x10^{-5}	1.27x10^{-4}	2.7318x10^{-4}	1.072x10^{-4}	2,520x10^{-6}

Table 2. Detail of free-stream conditions and grids used

interactions can cause high localized heat transfer rates as well as local oscillations of the pressure fields on the vehicle surface. These can alter the aerodynamic characteristics of the vehicle leading to catastrophic failures.

The geometrical constraints and flow field parameters are chosen such as flow generates an interaction of type IVr. After study grid independence, the number of grid points used here is 400x500. The initial conditions are: $L_1, L_2 = 0.05$ (m); $\theta_1 = 15$; $\theta_2 = 60$; $P_\infty = 10^5(P_a)$; T_∞, $T_{V_{N_2}}$, $T_{V_{O_2}} = 300$ (K);

$U_\infty = 3130.137$ (m/s); $M_\infty = 9.0$; $R_{e_\infty} = 2.0016\text{x}10^5$/m, $Y_{N_2} = 0.766$; $Y_{O_2} = 0.233$.

In this study, a value of $\gamma = 1.4$ is used to initialize calculation because the free-stream flow is assumed as a perfect gas frozen.

The choice of the chemical model (in equilibrium or in nonequilibrium) for use in the numerical simulations of flow around double-wedge can lead to different results. Although widely validated for flow around blunt bodies [15], very large uncertainties remain for the double-wedge. Figure 17 reflects the subtle differences that can be achieved between two kinetics (slow or fast) at thermal equilibrium. Gardiner and Park(93) chemical kinetics are used. A remark that emerges from the different results obtained with the models of Park and Gardiner show that the disociation rates at the equilibrium appear not to be well known for a flow around a double-wedge.

With Park Consts.	δ/R (%)
Park (93)	8.321
Gardiner	9.349
Moss	9.35
Modified Dunn and Kang	9.930
Dunn and Kang	7.650
Lobb - Experiment[30]	**8.7**$\pm$**0.5**

Table 3. Shock stand-off distance, $M_\infty = 15.3$

Chemical kinetics with K_{eq}:	Park	Gupta
Park (93)	44.95	45.5
Gardiner	47.0	47.4
Moss	48.5	49.2
Modified Dunn and Kang	49.0	50.1
Dunn and Kang		32.5
Fay-Ridell[35]		**45.0**
Experiment- Rose et al.[31]		**46.0**

Table 4. Stagantion heat flux Q_w (MW/m^2), with Park CVD coupling, Mach 18

Source	Chemical model	Q_w (MW/m^2)
Present	D & K	0.58
Present	Park (93)	1.02
Present	Gardiner	1.91
Present	D & K + Gupta	2.42
Present	Moss	2.63
Walpot[36]	Park(93)	**0.74**
Walpot[36]	D & K	**0.90**
Soubrié et al.[37]	Park(93)	**0.96**
Candler et al.[38]	Park & Wray	**2.5**

Table 5. Stagnation heat flux, RAM-CII, Mach 25.9

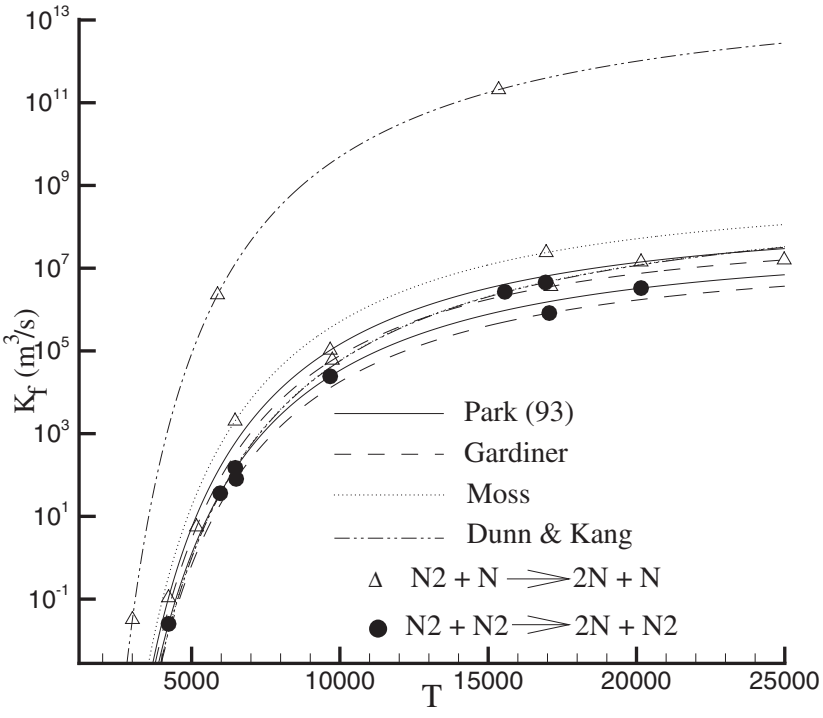


Fig. 1. Forward reactions rate

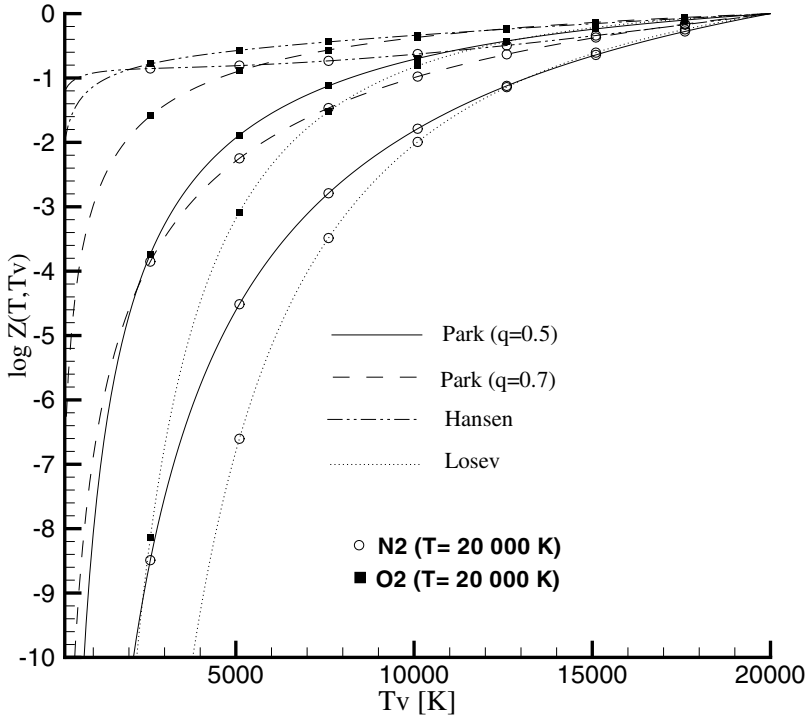


Fig. 2. Coupling factor Z(T,T_v) for O2 and N2

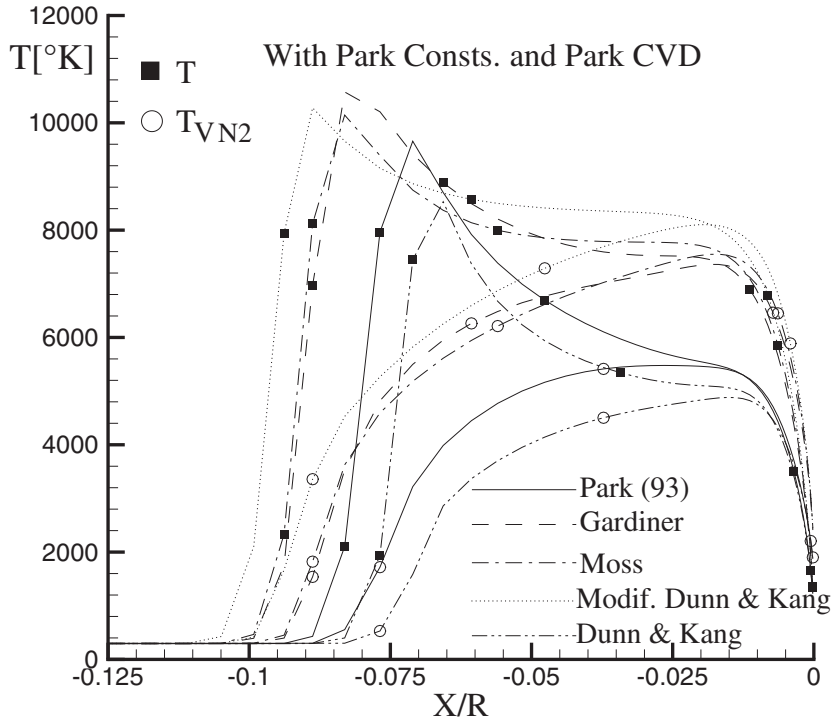


Fig. 3. Temperatures along the stagnation line, M_∞=15.3

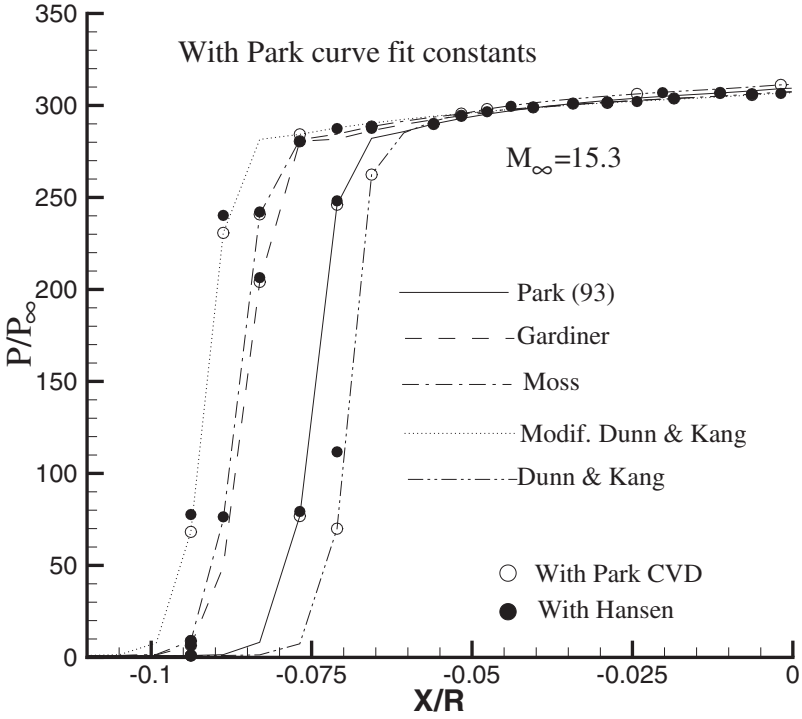


Fig. 4. Pressure distribution along the stagnation line, M_∞=15.3

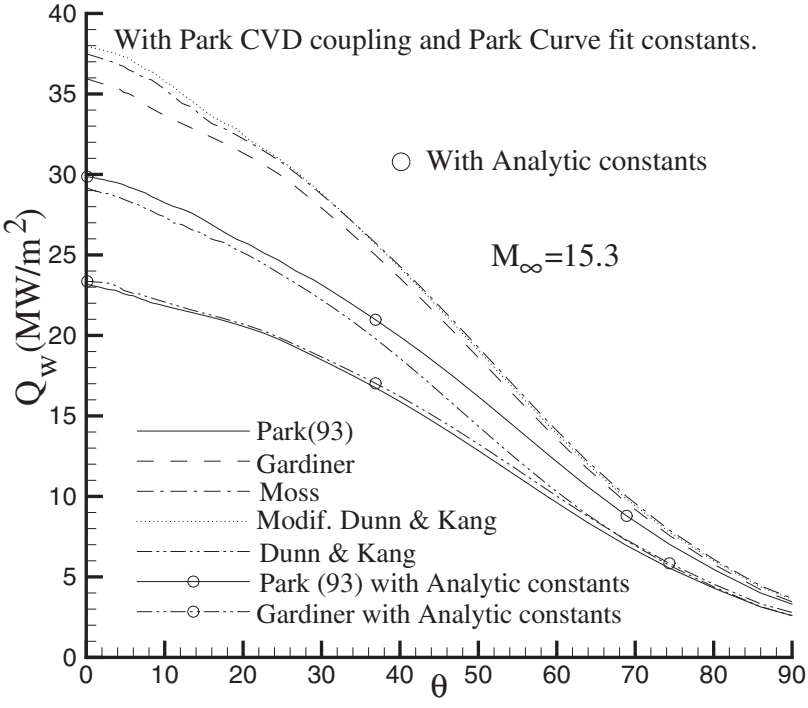


Fig. 5. Wall heat flux, M_∞=15.3

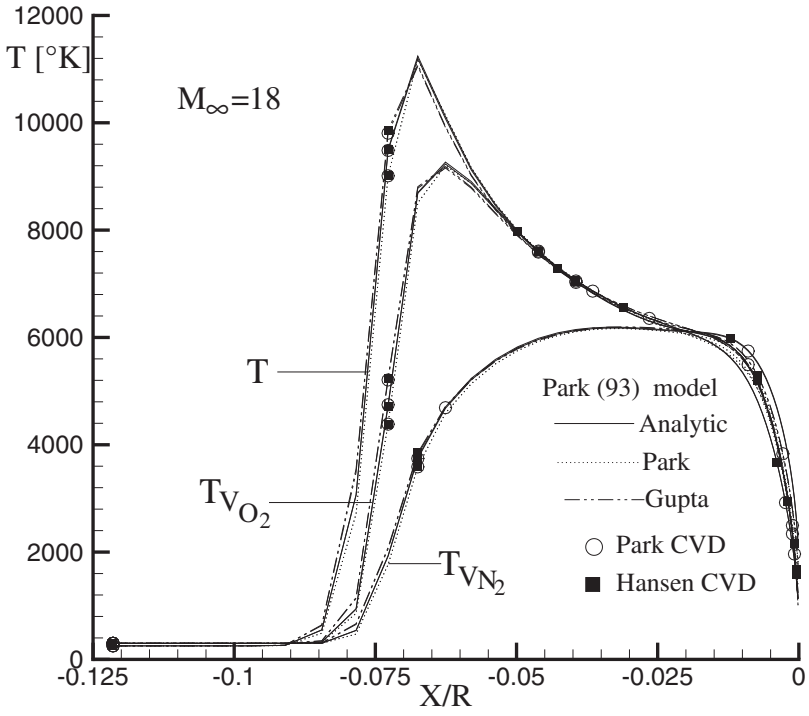


Fig. 6. Temperatures with Park(93) model, M_∞=18

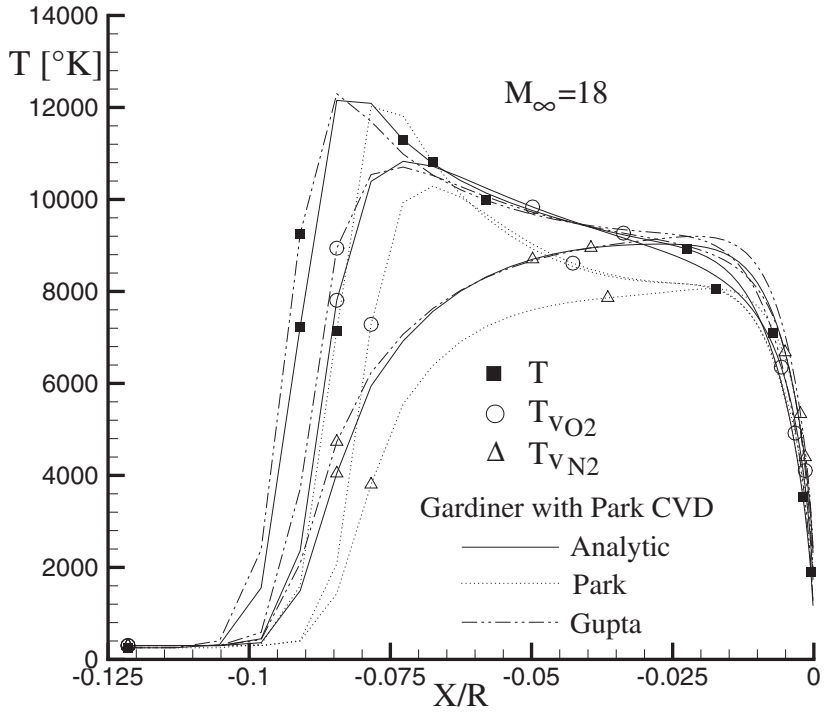


Fig. 7. Temperatures profile with Gardiner model, M_∞=18

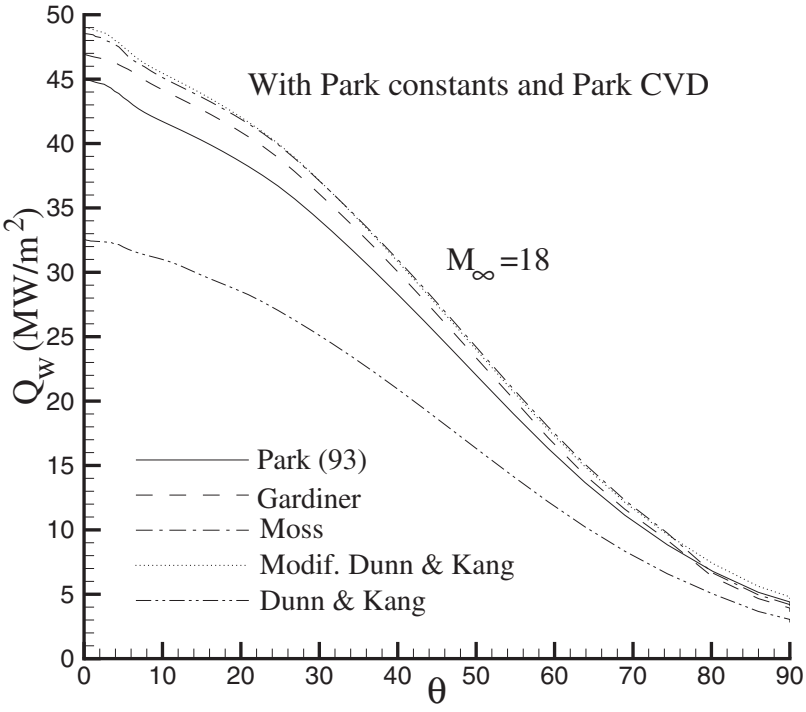


Fig. 8. Wall heat flux, M_∞=18

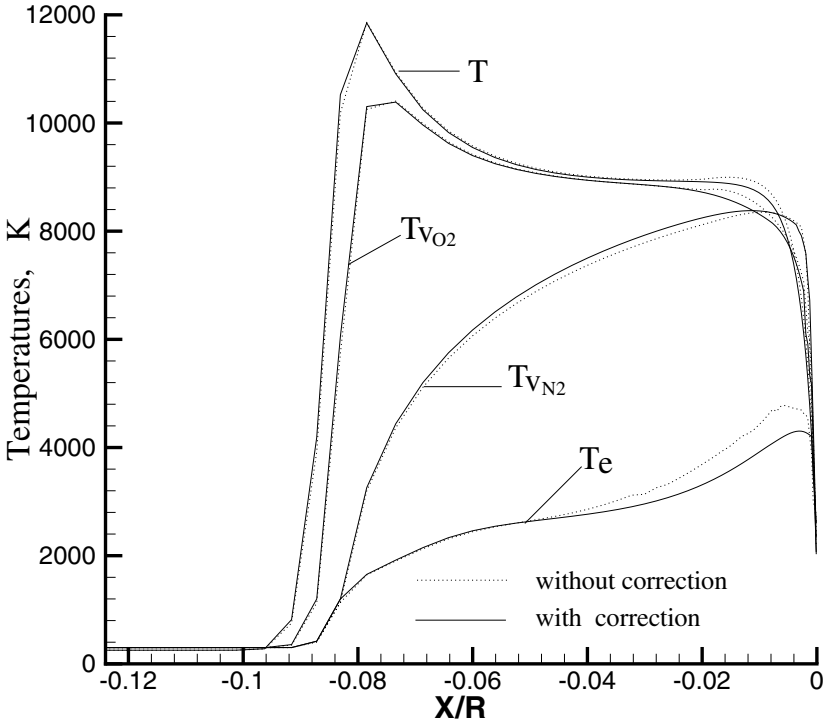


Fig. 9. Temperature along stagnation line with weakly ionized gas; $M_\infty = 18$

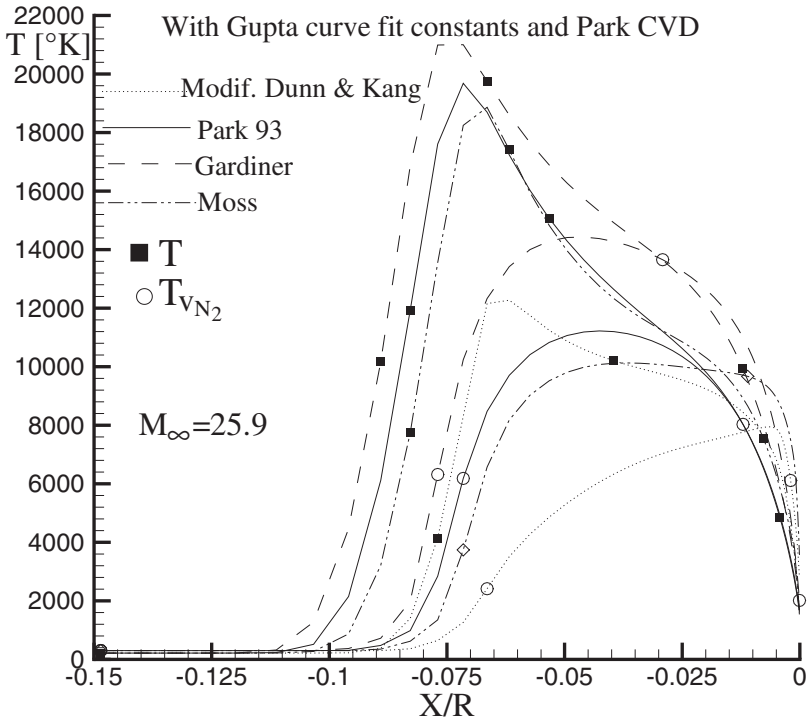


Fig. 10. Temperature profile with Gupta curve fit constants, M_∞=25.9

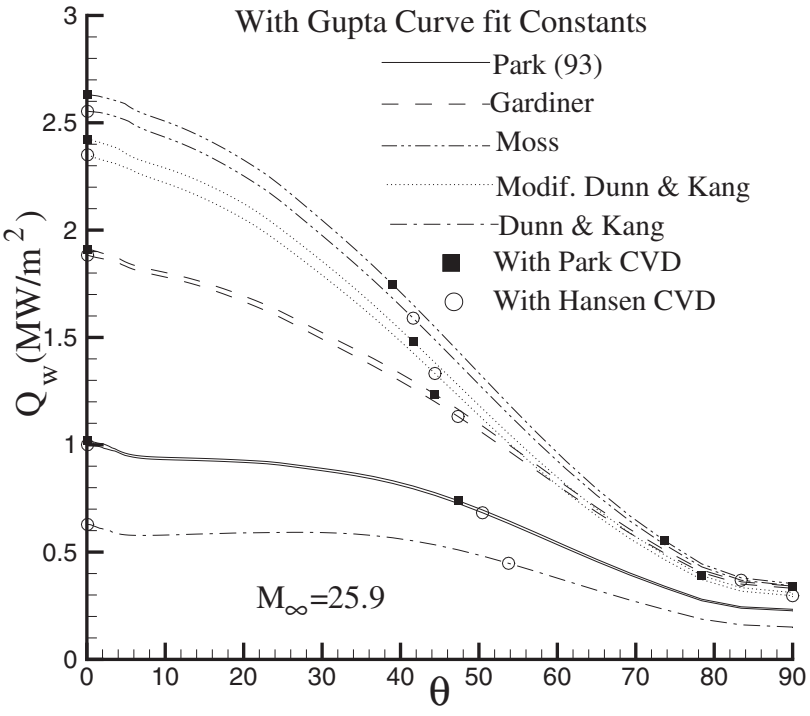


Fig. 11. Stagnation heat flux, M_∞=25.9

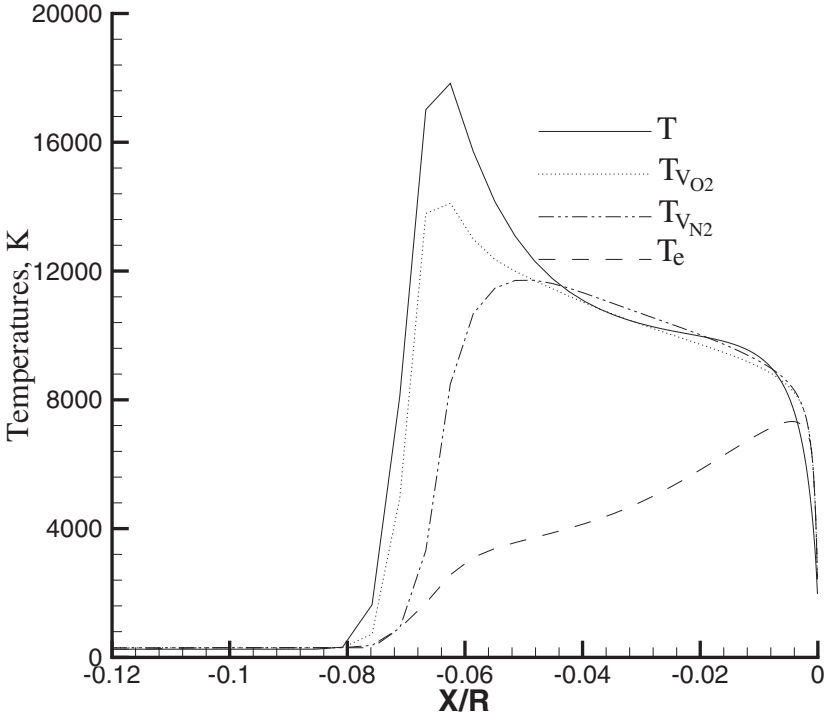


Fig. 12. Temperatures distribution along the stagnation line, M_∞=23.9

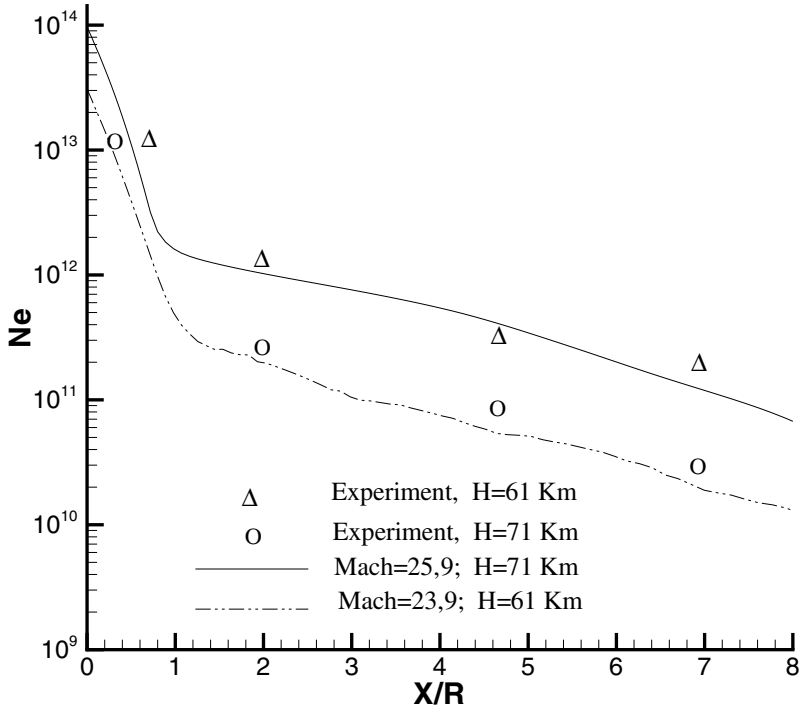


Fig. 13. Comparison with experiment of the peak of electron number density following the axial distance

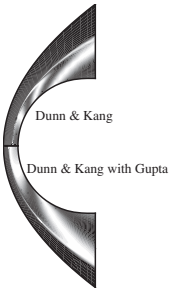


Fig. 14. Dunn and Kang, and Modified Dunn and Kang Interferograms computed

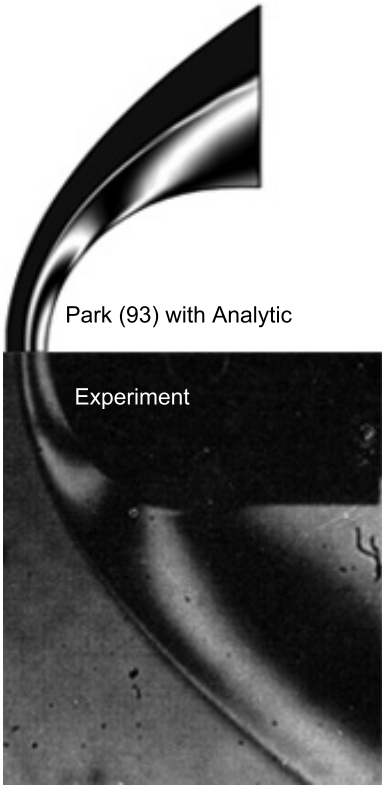


Fig. 15. Fringe patterns on 1 in diameter cylinder with Park(93) model and Exact equilibrium constant

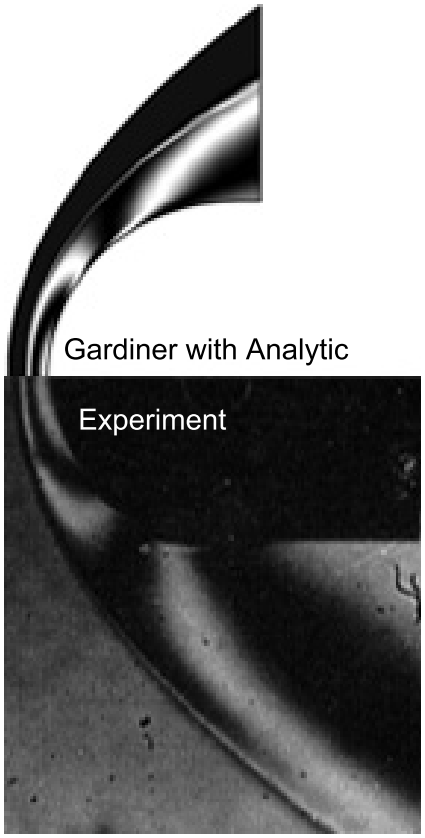


Fig. 16. Fringe patterns on 1 in diameter cylinder with Gardiner model and Exact equilibrium constant

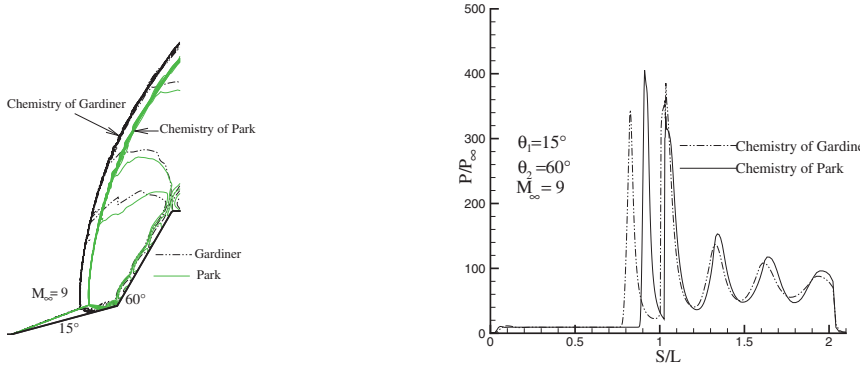


Fig. 17. Effect of chemical kinetics: a) Contours Mach number b) Surface pressure

6. Conclusion

The hypersonic flow past blunt bodies with thermo-chemical nonequilibrium were numerically simulated. The dependence of solutions on available chemical models, allowing to assess the accuracy of finite rate chemical processes has been examined. The present results were successfully validated with the theoretical and experimental work for shock-standoff distances, stagnation point heat transfer and interferograms of the flow. Although all model describe the essential aspects of the nonequilibrium zone behind the shock, they are not accurate for the evaluation of the aerothermodynamic parameters. A comparative study of various kinetic air models is carried out to identify the reliable models for applications with a wide range of Mach number.

The present study has shown that the prediction of hypersonic flowfield structures, shock shapes, and vehicle surface properties are very sensitive to the choice of the kinetic model. The large dispersion in the wall heat flux reaches 60 % as observed in the RAM-CII case. The manner in which the backward reaction rates are computed is quite important as indicated by the interferograms that were obtained. The model of Park (93) gives a better prediction of hypersonic flowfield around blunt bodies. Park(93) is identify as the model for hypersonic flow around blunt bodies with a confidence acceptable to a wide range of Mach number. There is also great sensitivity to the choice of chemical kinetics in flowfield around double-wedge.

More numerical simulations compared with experiments need to be conducted to improve the knowledge of the thermochemical model of air flow around double-wedge.

7. References

[1] Park, C. (1990). *Nonequilibrium Hypersonic Aerothermodynamics*. New York, Wiley.

[2] Gupta, R. N., Yos, J. M., Thompson, R. A., and Lee, K. P. (1990). A Review of Reaction Rates and Thermodynamic and Transport Properties for an 11-Species Air Model for Chemical and Thermal Nonequilibrium Calculations to 30 000K", *NASA RP-1232*.

[3] Vincenti, W. G., Jr.Kruger, C. H. (1965). *Introduction to physical Gas Dynamics*. Krieger, FL.

[4] Gardiner, W. C. (1984). *Combustion Chemistry*, Springer-Verlag, Berlin.

[5] Shinn, J. L., Moss, J. N., Simmonds, A. L. (1982). Viscous Shock Layer Heating Analysis for the Shuttle Winward Plane Finite Recombination Rates, AIAA 82-0842.
[6] Dunn,M. G., Kang, S. W. (1973). Theoretical and Experimental Studies of re-entry Plasma, NASA CR-2232.
[7] Park, C. (1993). Review of Chemical-Kinetic Problems of Future NASA Mission, I: Earth Entries", *Journal of Thermophysics and Heat Transfer.* Vol. 7, No. 3, pp. 385-398.
[8] Hansen, C. F. (1993). Vibrational Nonequilibrium Effect on Diatomic Dissociation Rates". *AIAA Journal*, Vol. 31, No. 11, pp. 2047-2051.
[9] Macrossan, N. M. (1990). Hypervelocity flow of dissociating nitrogen donwnstream of a blunt nose. *Journal of Fluid Mechanics*, Vol. 27, pp. 167-202.
[10] Josyula, E. (2001). Oxygen atoms effect on vibrational relaxation of nitrogen in blunt body flows. *Journal of Thermophysics and Heat Transfer*, Vol. 15, No. 1, pp. 106-115.
[11] Peter, A. G. and Roop, N. G. and Judy, L. S. (1989). Conservation Equation and Physical Models for Hypersonic Air flows in Thermal and Chemical Nonequilibrium. NASA TP 2867.
[12] Knab, O. and Fruaudf, HH. and Messerschmid, EW. (1995). Theory and validation of the the physically consistent coupled vibration-chemistry-vibration model. *J. Thermophys Heat Transfer*, nř9, Vol.2, pp.219-226.
[13] Tchuen G., Burtschell Y., and Zeitoun E. D. (2008). Computation of non-equilibrium hypersonic flow with Artificially Upstream Flux vector Splitting (AUFS) schemes. *International Journal of Computational Fluid Dynamics*, Vol. 22, Nř4, pp. 209 - 220.
[14] Tchuen G., and Zeitoun D. E. (2008). Computation of thermo-chemical nonequilibrium weakly ionized air flow over sphere cones. *International journal of heat and fluid flow*, Vol.29, Issue 5, pp.1393 - 1401.
[15] Tchuen G., and Zeitoun D. E. (2009). Effects of chemistry in nonequilibrium hypersonic flow around blunt bodies. *Journal of Thermophysics and Heat Transfer*, Vol. 23, Nř 3, pp.433-442.
[16] Burtschell Y., Tchuen G., and Zeitoun E. D. (2010). H2 injection and combustion in a Mach 5 air inlet through a Viscous Mach Interaction. *European Journal of Mechanics B/fluid*, Vol. 29, Issue 5, pp. 351-356.
[17] Lee Jong-Hun. (1985). Basic Governing Equations for the Flight Regimes of Aeroassisted Orbital Transfer Vehicles. *Thermal Design of Aeroassisted Orbital Transfer Vehicles, H. F. Nelson, ed., Volume 96 of progress in Astronautics and Aeronautics, American Inst. of Aeronautics and Astronautics*, Vol. 96, pp. 3-53.
[18] Appleton, J. P. and Bray, K. N. C. (1964). The Conservation Equations for a Nonequilibrium Plasma. *J. Fluid Mech*. Vol. 20, No. 4, pp. 659-672.
[19] Tchuen, G., Burtschell, Y., and Zeitoun, E. D., 2005. Numerical study of nonequilibrium weakly ionized air flow past blunt body. Int. J. of Numerical Methods for heat and fluid flow, 15 (6), 588 - 610.
[20] Blottner, F. G., Johnson, M., and Ellis, M., 1971. Chemically Reacting Viscous Flow Program for Multi-Component Gas Mixtures. Sandia Laboratories, Albuquerque, NM, Rept. Sc-RR-70-754.
[21] Wilke, C. R., 1950. A viscosity Equation for Gas Mixture. J. of Chem. Phys. 18 (4), 517-519.
[22] Ramsaw JD., and Chang CH., 1993. Ambipolar diffusion in two temperature multicomponent plasma. Plasma Chem Plasma process. 13 (3), 489-498.

[23] Masson, E. A., and Monchick, 1962. Heat Conductivity of Polyatomic and Polar Gases. The Journal of Chemical. 36 (6), 1622-1640.
[24] Ahtye, W. F., 1972. Thermal Conductivity in Vibrationnally Excited Gases. Journal of Chemical Physics, 57, 5542-5555.
[25] Candler, G. V., and MacCormarck, R. W., 1991. Computation of weakly ionized hypersonic flows in thermochemical nonequilibrium. Journal of Thermophysics and heat transfer, 5 (3), 266-273.
[26] Taylor R., Camac, M., and Feinberg, M., 1967. Measurement of vibration-vibration coupling in gas mixtures, In Proceeding of the 11th Intenational Symposium on combustion, Pittsburg, PA, 49-65.
[27] Sharma, S. P., Huo, W. M., and Park, C., 1988. The Rate Parameters for Coupled Vibration-Dissociation in a Generalized SSH Approximation Flows. AIAA-88-2714.
[28] Shatalov, O. P., and Losev, S. A., "Modeling of diatomic molecules dissociation under quasistationary conditions", AIAA 97-2579, 1997.
[29] Roe, P., 1983. Approximate Riemann Solvers, Parameters vectors and difference schemes. Journal of Computational Physics, Vol. 43, 357-372.
[30] Lobb, K., "Experimental measurement of shock detachment distance on sphere fired in air at hypervelocities", *in The High Temperature Aspect of Hypersonic Flow.* ed. Nelson W. C., Pergamon Press, Macmillan Co., New York, 1964.
[31] Rose, P. H., Stankevics, J. O., "Stagnation-Point Heat Transfer Measurements in Partially Ionized Air". *AIAA Journal*, Vol. 1, No. 12, 1963, pp. 2752-2763.
[32] Hornung, H. G., "Non-equilibrium dissociating nitrogen flow over spheres and circular cylinders". *Journal of Fluid Mechanics*, Vol. 53, 1972, pp. 149-176.
[33] Joly, V., Coquel, F., Marmignon, C., Aretz, W., Metz, S., and Wilhelmi, H., "Numerical modelling of heat transfer and relaxation in nonequilibrium air at hypersonic speeds", *La Recherche Aérospatiale*, Vol.3, 1994, pp. 219-234.
[34] Séror, S., Schall, E., and Zeitoun, E. D., "Comparison between coupled euler/defect boundary-layer and navier-stokes computations for nonequilibrium hypersonic flows, *Computers & Fluids*, Vol.27, 1998, pp. 381-406.
[35] Fay, J. A., Riddell, F. R., "Theory of stagnation point heat transfer in dissociated air". *J. Aero. Sciences*, Vol. 25, 1958, pp. 73-85.
[36] Walpot, L. M., "Development and Application of a Hypersonic Flow Solver". PhD thesis, TU Delft, 2002.
[37] Soubrié, T., Rouzaud, O., Zeitoun, E. D., "Computation of weakly multi-ionized gases for atmospheric entry using an extended Roe scheme". ECCOMAS, Jyväskylä, 2004.
[38] Candler, V. G., MacCormack, W. R., "The computation of hypersonic ionized flows in chemical and thermal nonequilibrium". AIAA 88-0511, 1988.

6

A Frequency-Domain Linearized Euler Model for Noise Radiation

Andrea Iob, Roberto Della Ratta Rinaldi and Renzo Arina
Politecnico di Torino, DIASP, 10129 Torino
Italy

1. Introduction

Aeroacoustics is the branch of Fluid Mechanics studying the mechanism of generation of noise by fluid motions and its propagation. The noise generation is associated with turbulent and unsteady vortical flows, including the effects of any solid boundary in the flow. Experimental studies in this field are very difficult, requiring anechoic wind tunnels and very sensitive instruments able to capture high frequency, low amplitude, pressure fluctuations. Computational Aeroacoustics (CAA) can be a powerful tool to simulate the aerodynamic noise associated to complex turbulent flow fields. As sound production represents only a very minute fraction of the energy associated to the flow motion, CAA methods for acoustic propagation have to be more accurate compared to the solution schemes normally used in Computational Fluid Dynamics (CFD).

A direct approach to aeroacoustic problems would imply to solve numerically the full Navier-Stokes equations for three-dimensional, unsteady, compressible flows. Sound would then be that part of the flow field which dominates at large distances from the region characterized by intense hydrodynamic fluctuations, propagating at the local sound speed. The direct noise simulation is hardly achievable in practice, except for very simple configurations of academic interest, and in a limited region of space. Even if the small amplitude of the fluctuations allow a linearization, the equations of acoustic disturbances on an arbitrary base flow are very complicated and their solution is not straightforward. To solve aeroacoustic problems of practical interest some simplifying approximations are necessary. One way to obtain realistic solutions is provided by the hybrid approach. It decouples the computation of the flow field from the computation of the acoustic field. When using a hybrid method the aeroacoustic problem is solved in two steps: in the first step, the hydrodynamic flow field is solved using a CFD method, then the noise sources are identified and the acoustic field is obtained. Extraction of noise sources from the fluid dynamic field can be done using an aeroacoustic theory such as the Lighthill's analogy [Crighton (1975); Goldstein (1976)]. A hybrid approach is based on the fundamental assumption that there is a one-way coupling of mean flow and sound, i.e., the unsteady mean flow generates sound and modifies its propagation, but sound waves do not affect the mean flow in any significant way. This assumption is not so restrictive, because acoustic feedback is possible only when the mechanical energy in the unsteady mean flow is weak enough to be influenced by acoustic disturbances. This occurs principally in the vicinity of a starting point for flow instability (for instance, upstream edges of cavities or initial areas of shear layers). Since the fluid-dynamic

field and the acoustic field are computed separately, numerical accuracy for the mean flow simulations used as an input of hybrid methods is less critical than in direct computation. Simpler, more flexible and lower-resolution schemes are applicable provided that numerical dissipation is carefully controlled to prevent the artificial damping of high-frequency source components. Incompressible flow solutions can be adequate for evaluating acoustic source terms based on the low Mach numbers approximation. Time-accurate turbulence simulation approaches such as DNS, LES, DES and unsteady RANS methods can be used to compute the space-time history of the flow field, from which acoustic sources are extracted. Because of the high computational cost of the time-accurate simulations, there have been efforts to use steady RANS calculations in conjunction with a statistical model to generate the turbulent acoustic terms.
Once the acoustics sources have been evaluated, the generated noise has to be propagated in the surrounding region with linearized propagation models. The main focus of the present chapter is the description of a computational method for noise propagation in turbomachinery applications. In the next Section a linearized model is presented. Section 3 describes the numerical algorithm based on a Discontinuous Galerkin approximation on unstructured grids, and in Section 4 several applications are presented.

2. Governing equations

In principle the propagation of acoustic waves could be directly studied using the equations of the fluid motion, i.e. the Navier Stokes equations. However, it is possible to introduce some approximations in the Navier Stokes equations in order to obtain equations more suitable for aeroacoustics. At frequencies of most practical interest, viscous effects are negligible in the acoustic field because the pressure represents a far greater stress field than the viscous stresses. Moreover, these disturbances are always small, also for very loudly acoustic waves. The threshold of pain, i.e. the maximum Sound Pressure Level (SPL) which a human can endure for a very short period of time without the risk of permanent ear damage, is equal to 140 dB, which corresponds to pressure fluctuations of amplitude equal to

$$A = \sqrt{2} p_{\text{ref}} 10^{(SPL/20)} \approx 90 \text{ Pa}\,, \tag{1}$$

where p_{ref} is the reference pressure corresponding to the threshold of hearing at 1 kHz for a typical human hear. For sound propagating in gases it is equal to $p_{\text{ref}} = 2 \times 10^{-5}$ Pa. The atmospheric pressure of the standard air is equal to $p_0 = 101325$ Pa, which is 10^3 greater than the pressure variation associated with an acoustic wave at the threshold of pain, i.e., $p'/p_0 = O\left(10^{-3}\right)$, where the superscript $(.)'$ denotes acoustic quantities and the subscript $(.)_0$ denotes mean flow quantities. The corresponding density fluctuations of a progressive plane wave are

$$\frac{\rho'}{\rho_0} = \frac{p'}{\rho_0 c_0^2}\,, \tag{2}$$

also of the order of 10^{-3}, because in air $\rho_0 c_0^2 / p_0 = \gamma = c_p / c_v = 1.4$ with c_0 being the speed of sound. These estimates demonstrate that the flow perturbations involved in acoustic waves are very small compared to the mean-flow quantities: the acoustic field can be considered as a small perturbation of the mean flow field. Therefore it is possible to linearize the equations of motion. Considering acoustic waves as a perturbation of the mean flow field, defining

$$p' = p - p_0, \quad \rho' = \rho - \rho_0, \quad \boldsymbol{v}' = \boldsymbol{v} - \boldsymbol{v}_0\,, \tag{3}$$

and assuming small perturbations, it is possible to obtain the equations for the propagation of the sound waves, i.e., the Linearized Euler Equations (LEE). For a two-dimensional problem and a steady mean flow field, LEE are formulated as

$$\frac{\partial \boldsymbol{u}}{\partial t} + \frac{\partial \boldsymbol{F}_x}{\partial x} + \frac{\partial \boldsymbol{F}_y}{\partial y} + \boldsymbol{H} = \boldsymbol{S} \,, \tag{4}$$

where $\boldsymbol{u} = [\rho',\, u',\, v',\, p']^T$ is the acoustic perturbation vector, $\boldsymbol{F}_x$ and $\boldsymbol{F}_y$ are the fluxes along x and y directions respectively, $\boldsymbol{H}$ contains the mean flow derivatives and $\boldsymbol{S}$ represents the acoustic sources. The fluxes, $\boldsymbol{F}_x$ and $\boldsymbol{F}_y$, and the term $\boldsymbol{H}$ have the following expressions

$$\boldsymbol{F}_x = \begin{pmatrix} \rho' u_0 + \rho_0 u' \\ u_0 \rho_0 u' + p' \\ u_0 \rho_0 v' \\ u_0 p' + \gamma p_0 u' \end{pmatrix} \,, \qquad \boldsymbol{F}_y = \begin{pmatrix} \rho' v_0 + \rho_0 v' \\ v_0 \rho_0 u' \\ v_0 \rho_0 v' + p' \\ v_0 p' + \gamma p_0 v' \end{pmatrix} \,,$$

$$\boldsymbol{H} = \begin{pmatrix} 0 \\ (\rho_0 u' + \rho' u_0)\, \frac{\partial u_0}{\partial x} + (\rho_0 v' + \rho' v_0)\, \frac{\partial u_0}{\partial y} \\ (\rho_0 u' + \rho' u_0)\, \frac{\partial v_0}{\partial x} + (\rho_0 v' + \rho' v_0)\, \frac{\partial v_0}{\partial y} \\ (\gamma - 1) \left(p' \frac{\partial u_0}{\partial x} + p' \frac{\partial v_0}{\partial y} - u' \frac{\partial p_0}{\partial x} - v' \frac{\partial p_0}{\partial y} \right) \end{pmatrix} \,. \tag{5}$$

It is evident from Eqs.(5) that, in order to solve the LEE, the mean flow field must be known in advance.

For turbomachinery tonal noise propagation it is better to express the LEE in a cylindrical coordinate system. Given the Cartesian coordinate system $(x,\, y,\, z)$, the cylindrical system $(r,\, z,\, \theta)$ is defined as

$$\begin{cases} x = r \cos\theta \\ y = r \sin\theta \\ z = z \end{cases} \,. \tag{6}$$

With respect to this reference frame, the LEE for a three-dimensional problem read

$$\frac{\partial \boldsymbol{u}}{\partial t} + \frac{\partial \boldsymbol{F}_z^{\mathrm{AX}}}{\partial z} + \frac{\partial \boldsymbol{F}_r^{\mathrm{AX}}}{\partial r} + \frac{\partial \boldsymbol{F}_\theta^{\mathrm{AX}}}{\partial \theta} + \boldsymbol{H}^{\mathrm{AX}} = \boldsymbol{S}^{\mathrm{AX}} \,, \tag{7}$$

where $\boldsymbol{u} = [\rho',\, u,'\, v',\, w',\, p']^T$ is the acoustic perturbation vector expressed in the cylindrical coordinate system, i.e. u', v', and w' are the velocity components in (z, r, θ) directions respectively. $\boldsymbol{F}_z^{\mathrm{AX}}$, $\boldsymbol{F}_r^{\mathrm{AX}}$, $\boldsymbol{F}_\theta^{\mathrm{AX}}$ are the fluxes along z, r, and θ directions respectively, $\boldsymbol{H}^{\mathrm{AX}}$ contains the terms due to the cylindrical reference frame and to the mean flow derivatives and $\boldsymbol{S}^{\mathrm{AX}}$ represents the acoustic sources. In Section 4.3 it will be shown that a generic turbomachinery tonal wave can be expanded in a sum of complex duct modes, having the form $\hat{f}\,(z, r, t) \cdot \exp\left(\mathcal{I} m \theta\right)$ where m is an integer number which identifies the azimuthal mode and $\mathcal{I}$ the imaginary unit. Therefore, using the dependence of the acoustic field on θ, the problem can be reduced, from a three-dimensional problem, to a two-dimensional one in (r, z). For a single duct-mode the LEE become

$$\frac{\partial \hat{\boldsymbol{u}}}{\partial t} + \frac{\partial \hat{\boldsymbol{F}}_z^{\mathrm{AX}}}{\partial z} + \frac{\partial \hat{\boldsymbol{F}}_r^{\mathrm{AX}}}{\partial r} + \mathcal{I} m \hat{\boldsymbol{F}}_\theta^{\mathrm{AX}} + \hat{\boldsymbol{H}}^{\mathrm{AX}} = \hat{\boldsymbol{S}}^{\mathrm{AX}} \,, \tag{8}$$

where the superscript $\hat{(.)}$ reminds that the variable comes from a mode expansion. Assuming that the mean flow is axial-symmetric, i.e. the azimuthal component of the mean flow velocity is zero, $w_0 = 0$, the fluxes, $\hat{\boldsymbol{F}}_z^{\mathrm{AX}}$, $\hat{\boldsymbol{F}}_r^{\mathrm{AX}}$, and $\hat{\boldsymbol{F}}_\theta^{\mathrm{AX}}$ have the following expressions

$$\hat{\boldsymbol{F}}_z^{\mathrm{AX}} = \begin{pmatrix} \hat{\rho}' u_0 + \rho_0 \hat{u}' \\ u_0 \rho_0 \hat{u}' + \hat{p}' \\ u_0 \rho_0 \hat{v}' \\ u_0 \rho_0 \hat{w}' \\ u_0 \hat{p}' + \gamma p_0 \hat{u}' \end{pmatrix}, \quad \hat{\boldsymbol{F}}_r^{\mathrm{AX}} = \begin{pmatrix} \hat{\rho}' v_0 + \rho_0 \hat{v}' \\ v_0 \rho_0 \hat{u}' \\ v_0 \rho_0 \hat{v}' + \hat{p}' \\ v_0 \rho_0 \hat{w}' \\ v_0 p' + \gamma p_0 \hat{v}' \end{pmatrix}, \quad \hat{\boldsymbol{F}}_\theta^{\mathrm{AX}} = \frac{1}{r} \begin{pmatrix} \rho_0 \hat{w}' \\ 0 \\ 0 \\ \hat{p}' \\ \gamma p_0 \hat{w}' \end{pmatrix}, \tag{9}$$

whereas the term $\hat{\boldsymbol{H}}^{\mathrm{AX}}$ is given by

$$\hat{\boldsymbol{H}}^{\mathrm{AX}} = \frac{1}{r} \begin{pmatrix} -\hat{\rho}' v_0 \\ -\hat{u}' v_0 \\ -\hat{v}' v_0 - \frac{\hat{p}'}{\rho_0} \\ 0 \\ (\gamma - 1)\, \hat{p}' v_0 \end{pmatrix} + \begin{pmatrix} 0 \\ (\rho_0 \hat{u}' + \hat{\rho}' u_0) \frac{\partial u_0}{\partial z} + (\rho_0 \hat{v}' + \hat{\rho}' v_0) \frac{\partial u_0}{\partial r} \\ (\rho_0 \hat{u}' + \hat{\rho}' u_0) \frac{\partial v_0}{\partial z} + (\rho_0 \hat{v}' + \hat{\rho}' v_0) \frac{\partial v_0}{\partial r} \\ -\hat{w}' \frac{\partial u_0}{\partial z} - \hat{w}' \frac{\partial v_0}{\partial r} \\ (\gamma - 1) \left(\hat{p}' \frac{\partial u_0}{\partial z} + \hat{p}' \frac{\partial v_0}{\partial z} - \hat{u}' \frac{\partial p_0}{\partial r} - \hat{v}' \frac{\partial p_0}{\partial r} \right) \end{pmatrix}. \tag{10}$$

2.1 Frequency domain approach

The linearized Euler equations, beside acoustic waves, support also instability waves that, for a mean flow with shear-layers, are the well-known Kelvin-Helmholtz instabilities. In the complete physical problem this instabilities are limited and modified by non-linear and viscous effects. Indeed, in the linearized Euler equations, these two effects are not present. Therefore when solving LEE in presence of a shear-layer type mean flow, Kelvin-Helmholtz instabilities can grow indefinitely as they propagate down-stream from the point of introduction and the acoustic solution may be obscured by the non-physical instabilities [Agarwal et al. (2004); Özyörük (2009)]. By using a Fourier decomposition of the acoustics sources and solving the linearized Euler equations in the frequency domain one can, in principle, avoid the unbounded growth of the shear-layer type instability, since the acoustic and instability modes correspond to different values of complex frequency [Rao & Morris (2006)]. However, this could be accomplished in practice only if the discretized form of the equations is solved using a direct solver. The use of iterative techniques to solve the resulting global matrix has been discussed by Agarwal et al. (2004). It is proved that the use of any iterative technique to solve the global matrix is equivalent to a pseudo-time marching method, and hence, produces an instability wave solution. Therefore, the solution of the global matrix needs to be sought by using direct methods such as Gaussian elimination or LU decomposition techniques.

2.2 GTS-like approximation

In order to reduce computational time and memory requirements, the pressure gradients in the momentum equations are neglected. A similar approximation, termed Gradient Terms Suppression (GTS), is often used to overcome instability problems that prevent convergence of time domain algorithms for the LEE [Tester et al. (2008); Zhang et al. (2003)]. While the original GTS approximation suppresses all mean-flow gradients, which are likely to be small in the considered subsonic flows, in the present case, being interested in reducing the computational time, only the density mean-flow gradients in momentum equations are neglected. This allows to decouple the continuity equation and to solve only momentum

and energy equations. For an axial-symmetric problem the number of total unknowns is thus reduced by a factor $\mathcal{T} = 5/4$, whereas the non-zero terms in the coefficient matrix of the linear system associated with the discretized form of Eqs. (8) is reduced by a factor $\mathcal{T}^2$. Indeed, a smaller linear system can be solved faster, and, more important, its resolution requires less memory.

2.3 Boundary conditions

When a problem is solved numerically, the governing equations must be solved only within the domain, whereas on its borders appropriate conditions, called boundary conditions, must be imposed. In many situations the boundary conditions associated with the continuous problem do not completely supply the discrete problem, and numerical boundary conditions must be added.

Rigid Walls

If walls are assumed impermeable and acoustically rigid, no flow passes through the boundary and acoustic waves are totally reflected. Assuming that the mean flow satisfies the slip flow boundary condition, an analogous slip flow condition must be imposed on the velocity fluctuations

$$\boldsymbol{u}' \cdot \boldsymbol{n} = 0\,, \tag{11}$$

where $\boldsymbol{u}'$ is the acoustic velocity and $\boldsymbol{n}$ is the normal vector to the wall. To apply this condition, Eq. (11) is used to express one of the velocity components in terms of the others.

Axial symmetry

When dealing with axial-symmetric problems, the equations could be solved only for $r \geq 0$ if an appropriate boundary condition is applied on the symmetry axis. Along that boundary, the acoustic velocity should be aligned with the $r = 0$ axis, this can be achieved applying a wall type boundary condition.

Far-field boundary

One of the major issues in CAA is to truncate the far-field domain preserving a physically meaning solution. This leads to the necessity to have accurate and robust non-reflecting far-field boundary conditions. A large number of families of non-reflecting boundary conditions has been derived in literature. The most widely used for the Euler equations are the characteristics-based boundary conditions [Giles (1990); Thompson (1990)]. These methods are derived applying the one-dimensional characteristic-variable splitting in the boundary-normal direction. This technique is usually efficient and robust. The main drawback is that reflections are prevented only for waves that are traveling in the boundary-normal direction. Not negligible reflections can be seen for waves that hit the boundary with other angles. Another class of non-reflecting boundary conditions is based on the asymptotic solutions of the wave equation [Bayliss & Turkel (1980); Tam & Webb (1993)]. In this case, the governing equations are replaced in the far field by an analytic solution obtained imposing an asymptotic behavior to the system. These conditions can be very accurate. Unfortunately, the asymptotic solution can be achieved only in a limited number of cases, reducing the applicability of this model to test cases.
Another family of non-reflective boundary conditions is composed by the buffer zone technique [Bodony (2006); Hu (2004)]. In this case, an extra zone is added to damp the reflected waves. The damping can be introduced as a low-pass filter, grid stretching or accelerating the mean flow to supersonic speed. The main drawback of these techniques is

represented by the increase of the computational cost, as the thickness of the buffer zone could be important to achieve a good level of accuracy.
More recently, the Perfectly Matched Layer (PML) technique has been developed as a new class of non-reflective boundary conditions. The basic idea of the PML approach is to modify the governing equations in order to absorb the out-going waves in the buffer region. The advantage of this technique is that the absorbing layer is theoretically capable to damp waves of any direction and frequency, resulting in thinner layers with respect to other buffer zone approaches, with benefits on the efficiency and the accuracy of the solution. Originally proposed by Berenger (1994) for the solution of the Maxwell equations, the PML technique was extended to CAA applying the split physical variable formulation to the linearized Euler equations with uniform mean flow [Hu (1996)]. It was shown that the PML absorbing zone is theoretically reflectionless to the acoustic, vorticity and entropy waves. Nonetheless, numerical instability arises in this formulation, and in Tam et al. (1998) the presence of instability waves is demonstrated. In Hu (2001), it was shown that the instability of the split formulation is due to an inconsistency of the phase and group velocities of the acoustic waves in presence of a mean flow, and a stable PML formulation for the linearized Euler equation was proposed, based on an unsplit physical variable formulation.
The PML technique can be seen as a change of variable in the frequency domain, for example, considering the vertical layer, this change of variable can be written as

$$x \to x + \frac{i}{\omega} \int_{x_0}^{x} \sigma_x dx \,, \tag{12}$$

where $\sigma_x > 0$ is the absorption coefficient and x_0 is the location of the PML/LEE interface. To avoid instabilities, a proper space-time transformation must be used before applying the PML change of variable, so that in the transformed coordinates all linear waves supported by the LEE have consistent phase and group velocities. Assuming that the mean flow in the absorbing layer is uniform and parallel to the x axis, the proper space-time transform involves a transformation in time of the form [Hu (2001)]

$$\bar{t} = t + \frac{M_0}{c_0 \left(1 - M_0^2\right)} x \,, \tag{13}$$

where $M_0 = u_0 / c_0$.
A stable PML formulation for the two-dimensional LEE can be obtained applying the space-time transformation Eq. (13) to Eqs.(8) and then using the PML change of variable Eq. (12) in the transformed coordinates. Expressing the formulation in the original (x, y) coordinates the PML formulation becomes

$$\frac{\partial \boldsymbol{F}_x^{\mathrm{PML}}}{\partial x} + \frac{\partial \boldsymbol{F}_y^{\mathrm{PML}}}{\partial y} + \boldsymbol{H}^{\mathrm{PML}} = 0 \,. \tag{14}$$

The terms $\boldsymbol{F}_x^{\mathrm{PML}}$, $\boldsymbol{F}_y^{\mathrm{PML}}$, and $\boldsymbol{H}^{\mathrm{PML}}$ of Eq. (14) are defined as follow

$$\boldsymbol{F}_x^{\mathrm{PML}} = \alpha_y \tilde{\boldsymbol{F}}_x \,; \quad \boldsymbol{F}_y^{\mathrm{PML}} = \alpha_x \tilde{\boldsymbol{F}}_y \,; \quad \boldsymbol{H}^{\mathrm{PML}} = \left(\alpha_x \alpha_y \tilde{\boldsymbol{H}} + \alpha_x \sigma_y \frac{M_0}{c_0 \left(1 - M_0^2\right)} \tilde{\boldsymbol{F}}_x \right) , \tag{15}$$

where $\alpha_x = \left(1 + \frac{\sigma_x}{i\omega}\right)$, $\alpha_y = \left(1 + \frac{\sigma_y}{i\omega}\right)$ and $\tilde{\boldsymbol{F}}_x$, $\tilde{\boldsymbol{F}}_y$, and $\tilde{\boldsymbol{H}}$ are the terms of Eq. (5) under the assumption that the mean flow is uniform and parallel to the x axis. The damping constants

σ_x and σ_y have the following expressions

$$\sigma_x = \sigma_{max}\left(1 - M_0^2\right)\left|\frac{x - x_l}{D_x}\right|^{\beta}, \qquad \sigma_y = \sigma_{max}\left|\frac{y - y_l}{D_y}\right|^{\beta}, \tag{16}$$

where D_x and D_y are the widths of the absorbing layers in the x and y directions respectively and x_l and y_l are the positions of the interfaces between the PML region and the physical domain. The maximum value of the damping σ_{max} is usually taken as $2c_0/\Delta x$ and the coefficient β is set to 2 [Hu (2001)]. At the end of the PML domain, no special boundary conditions are needed except those that are necessary to maintain the numerical stability of the scheme. For this reason at the external boundary of the absorbing layer wall boundary conditions are applied.
Applying the same PML formulation to the LEE written for the turbomachinery duct modes, the following system is obtained

$$\frac{\partial \boldsymbol{F}_z^{\mathrm{PML}}}{\partial z} + \frac{\partial \boldsymbol{F}_r^{\mathrm{PML}}}{\partial r} + \mathcal{I}m\boldsymbol{F}_\theta^{\mathrm{PML}} + \boldsymbol{H}^{\mathrm{PML}} = 0\,. \tag{17}$$

The terms $\boldsymbol{F}_z^{\mathrm{PML}}$, $\boldsymbol{F}_r^{\mathrm{PML}}$, $\boldsymbol{F}_\theta^{\mathrm{PML}}$, and $\boldsymbol{H}^{\mathrm{PML}}$ of Eq. (17) are defined as follow

$$\boldsymbol{F}_z^{\mathrm{PML}} = \alpha_r \tilde{\boldsymbol{F}}_z^{\mathrm{AX}}, \quad \boldsymbol{F}_r^{\mathrm{PML}} = \alpha_z \tilde{\boldsymbol{F}}_r^{\mathrm{AX}}, \quad \boldsymbol{F}_\theta^{\mathrm{PML}} = \alpha_z\alpha_r \tilde{\boldsymbol{F}}_\theta^{\mathrm{AX}}, \tag{18}$$

$$\boldsymbol{H}^{\mathrm{PML}} = \left(\alpha_z\alpha_r \tilde{\boldsymbol{H}}^{\mathrm{AX}} + \alpha_z\sigma_r \frac{M_0}{c_0\left(1 - M_0^2\right)} \tilde{\boldsymbol{F}}_z^{\mathrm{AX}}\right), \tag{19}$$

where α_z and α_r are defined as in the two-dimensional case and $\tilde{\boldsymbol{F}}_z$, $\tilde{\boldsymbol{F}}_r$, $\tilde{\boldsymbol{F}}_\theta$, and $\tilde{\boldsymbol{H}}$ are the terms of Eq. (9) and Eq. (10) under the assumption that the mean flow is uniform and parallel to the z axis.

Acoustic inlet

The PML formulation is also used to impose incoming waves at acoustic inlet boundaries. On those boundaries incoming waves should be specified, but at the same time outgoing waves should leave the computational domain without reflections. This can be achieved applying the PML equations to the reflected wave, $\boldsymbol{u}_{\mathrm{re}}$ [Özyörük (2009)], which can be expressed as the total acoustic field, $\boldsymbol{u}$, minus the incoming prescribed acoustic wave, $\boldsymbol{u}_{\mathrm{in}}$

$$\boldsymbol{u}_{\mathrm{re}} = \boldsymbol{u} - \boldsymbol{u}_{\mathrm{in}}\,. \tag{20}$$

Considering the two-dimensional problem and substituting Eq. (20) into Eq. (14) the equation for the inlet PML domain reads

$$\frac{\partial}{\partial x}\left(\boldsymbol{F}_x^{\mathrm{PML}}\left(\boldsymbol{u} - \boldsymbol{u}_{\mathrm{in}}\right)\right) + \frac{\partial}{\partial y}\left(\boldsymbol{F}_y^{\mathrm{PML}}\left(\boldsymbol{u} - \boldsymbol{u}_{\mathrm{in}}\right)\right) + \boldsymbol{H}^{\mathrm{PML}}\left(\boldsymbol{u} - \boldsymbol{u}_{\mathrm{in}}\right) = 0\,. \tag{21}$$

Since the linearized Euler flux functions are linear, Eq. (21) becomes

$$\begin{aligned}\frac{\partial}{\partial x}\left(\boldsymbol{F}_x^{\mathrm{PML}}\left(\boldsymbol{u}\right)\right) + \frac{\partial}{\partial y}\left(\boldsymbol{F}_y^{\mathrm{PML}}\left(\boldsymbol{u}\right)\right) + \boldsymbol{H}^{\mathrm{PML}}\left(\boldsymbol{u}\right) = \\ = \frac{\partial}{\partial x}\left(\boldsymbol{F}_x^{\mathrm{PML}}\left(\boldsymbol{u}_{\mathrm{in}}\right)\right) + \frac{\partial}{\partial y}\left(\boldsymbol{F}_y^{\mathrm{PML}}\left(\boldsymbol{u}_{\mathrm{in}}\right)\right) + \boldsymbol{H}^{\mathrm{PML}}\left(\boldsymbol{u}_{\mathrm{in}}\right)\,.\end{aligned} \tag{22}$$

The same procedure can be used for the acoustic inlet boundaries of the axial-symmetric LEE.

3. Numerical methods

The numerical solution of the LEE requires highly accurate and efficient algorithms able to mimic the non-dispersive and non-diffusive nature of the acoustic waves propagating over long distances. One of the most popular numerical scheme in CAA is the Dispersion Relation Preserving (DRP) algorithm originally proposed by Tam & Webb (1993). The DRP scheme is designed for Cartesian or highly regular curvilinear coordinates. However, in many practical applications, complex geometries must be considered and unstructured grids may be necessary. One of the most promising numerical scheme able to fulfill all the above requirements is the Discontinuous Galerkin method (DGM or DG method).
The DGM was firstly proposed in the early seventies by Reed and Hill in the frame of the neutron transport [Reed & Hill (1973)]. Since then, the method has found its use in many different computational models. In the last years, in the context of CFD, DGM has gained an increasing popularity because of its superior properties with respect to more traditional schemes in terms of accuracy and intrinsic stability [Cockburn et al. (2000)].
The DG method displays many interesting properties. It is compact: regardless of the order of the element, data are only exchanged between neighboring elements. It is well suited for complex geometries because the expected dispersion and dissipation properties are retained also on unstructured grids. Furthermore in the framework of DGM it is straightforward to implement the boundary conditions, since only the flux needs to be specified at the boundary. The main disadvantage of the DGM is its computational cost. Because of the discontinuous character, there are extra degrees of freedom at cell boundaries in comparison to the continuous finite elements, demanding more computational resources. This drawback can be partially reduced with a static condensation technique and with a parallel implementation of the algorithm, operations which are made easier by the compactness of the scheme [Bernacki et al. (2006)].

3.1 Discontinuous Galerkin formulation

The DGM will be initially presented for the scalar problem of finding the solution u of the hyperbolic conservation equation

$$\frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{F}(u) + Hu - S = 0\,, \tag{23}$$

where $\boldsymbol{F}(u)$ is the flux vector, H is the source term and S is the forcing term. Defining a test function vector space, W, the weak form of the problem (23) over the domain Ω consists in finding $u \in W$ such that

$$\int_{\Omega} w \left(\frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{F}(u) + Hu - S \right) d\Omega = 0 \quad \forall w \in W\,. \tag{24}$$

The discontinuous Galerkin formulation is based on the idea of discretizing the domain Ω into a set of E non-overlapping elements Ω_e. Introducing the notations

$$\int_{\Omega'} (.) \stackrel{def}{=} \sum_{e=1}^{E} \int_{\Omega_e} (.)\, d\Omega\,, \qquad \int_{\partial\Omega'} (.) \stackrel{def}{=} \sum_{e=1}^{E} \int_{\partial\Omega_e} (.)\, d\Sigma\,, \tag{25}$$

the weak form can be rewritten as

$$\int_{\Omega'} w \left(\frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{F}(u) + Hu - S \right) = 0 \quad \forall w \in W\,. \tag{26}$$

To obtain an expression which explicitly contains the flux at the element interfaces, the divergence term in Eq. (26) is integrated by parts

$$\int_{\Omega'} \left[w \left(\frac{\partial u}{\partial t} + Hu - S \right) - \nabla w \cdot \boldsymbol{F}(u) \right] + \int_{\partial\Omega'} w \boldsymbol{F}(u) \cdot \boldsymbol{n} = 0 \,, \tag{27}$$

where $\boldsymbol{n}$ is the outward-pointing normal versor referred to each element edge. For interfaces on the domain borders, the normal flux vector is evaluated using appropriate boundary conditions. In the general case a boundary condition defines the normal flux as $\boldsymbol{F}(u) \cdot \boldsymbol{n} = \boldsymbol{F}^{\mathrm{BC}}(u) + \boldsymbol{G}^{\mathrm{BC}}$. On internal interfaces, $\boldsymbol{F}(u) \cdot \boldsymbol{n}$ is evaluated from the values of u. In order for the formulation to be consistent, the normal flux vector evaluated on right side of an internal interface must be equal to minus the normal flux vector evaluated on the left side of the same interface. Since one of the key feature of the DGM is the discontinuity of the solution among the elements, the consistency is not automatically guaranteed by the formulation. Therefore the normal flux $\boldsymbol{F}(u) \cdot \boldsymbol{n}$ is replaced by a numerical flux $\boldsymbol{F}^R(u)$ which is uniquely defined no matter of the side on which it is evaluated (see section 3.2). For ease of notation it is convenient to introduce the following definition

$$\boldsymbol{F}(u) \cdot \boldsymbol{n} \stackrel{def}{=} \boldsymbol{F}^{\partial}(u) + \boldsymbol{G}^{\partial} \,, \tag{28}$$

where $\boldsymbol{F}^{\partial} + \boldsymbol{G}^{\partial}$ is equal to $\boldsymbol{F}^{\mathrm{BC}} + \boldsymbol{G}^{\mathrm{BC}}$ for interfaces on the domain borders and is equal to to $\boldsymbol{F}^{\mathrm{R}}$ for internal ones. Furthermore, assuming that the flux vector is a linear function of the unknown, yields

$$\boldsymbol{F}(u) = \boldsymbol{A}u \,, \qquad \boldsymbol{F}(u) \cdot \boldsymbol{n} = \boldsymbol{F}^{\partial}(u) + \boldsymbol{G}^{\partial} = \boldsymbol{A}^{\partial}u + \boldsymbol{G}^{\partial} \,, \tag{29}$$

where $\boldsymbol{A}$ and $\boldsymbol{A}^{\partial}$ are two matrices representing the Jacobian of the physical flux and the Jacobian of the numerical flux respectively. Using Eq. (28) and Eqs. (29), the weak formulation reads

$$\int_{\Omega'} \left[w \left(\frac{\partial u}{\partial t} + Hu - S \right) - \nabla w \cdot (\boldsymbol{A}u) \right] + \int_{\partial\Omega'} w \left(\boldsymbol{A}^{\partial}u + \boldsymbol{G}^{\partial} \right) = 0 \,. \tag{30}$$

Given Eq. (30), the discontinuous Galerkin approximation is obtained considering a finite element space, W_h, to approximate W. On each element, a set of points called nodes or degrees of freedom is identified. The number and the position of the nodes depend on the type of approximation used. The set of nodes is chosen to be the same on each element, in this way, on element's borders, there is a direct correspondence among the nodes defined on neighboring elements. The nodes are numbered globally using the index $j_{\mathrm{glob}} = 1, 2, \ldots, n_{\mathrm{glob}}$ with n_{glob} being the global number of degrees of freedom. Beside the global numbering, there is a local numbering. On each element the nodes are identified using the index $j^e_{\mathrm{loc}} = 1, 2, \ldots, n^e_{\mathrm{loc}}$ where n^e_{loc} is the number of degrees of freedom of the e-th element. The correspondence between local node numbers and global node numbers can be expressed through a matrix called connectivity matrix

$$j_{\mathrm{glob}} = \boldsymbol{\mathcal{C}}_{e, j^e_{\mathrm{loc}}} \,. \tag{31}$$

The nodes of the discretization are used to define the finite element space W_h: the vector space W_h is generated by the Lagrangian polynomials defined on the nodes of the discretization. The variable $u \in W$ is therefore approximated in the W_h space with an interpolation of its nodal values

$$u \approx u_h = \sum_{j=1}^{n_{\mathrm{glob}}} u_j(t) \, \Phi_j(x, y) \,, \tag{32}$$

where $u_j(t)$ is the value of u in the j-th global node $\left(x_j, y_j\right)$ at the time t and Φ_j is the Lagrangian polynomial defined on the j-th global node with the property

$$\Phi_i\left(x_j, y_j\right) = \delta_{ij} \quad i, j = 1, 2, \ldots, n_{\text{glob}}\,. \tag{33}$$

Although in this work Lagrangian interpolation functions are used, other types of interpolation are possible. Considering the vector space W_h, the discrete weak form of problem (23) consists in finding $u_h \in W_h$ such that

$$\int_{\Omega'}\left[w_h\left(\frac{\partial u}{\partial t} + Hu - S\right) - \left(\nabla w_h \cdot \boldsymbol{A}\right)u\right] + \int_{\partial\Omega'} w_h\left(\boldsymbol{A}^{\partial}u - \boldsymbol{G}^{\partial}\right) = 0 \quad \forall w_h \in W_h\,. \tag{34}$$

Substituting Eq. (32) into the discrete weak form (34) leads to

$$\sum_{j=1}^{n_{\text{glob}}}\left[\left(\int_{\Omega'} w_h\Phi_j\right)\frac{\partial u_j}{\partial t} + \left(\int_{\Omega'} Hw_h\Phi_j\right)u_j - \left(\int_{\Omega'}\left(\nabla w_h \cdot \boldsymbol{A}\right)\Phi_j\right)u_j\right] + \\ + \sum_{j=1}^{n_{\text{glob}}}\left(\int_{\partial\Omega'} w_h\boldsymbol{A}^{\partial}\Phi_j\right)u_j = \int_{\Omega'} w_hS - \int_{\partial\Omega'} w_h\boldsymbol{G}^{\partial} \quad \forall w_h \in W_h\,. \tag{35}$$

This equation must hold for every admissible choice of weight functions w_h, therefore it is sufficient to test it for the n_{glob} linearly independent functions of a base of W_h. In this way it is possible to obtain n_{glob} independent algebraic equations to solve for the n_{glob} unknowns u_j. The vector space W_h is defined as the space formed by the Lagrangian polynomials Φ_i, therefore the functions Φ_i form a base for W_h. The i-th algebraic equation is obtained substituting $w_h = \Phi_i$ into Eq. (35)

$$\sum_{j=1}^{n_{\text{glob}}}\left[\left(\int_{\Omega'}\Phi_i\Phi_j\right)\frac{\partial u_j}{\partial t} + \left(\int_{\Omega'} H\Phi_i\Phi_j\right)u_j - \left(\int_{\Omega'}\left(\nabla\Phi_i \cdot \boldsymbol{A}\right)\Phi_j\right)u_j\right] + \\ + \sum_{j=1}^{n_{\text{glob}}}\left[\int_{\partial\Omega'}\Phi_i\boldsymbol{A}^{\partial}\Phi_j\right]u_j = \int_{\Omega'}\Phi_iS - \int_{\partial\Omega'}\Phi_i\boldsymbol{G}^{\partial}\,. \tag{36}$$

Taking the Fourier transform of Eq. (36), the weak formulation associated with the l-th mode can be written as

$$\sum_{j=1}^{n_{\text{glob}}}\left[\left(\int_{\Omega'}\Phi_i\Phi_j\right)\mathcal{I}\omega^{(l)} + \left(\int_{\Omega'} H\Phi_i\Phi_j\right) - \left(\int_{\Omega'}\left(\nabla\Phi_i \cdot \boldsymbol{A}\right)\Phi_j\right)\right]\hat{u}_j^{(l)} + \\ + \sum_{j=1}^{n_{\text{glob}}}\left[\int_{\partial\Omega'}\Phi_i\boldsymbol{A}^{\partial}\Phi_j\right]\hat{u}_j^{(l)} = \int_{\Omega'}\Phi_i\hat{S}^{(l)} - \int_{\partial\Omega'}\Phi_i\boldsymbol{G}^{\partial}\,, \tag{37}$$

where $\hat{(.)}_j^{(l)}$ is the l-th component of the Fourier transform of $(.)$, $\omega^{(l)}$ is the angular frequency of the l-th Fourier mode and $\mathcal{I}$ is the imaginary unit. Equation (37) represents the weak-form discontinuous Galerkin model for a scalar hyperbolic problem in the frequency domain. It can also be written in matrix notation as

$$\boldsymbol{K}\boldsymbol{u}^{(l)} = \boldsymbol{f}^{(l)}\,, \tag{38}$$

where

$$K_{ij}^{(l)} = \int_{\Omega'} \left[\left(\mathcal{I}\omega^{(l)} + H \right) \Phi_i \Phi_j - \left(\nabla \Phi_i \cdot \boldsymbol{A} \right) \Phi_j \right] + \int_{\partial\Omega'} \Phi_i \boldsymbol{A}^{\partial} \Phi_j \,, \tag{39}$$

$$f_i^{(l)} = \int_{\Omega'} \Phi_i \hat{S}^{(l)} - \int_{\partial\Omega'} \Phi_i \boldsymbol{G}^{\partial} \,. \tag{40}$$

Solving this linear system it is possible to obtain the nodal values of the l-th Fourier mode, $\boldsymbol{u}^{(l)}$.
The same formulation can be applied to the vectorial problem of finding the solution $\boldsymbol{u}$ of the system of equations

$$\frac{\partial \boldsymbol{u}}{\partial t} + \nabla \cdot \overline{\boldsymbol{F}}\left(\boldsymbol{u}\right) + \boldsymbol{H}\boldsymbol{u} - \boldsymbol{S} = 0 \,, \tag{41}$$

where $\boldsymbol{u}$ is a vector of n_{vars} unknowns, $\overline{\boldsymbol{F}}\left(\boldsymbol{u}\right)$ is the flux tensor, $\boldsymbol{H}$ is the source term, and $\boldsymbol{S}$ is the forcing term. For the vector problem the weak formulation consists in finding $\boldsymbol{u} \in \boldsymbol{W}$ such that

$$\int_{\Omega'} \left[\boldsymbol{w}^T \left(\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{H}\boldsymbol{u} - \boldsymbol{S} \right) - \nabla \boldsymbol{w}^T \cdot \overline{\boldsymbol{F}}\left(\boldsymbol{u}\right) \right] + \int_{\partial\Omega'} \boldsymbol{w}^T \overline{\boldsymbol{F}}\left(\boldsymbol{u}\right) \cdot \boldsymbol{n} = 0 \quad \forall \boldsymbol{w} \in \boldsymbol{W} \,. \tag{42}$$

Assuming a linear flux function, i.e., $\overline{\boldsymbol{F}}\left(\boldsymbol{u}\right) = \overline{\boldsymbol{A}}\boldsymbol{u}$ and $\overline{\boldsymbol{F}}\left(\boldsymbol{u}\right) \cdot \boldsymbol{n} = \overline{\boldsymbol{A}}^{\partial}\boldsymbol{u} + \overline{\boldsymbol{G}}^{\partial}$, the Discontinuous Galerkin approximation in the frequency domain leads to the following linear system

$$\overline{\boldsymbol{K}}\overline{\boldsymbol{u}}^{(l)} = \overline{\boldsymbol{f}}^{(l)} \,, \tag{43}$$

where $\overline{\boldsymbol{u}}^{(l)}$ contains the vectorial nodal values of the l-th Fourier mode and the system is defined as

$$\boldsymbol{K}_{ij}^{(l)} = \int_{\Omega'} \left[\boldsymbol{\Phi}_i^T \left(\mathcal{I}\omega^{(l)} \boldsymbol{I} + \boldsymbol{H} \right) \boldsymbol{\Phi}_j - \left(\nabla \boldsymbol{\Phi}_i^T \cdot \overline{\boldsymbol{A}} \right) \boldsymbol{\Phi}_j \right] + \int_{\partial\Omega'} \boldsymbol{\Phi}_i^T \overline{\boldsymbol{A}}^{\partial} \boldsymbol{\Phi}_j \,, \tag{44}$$

$$\boldsymbol{f}_i^{(l)} = \int_{\Omega'} \boldsymbol{\Phi}_i^T \hat{\boldsymbol{S}}^{(l)} - \int_{\partial\Omega'} \boldsymbol{\Phi}_i^T \overline{\boldsymbol{G}}^{\partial} \,, \tag{45}$$

with $\boldsymbol{I}$ being the identity matrix.

3.2 Interface flux

The flux through an interface has to be uniquely computed, but, due to the discontinuous function approximation, flux terms are not uniquely defined at element interfaces. Therefore, to evaluate the flux at element interfaces, a technique traditionally used in finite volume schemes is borrowed by the discontinuous Galerkin formulation: the flux function $\overline{\boldsymbol{F}}\left(\boldsymbol{u}\right) \cdot \boldsymbol{n}$ of the vector weak form, Eq. (42), is replaced by a numerical flux function, called Riemann flux, $\boldsymbol{F}^R\left(\boldsymbol{u}\right)$. Arbitrarily designating one element of the interface to be on the left , l, and the other to be on the right, r, the numerical flux depends only on the internal interface state, $\boldsymbol{u}_l$, on the neighboring element interface state, $\boldsymbol{u}_r$, and on the direction $\boldsymbol{n}$ normal to the interface, i.e. $\boldsymbol{F}^R\left(\boldsymbol{u}\right) = \boldsymbol{F}^R\left(\boldsymbol{u}_r, \boldsymbol{u}_l, \boldsymbol{n}\right)$. In order to guarantee the formal consistency of the scheme, $\boldsymbol{F}^R$ is required to satisfy the relations

$$\boldsymbol{F}^R\left(\boldsymbol{u}_r, \boldsymbol{u}_l\right) = \overline{\boldsymbol{F}}\left(\boldsymbol{u}\right) \cdot \boldsymbol{n} \,, \quad \boldsymbol{F}^R\left(\boldsymbol{u}_r, \boldsymbol{u}_l\right) = -\boldsymbol{F}^R\left(\boldsymbol{u}_l, \boldsymbol{u}_r\right) \,, \tag{46}$$

which are the consistency and the conservative conditions respectively. In the present work, the Riemann flux $\boldsymbol{F}^R$ is approximated by the Osher flux. This approach is based on the diagonalization of the Jacobian matrix [Toro (1999)]. Assuming a linear dependence of the flux function on the unknown $\boldsymbol{u}$, the flux along the interface normal direction can be written as

$$\overline{\boldsymbol{F}}\left(\boldsymbol{u}\right)\cdot\boldsymbol{n}=\left(\overline{\boldsymbol{A}}\boldsymbol{u}\right)\cdot\boldsymbol{n}=\boldsymbol{A}_n\boldsymbol{u}\,, \tag{47}$$

where $\boldsymbol{A}_n=\overline{\boldsymbol{A}}\cdot\boldsymbol{n}$. The numerical method is applied to a hyperbolic system, i.e., LEE, which has a diagonalizable Jacobian matrix $\boldsymbol{A}_n$, that is

$$\boldsymbol{A}_n=\boldsymbol{K}\boldsymbol{\Lambda}\boldsymbol{K}^{-1}\,, \tag{48}$$

where $\boldsymbol{K}$ is the non-singular matrix whose columns are the right eigenvectors of $\boldsymbol{A}_n$

$$\boldsymbol{K}=\left[\boldsymbol{K}^1;\boldsymbol{K}^1;\ldots;\boldsymbol{K}^{n_{\text{vars}}};\right]\,, \tag{49}$$

and $\boldsymbol{\Lambda}$ is the diagonal matrix formed by the eigenvalues λ_i

$$\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & \ldots & 0\\ \vdots & \ddots & \vdots\\ 0 & \ldots & \lambda_m\end{pmatrix}\quad. \tag{50}$$

Given the diagonalization of $\boldsymbol{A}_n$ it is convenient to introduce the diagonal matrix formed by the absolute eigenvalues, $|\boldsymbol{\Lambda}|$, and the corresponding absolute flux matrix

$$|\boldsymbol{A}_n|=\boldsymbol{K}\,|\boldsymbol{\Lambda}|\,\boldsymbol{K}^{-1}\,. \tag{51}$$

Using the Osher approach the numerical flux $\boldsymbol{F}^R$ can be written as

$$\boldsymbol{F}^R\left(\boldsymbol{u}_r,\boldsymbol{u}_l\right)=\frac{1}{2}\left[\overline{\boldsymbol{F}}\left(\boldsymbol{u}_r\right)+\overline{\boldsymbol{F}}\left(\boldsymbol{u}_l\right)\right]-\frac{1}{2}\int_{\boldsymbol{u}_l}^{\boldsymbol{u}_r}|\boldsymbol{A}_n|\,d\boldsymbol{u}\,. \tag{52}$$

Assuming that the Jacobian matrix does not depend upon the unknown and using the hypothesis of linear fluxes, Eq. (47), the numerical flux becomes

$$\boldsymbol{F}^R\left(\boldsymbol{u}_r,\boldsymbol{u}_l\right)=\frac{1}{2}\left(\overline{\boldsymbol{A}}-|\boldsymbol{A}_n|\right)\boldsymbol{u}_r+\frac{1}{2}\left(\overline{\boldsymbol{A}}+|\boldsymbol{A}_n|\right)\boldsymbol{u}_l=\boldsymbol{A}^R\boldsymbol{u}\,. \tag{53}$$

3.3 Numerical integration

Integrals of Eq. (39) and Eq. (40) can be evaluated numerically for every element of the mesh using Gauss quadrature formulae. However, it is not convenient to evaluate the integrals directly on the generic element: it is easier to transform (or map) every element of the finite element mesh, Ω_e, into a reference element, $\hat{\Omega}$, called master element and perform the numerical integration on this master element. The transformation between Ω_e and $\hat{\Omega}$ is accomplished by a coordinate transformation from the physical coordinates (x,y) to the reference coordinates (ξ,η)

$$x=\sum_{i=1}^{m}x_i\Psi_i^e\left(\xi,\eta\right)\;;\quad y=\sum_{i=1}^{m}y_i\Psi_i^e\left(\xi,\eta\right)\,, \tag{54}$$

where m is the number of parameters used to identify the transformation, (x_i, y_i) are the global coordinates of the points of the element used in the transformation and Ψ_i^e denote the interpolation functions used in the transformation. It is important to point out that the functions Ψ_i^e used for the approximation of the geometry differ from the functions Φ_i^e used for the interpolation of the dependent variables. In this work linear interpolation functions Ψ_i^e are used: on each element the number of parameters used to identify the transformation is equal to the number of vertexes, $n = n_{\text{vertex}}$, and the (x_i, y_i) points used in the transformation are the vertexes of the element.
Once integrands are expressed on the master element $\hat{\Omega}$, numerical integration is performed using Gauss quadrature formulae, in the form

$$\int_{\hat{\Omega}} F(\xi, \eta)\, d\hat{\Omega} \approx \sum_{i=1}^{M} F(\xi_i, \eta_i)\, W_i\,, \tag{55}$$

where M denotes the number of quadrature points, (ξ_i, η_i) are the Gauss points and W_i denotes the corresponding Gauss weights.

3.4 Interpolation functions

The basis $\{\Phi_i\}$ are also evaluated over the master elements: they are the Lagrangian polynomials defined on the node set $T_p = \{\boldsymbol{x}_i;\ i = 1, \ldots, N\}$, where N is the number of nodes in the node set. For rectangular elements the basis are obtained as the tensor product of the corresponding one-dimensional Lagrangian polynomials defined on the Gauss-Lobatto nodes. Given the one-dimensional polynomials $\phi_l(\xi)$ with $l = 1, \ldots, N_\xi$ and $\phi_r(\eta)$ with $r = 1, \ldots, N_\eta$, the two-dimensional ones are defined as

$$\Phi_i(\xi, \eta) = \phi_l(\xi) \cdot \phi_r(\eta)\,, \qquad i = 1, \ldots, N_\xi N_\eta\,. \tag{56}$$

For triangular elements the Lagrangian polynomials are constructed on a set of nodes which is defined in such a way that the internal-node positions are the solutions of a steady state, minimum energy electrostatics problem, whereas the nodes along the edges are specified as one-dimensional Gauss-Lobatto quadrature points [Hesthaven (1998)].

3.5 Static condensation

One of the main disadvantages of using the DGM for solving LEE in frequency domain is the requirement of a huge amount of memory. The method leads to a linear system of equations which, as explained above, has to be solved with a direct solver, thus requiring a great amount of memory. To partially overcome this problem, a static condensation method can be applied. Static condensation allows to assemble and solve a system matrix which contains only the degrees of freedom associated with the element boundary nodes [Karniadakis & Sherwin (2005)]. Distinguishing between the boundary and interior components of the vectors $\boldsymbol{u}^e$ and $\boldsymbol{f}^e$ using $\boldsymbol{u}_b^e$, $\boldsymbol{u}_i^e$ and $\boldsymbol{f}_b^e$, $\boldsymbol{f}_i^e$ respectively, that is

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_b \\ \boldsymbol{u}_i \end{bmatrix} \quad , \qquad \boldsymbol{f} = \begin{bmatrix} \boldsymbol{f}_b \\ \boldsymbol{f}_i \end{bmatrix}\,, \tag{57}$$

the DGM linear system (38) can be written as

$$\begin{bmatrix} \boldsymbol{K}_b & \boldsymbol{K}_{c1} \\ \boldsymbol{K}_{c2} & \boldsymbol{K}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_b \\ \boldsymbol{u}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}_b \\ \boldsymbol{f}_i \end{bmatrix}\,. \tag{58}$$

In this decomposition the block $\boldsymbol{K}_b$ corresponds to the global assembly of the elemental boundary-boundary basis interaction, $\boldsymbol{K}_{c1}$ and $\boldsymbol{K}_{c2}$ correspond to the global assembly of the elemental boundary-interior coupling and $\boldsymbol{K}_i$ corresponds to the interior-interior coupling. The static condensation of internal degrees of freedom consists in performing a block elimination by a pre-multiplication of the system by the matrix

$$\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{K}_{c1}\boldsymbol{K}_i^{-1} \\ 0 & \boldsymbol{I} \end{bmatrix}, \tag{59}$$

leading to

$$\begin{bmatrix} \boldsymbol{K}_b - \boldsymbol{K}_{c1}\boldsymbol{K}_i^{-1}\boldsymbol{K}_{c2} & 0 \\ \boldsymbol{K}_{c2} & \boldsymbol{K}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_b \\ \boldsymbol{u}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}_b - \boldsymbol{K}_{c1}\boldsymbol{K}_i^{-1}\boldsymbol{f}_i \\ \boldsymbol{f}_i \end{bmatrix}. \tag{60}$$

The elemental boundary unknowns can therefore be evaluated solving the linear system

$$\left(\boldsymbol{K}_b - \boldsymbol{K}_{c1}\boldsymbol{K}_i^{-1}\boldsymbol{K}_{c2}\right)\boldsymbol{u}_b = \boldsymbol{f}_b - \boldsymbol{K}_{c1}\boldsymbol{K}_i^{-1}\boldsymbol{f}_i\,. \tag{61}$$

From equation (61) it is evident that, using the static condensation, it is possible to assembly and solve a system that contains only the degrees of freedom associated to the boundary nodes. The therm $\left(\boldsymbol{K}_b - \boldsymbol{K}_{c1}\boldsymbol{K}_i^{-1}\boldsymbol{K}_{c2}\right)$ is the Schur complement of the full system matrix and can be globally assembled starting from the Schur complements of the elemental matrix

$$\left(\boldsymbol{K}_b - \boldsymbol{K}_{c1}\boldsymbol{K}_i^{-1}\boldsymbol{K}_{c2}\right) = \mathcal{A}_b^T\left[\underline{\boldsymbol{K}_b^e - \boldsymbol{K}_{c1}^e\left[\boldsymbol{K}_i^e\right]^{-1}\boldsymbol{K}_{c2}^e}\right]\mathcal{A}_b = \mathcal{A}_b^T\,\underline{M^e}\,\mathcal{A}_b\,, \tag{62}$$

where the superscript $(.)^e$ denotes elemental matrices and $\underline{M^e}$ is a block diagonal matrix which has been formed by the local matrices M^e with $e = 1, \ldots, E$. $\mathcal{A}_b$ is the matrix that performs the scattering from the global boundary degrees of freedom to the boundary degrees of freedom, that is

$$\begin{bmatrix} \boldsymbol{u}_b^1 & \boldsymbol{u}_b^2 & \ldots & \boldsymbol{u}_b^I \end{bmatrix}^T = \mathcal{A}_b\boldsymbol{u}_b\,, \tag{63}$$

where $\boldsymbol{u}_b^e$ contains the components of $\boldsymbol{u}_b^e$ in element e. Similarly, $\mathcal{A}_b^T$ is the matrix which performs the assembly process from local to global degrees of freedom.
Once the linear system of Eq. (61) is solved and the elemental boundary solution is known the solution for the interior elemental nodes is given by the second row of Eq. (60), i.e.,

$$\boldsymbol{u}_i = \boldsymbol{K}_i^{-1}\left(\boldsymbol{f}_i - \boldsymbol{K}_{c2}\boldsymbol{u}_b\right). \tag{64}$$

Since Eq. (60) involves matrix-vector product of known quantities, it can be evaluated locally within every element, leading to

$$\boldsymbol{u}_i^e = \left[\boldsymbol{K}_i^e\right]^{-1}\left(\boldsymbol{f}_i^e - \boldsymbol{K}_{c2}^e\boldsymbol{u}_b^e\right). \tag{65}$$

3.5.1 Linear system solver

The discrete problem leads to a complex matrix system where the complex Fourier coefficients of the acoustic fluctuations are the unknowns. As stated in Section 2.1, this system must be solved with a direct method in order to avoid the Kelvin-Helmholtz instabilities. For this purpose the MUMPS (MUltifrontal Massively Parallel Solver) package Amestoy et al. (2006) will be adopted. MUMPS uses a direct method based on a multifrontal approach which performs a direct factorization $\boldsymbol{K} = \boldsymbol{L}\boldsymbol{U}$ or $\boldsymbol{K} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{t}}$ depending on the symmetry of the matrix. In the multifrontal method the factorization of a sparse matrix is achieved through the partial factorization of many, smaller dense matrices (called frontal matrices).

4. Applications

4.1 Multi-geometry scattering problem

A typical test-case to assess the ability of an aeroacustical code to resolve complex geometries is the two-dimensional scattering of sound generated by a spatially distributed monopole source from two rigid circular cylinders, as defined in the Fourth Computational Aeroacoustics (CAA) Workshop on Benchmark Problems [Scott & Sherer (2004)]. The scattering problem is presented here in terms of non-dimensional quantities. Assuming a mean flow at rest, variables can be non-dimensionalized using the mean flow pressure p_0, density ρ_0, and speed of sound c_0. To generate a time-harmonic monopole, only the source term in the energy equation has to be different from zero, i.e. $S = [0, 0, 0, S_e]$. The forcing term S_e is a Gaussian function and can be written in a source-centered coordinate system as

$$S_e = \epsilon \cdot \exp\left[-\ln(2) \cdot \left(\frac{x_S^2 + y_S^2}{b^2}\right)\right] \sin(\omega t) \ , \tag{66}$$

where $\omega = 8\pi$, $b = 0.2$, $\epsilon = 0.4$. The cylinders have unequal diameters ($D_1 = 1.0$, $D_2 = 0.5$), with the source located on the $x-$axis and equidistant from the center of each cylinder. In the (x_S, y_S)-coordinate system centered on the source, the locations of the cylinders are given by $L_1 = (-4, 0)$, and $L_2 = (4, 0)$. Considering the symmetry of the problem, only the $y \geq 0$ half-domain can be considered if an appropriate symmetry boundary condition is applied on the $x-$axis. To obtain such a symmetry boundary condition it is sufficient to consider the $x-$axis as an acoustically rigid wall. The physical domain extends for $x \in [-10, 10]$, $y \in [0, 10]$ and is surrounded by a PML region with a thickness equal to 0.75. The domain is discretized with an unstructured grid, figure (1), of about 27,000 elements (both triangles and quadrangles) and on each element Lagrangian basis of degree $p = 4$ are used.

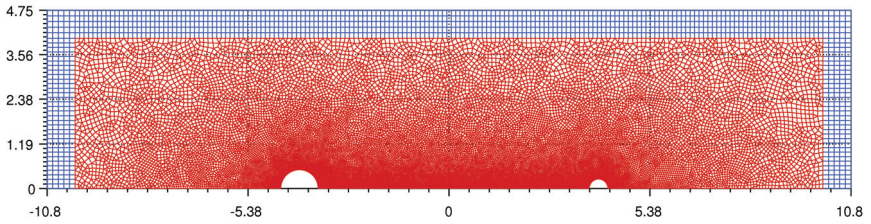


Fig. 1. Mesh of the internal (red) and PML (blue) domains

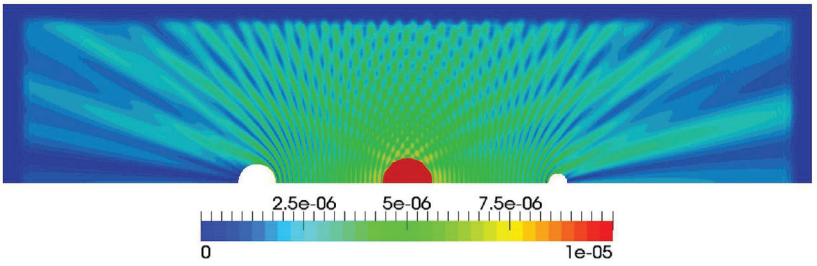


Fig. 2. RMS of the fluctuating pressure field

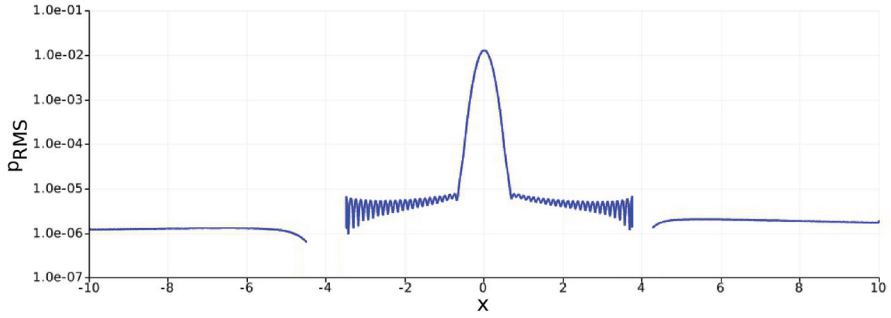


Fig. 3. RMS of the fluctuating pressure along the x axis

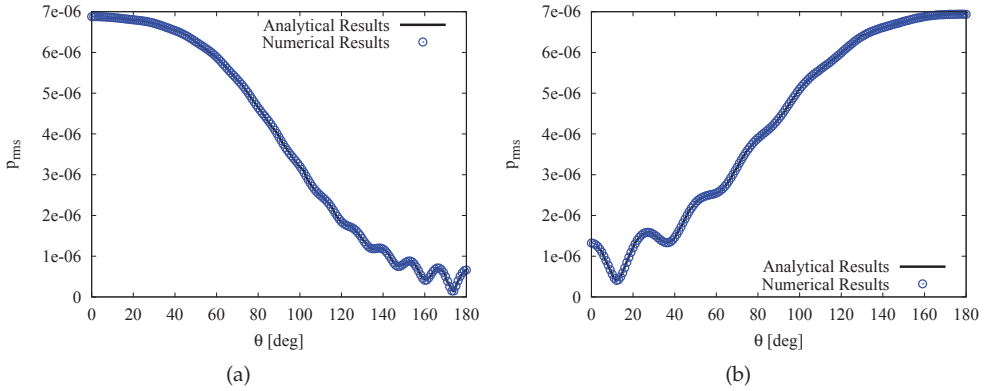


Fig. 4. RMS of the fluctuating pressure on the surface of the cylinders: ○ numerical, — analytical; (a) left cylinder; (b) right cylinder

Figure (2) shows the computed mean-square fluctuating-pressure field over the entire resolved portion of the computational domain, including the PML. In figures (3) and (4), the RMS of the fluctuating pressure is plotted along the center line and on the surface of the cylinders. Figure (4) shows a very good agreement between the numerical and the analytical solution [Scott & Sherer (2004)]. Computational time is about 5 minutes using 4 cores of a dual Intel Xeon quad-core computer and the calculation requires 5Gb of RAM.

4.2 Sound propagation around a high-lift airfoil

To assess the ability of the present method to simulate realistic geometries, the scattering of a monopole source from an high-lift airfoil is considered here . The airfoil is a three element airfoil based on the RA16SC1 profile, with the slat and flap deflected by 30 deg and 20 deg, respectively. The chord in fully retracted configuration is 0.480m. Mean flow is at rest with a speed of sound equal to $c_0 = 340.17$ m/s and a density equal to $\rho_0 = 1.225\,\mathrm{Kg/m^3}$. The computational domain extends for $(x, y) \in [-0.85\text{ m}; 0.85\text{ m}]$ and it is surrounded by vertical and horizontal PML layers with a thickness of 0.1 m. This domain is discretized with an unstructured grid with about $28,000$ elements refined in proximity of slat and flap, as shown in figure (5).

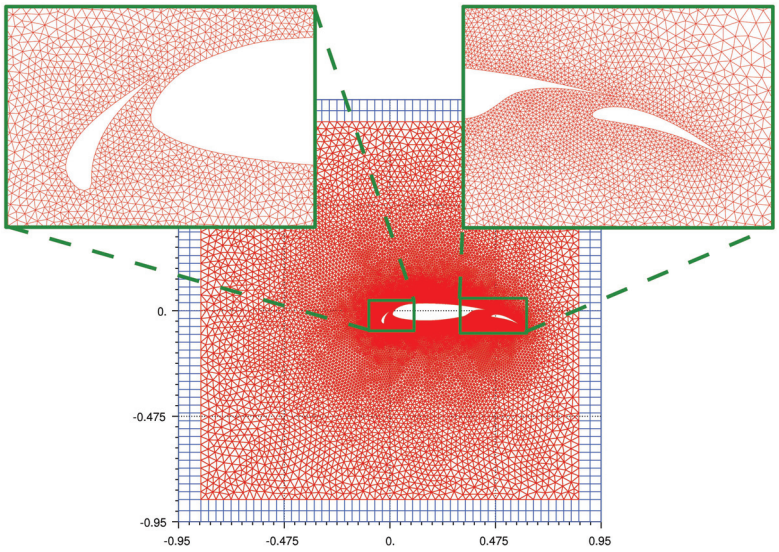


Fig. 5. High-lift airfoil scattering: mesh of the internal (red) and PML (blue) domains

A monopole source is placed near the trailing edge of the slat. As for the multi-geometry scattering problem of section 4.1, there is only the source for the energy equation, S_e. The expression of S_e is given by Eq. (66) with $\omega = 2\pi f$, $f = 5000\,\text{Hz}$, $\epsilon = 2\,\text{Pa}$, $b = 0.01\,\text{m}$ and $(x_S, y_S) = (0.02\,\text{m}, 0.02\,\text{m})$. Different computations are performed varying the interpolation polynomial order: starting from $p = 3$ up to $p = 6$. The solution obtained using polynomials of degree equal to 6 is taken as the reference solution. In figure (6) it is shown the root mean square (RMS) of the instantaneous pressure distribution for the reference solution. Computational time using 6^{th} order polynomials is less then 10 minutes using 4 cores of a dual Intel Xeon quad-core computer and the calculation requires about 6.5Gb of RAM. To compare the solutions obtained using different polynomial orders, RMS pressure values are extracted on a circle with radius of 0.7m centered in $(0, 0)$, see figure (7(a)). In figure (7(b)) it can be shown that as the interpolation polynomial order increases the directivity converges to the reference solution.

4.3 Turbomachinery noise

A turbomachine produces two type of noise: a broad-band noise, associated with vortex shedding, wake turbulence and blade vibrations, and a tonal noise, related to steady aerodynamic blade loading and to blade thickness effects. Indeed, this last kind of noise is generated because the rotor blade rotation and the aerodynamic interaction of rotor blades with stationary vane wakes generates a periodic pressure field. This field is the source of a narrow banded spectrum type acoustic field, known as spinning mode tones. For a duct having a constant and geometrically simple cross section, spinning mode tones can be studied analytically with the duct mode theory, see appendix (6). However, when the duct has a non-constant or non-simple cross section, like in the case of realistic engine geometries, no exact solution can be found and the problem must be solved numerically.

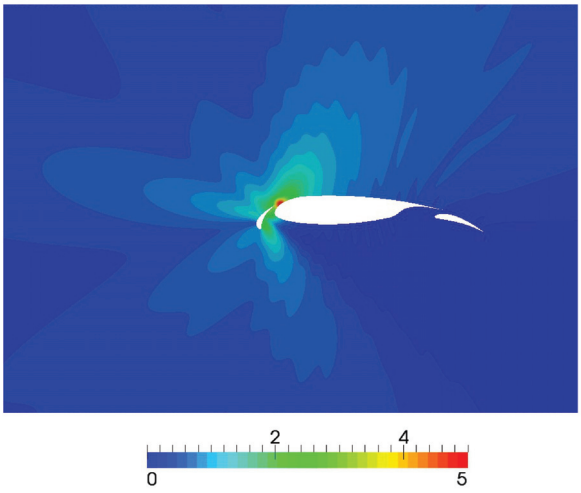


Fig. 6. RMS of the instantaneous pressure field near the profile using interpolation polynomials of degree $p = 6$, Pa

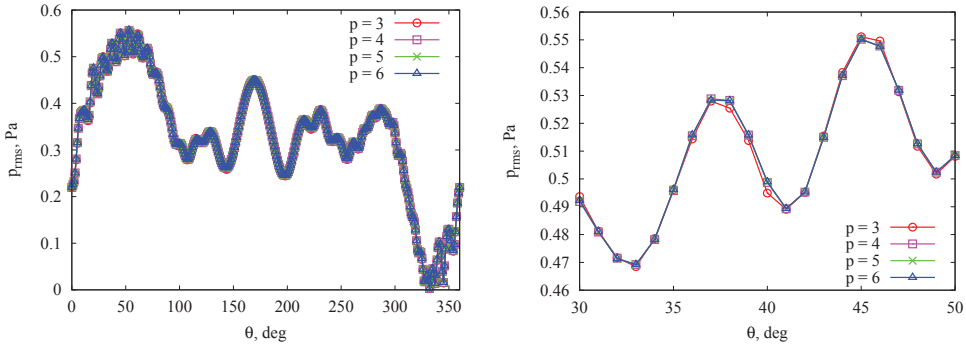


Fig. 7. Directivity of instantaneous pressure RMS values for interpolation polynomial of order $p = 3, 4, 5, 6$: (a) directivity on the whole circle; (b) close-up for $\theta \in [30, 50]$

While recent research programs brought to significant progress in reducing both the turbomachinery noise generation and the radiation of noise from the intake, there is still a lack of knowledge about the exhaust noise radiation problem and the need to develop accurate models for its prediction. This problem represents a challenge for CAA, due to the fact that radiated sound propagates through the shear layers separating core, bypass and free-stream fields.

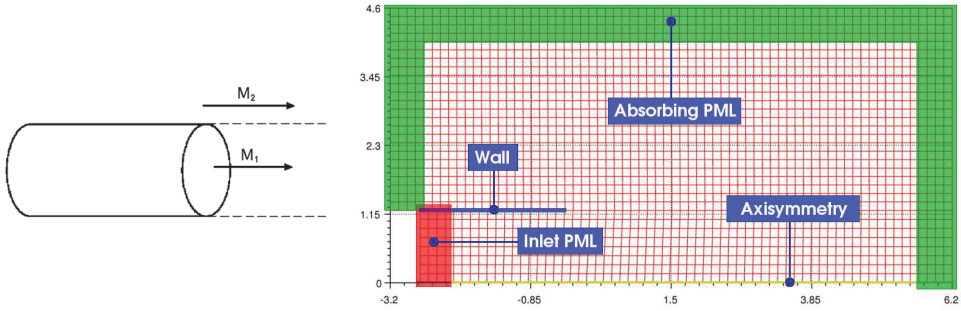
4.3.1 Circular duct propagation and radiation

In order to validate the propagation model, an idealized case for the propagation of sound waves inside the exhaust nozzle is studied. This idealized problem studies the propagation of a sound wave inside a semi-infinite circular cylinder and the subsequent radiation of the

wave outside the pipe. This idealized problem is the acoustic diffraction by a sound wave propagating out of a rigid semi-infinite cylindrical duct (figure 8(a)). The radius of the cylinder is equal to $r_1 = 1.212$ m and the duct wall has no thickness with acoustically rigid inner and outer surfaces. Inside the duct there is a uniform axial mean flow of density ρ_2, Mach M_2 and speed of sound c_2. In the outer region the flow is also uniform and axial, with density ρ_1, Mach M_1 and speed of sound c_1. There is no shear layer between the two flows, instead they are separated by a vortex sheet. For this problem, termed "Munt problem", the analytical solution has been found by Munt (1977) and it has been subsequently generalized for annular ducts and lined walls [Demir & Rienstra (2006); Gabard & Astley (2006); Rienstra (1984)]. Since the analytical solution is available only for points at great distance from the cylinder exit, it is not possible to compare directly the solution of LEE obtained using the DGM with the analytical one. Instead LEE are solved only in a small computational domain, the near-field domain, and then the far-field solution is evaluated for the near-field one using the three-dimensional integral formulation of the wave equation proposed by Ffowcs Williams and Hawkings [Iob et al. (2010)]. The far-field results are then compared with the analytical solution. The LEE computational domain extends for $z \in [-2.5\text{ m}; 5.5\text{ m}]$ and for $r \in [0.0\text{ m}; 3.9\text{ m}]$ and is surrounded by vertical and horizontal PML layers with a thickness of 0.7 m. This domain is discretized using a uniform structured grid with about $5,000$ rectangular elements with $p = 3$. The discretization and the boundary conditions for the near-field domain are shown in figure 8(b).

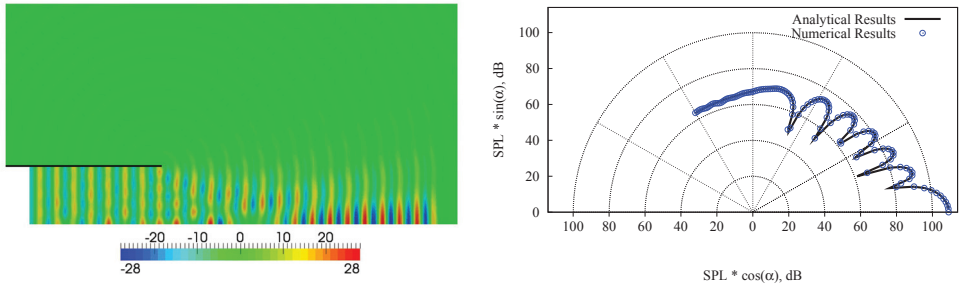
The first case considered is the "no-flow" condition: the mean flow is assumed to be at rest both inside and outside of the duct, with a speed of sound equal to $c_1 = c_2 = 340.17\text{ m/s}$ and a density equal to $\rho_1 = \rho_2 = 1.225\text{ Kg/m}^3$. In figure (9(a)) the instantaneous near-field pressure field for a plain wave, mode $(0, 1)$, at a frequency of 956 Hz is reported, whereas figure (9(b)) displays the corresponding Sound Pressure Level (SPL) directivity pattern in the far-field. The directivity is evaluated on an arc having the center defined at the center of the duct exit section and radius equal to $r = 46$ m. It can be seen that the agreement between the numerical and the analytical solutions is very good.

To study the effect of the mean flow, another flow condition is analyzed: in this condition there is a mean flow velocity inside the duct, with Mach number equal to 0.447, whereas outside the duct the fluid is at rest. In figure 10(a) and 10(b), instantaneous pressure field and directivity



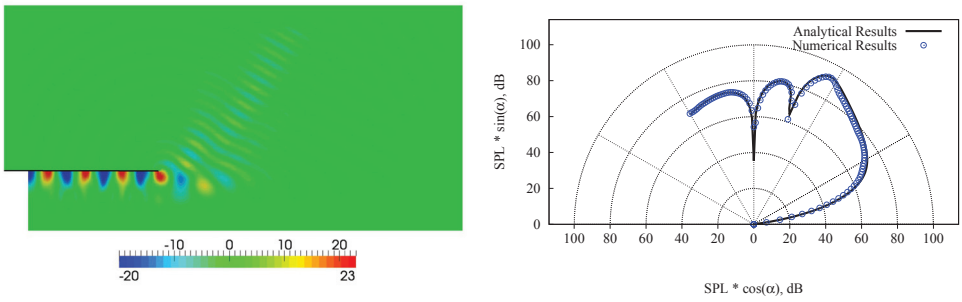
(a) (b)

Fig. 8. Circular duct; (a) geometry; (b) domain discretization and boundary conditions



(a) (b)

Fig. 9. Circular duct, $M_1 = 0.0$, $M_2 = 0.0$, mode $(0, 1)$, $f = 956\,Hz$; (a) near-field solution: real part of Fourier pressure coefficient, Pa; (b) far-field solution: SPL directivity



(a) (b)

Fig. 10. Circular duct, $M_1 = 0.0$, $M_2 = 0.447$, mode $(9, 1)$, $f = 866\,Hz$; (a) near-field solution: real part of Fourier pressure coefficient, Pa; (b) far-field solution: SPL directivity

for the mode $(9, 1)$, frequency 866 Hz are presented. Also in this case there is a good agreement with the analytical solution.
Computational times for all the conditions is about 2 minutes on an AMD Athlon 64 X2 Dual Core Processor and calculations require less than 1Gb of RAM.

4.3.2 Engine exhaust propagation and radiation

To study a more realistic test case, the propagation of duct modes inside a turbofan engine exhaust is considered. Numerical simulations on this geometry are carried out for two flow conditions. In the first condition, termed condition A, the mean flow is at rest with a speed of sound equal to $c_\infty = 340.17$ m/s and a total density equal to $\rho_{t\,\infty} = 1.225$ Kg/m^3. In the second one, condition B, the external flow is at rest, $c_\infty = 340.17$ m/s and $\rho_\infty = 1.225$ Kg/m^3, whereas there is a mean velocity both inside the by-pass duct and inside the core duct. Flow properties at duct's inlet are the following: $c_{\text{fan}} = 353.15$ m/s, $\rho_{\text{t fan}} = 1.327$ Kg/m^3, and $M_{\text{fan}} = 0.35$ for the by-pass duct and $c_{\text{turb}} = 508.75$ m/s, $\rho_{\text{t turb}} = 0.598$ Kg/m^3, and $M_{\text{turb}} = 0.29$ for the core duct.
Like for the "Munt" problem, LEE are solved only in the near-field domain, and then the far-field solution is evaluated using the Ffowcs Williams and Hawkings formulation. The

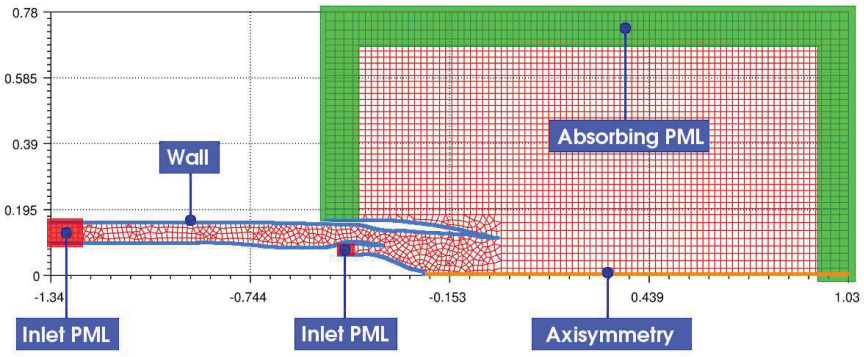
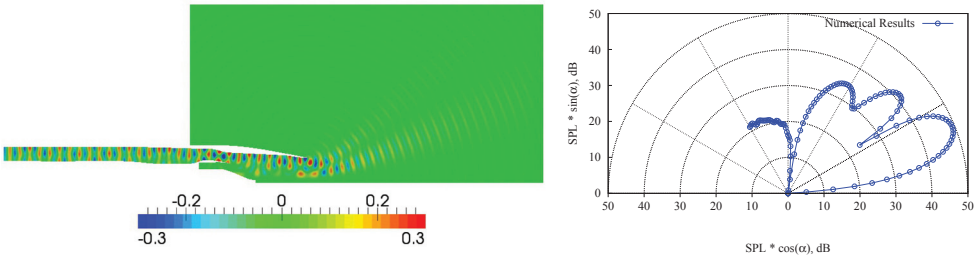


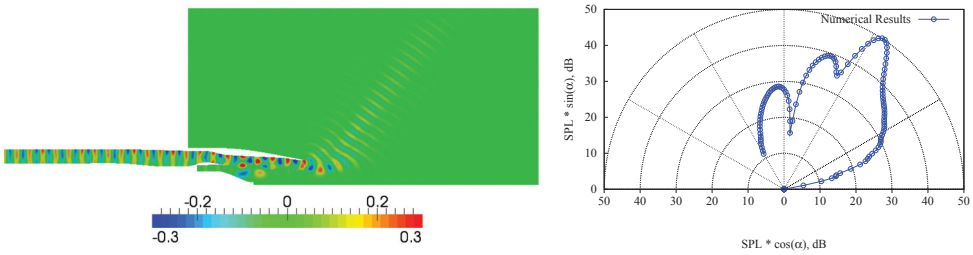
Fig. 11. Engine exhaust, domain discretization and boundary conditions



(a) (b)

Fig. 12. Engine exhaust, flow condition A, mode $(4,\,1)$, $f = 7981.25\,Hz$; (a) near-field solution: real part of Fourier pressure coefficient, Pa; (b) far-field solution: SPL directivity

near-field domain is discretized with an unstructured grid of $3,800$ elements and, in the external region, it extends for $z \in [-0.44\text{ m};\, 0.95\text{ m}]$ and for $r \in [0.0\text{ m};\, 0.7\text{ m}]$ and is surrounded by vertical and horizontal PML domains with a thickness of 0.08 m (see figure 11). For all the presented cases the incoming wave enters from the by-pass duct and consists in the duct mode $(4,1)$ at a frequency of 7981.25 Hz, which corresponds to a dimensionless frequency of $kr_{\text{fan}} = 16.22$ where $r_{\text{fan}} = 0.11$ m is the radius of the by-pass duct at the exit plane. Elements are discretized with interpolation polynomials of degree $p = 3$, therefore the discretization has about 60000 nodes. Near-field instantaneous pressure perturbations for the flow condition A are shown in figure 13(a), whereas the corresponding SPL directivity pattern is shown in figure 13(b).The far-field directivity is evaluated on an arc of radius $r = 12$ m having the center placed at the center of the by-pass duct exit section. Same results are presented for the flow condition B in figure 13(a) and figure 13(b). It can be shown that the presence of the flow field, in particular the presence of the shear layers between core flow, by-pass flow and external region, deflects the sound waves in the upward direction. When the flow is at rest the maximum far-field SPL level is located at a directivity angle of 21°, whereas this maximum moves at an angle of 55° when there is a mean flow inside the by-pass and the core duct.



(a) (b)

Fig. 13. Engine exhaust, flow condition B, mode $(4,\,1)$, $f = 7981.25\,Hz$; (a) near-field solution: real part of Fourier pressure coefficient, Pa; (b) far-field solution: SPL directivity

5. Conclusions

A numerical model based on the solution of the LEE in the frequency domain with a discontinuous Galerkin method has proved to be a valuable tool for the analysis of the propagation of sound waves. With the LEE it is possible to study acoustic wave propagation in the presence of rotational mean flows. The frequency domain approach suppress the Kelvin-Helmholtz instability waves which pollute LEE solutions in time-domain calculations. The static condensation technique greatly reduces the memory requirements of the DGM and make feasible the frequency domain approach. Moreover each single calculation, limited to a single frequency, is well suited to the exhaust noise radiation problem where the incoming wave can be treated as a superposition of the elementary duct modes. The model is well suited for design and optimization processes. The model has been successfully validated with the analytical solution of the Munt problem. In the case of realistic configurations, the numerical results reproduce the main expected features.

6. Appendix

A. Duct mode theory

According to the duct mode theory the acoustic field inside a duct may be expressed by means of a series expansion in a particular family of solutions, called duct modes. Duct modes are interesting mathematically because they form a complete basis by which any solution can be represented. Nonetheless, modes are also interesting from a more physical point of view because they are solutions in their own right, not just mathematical building blocks. The simple structure of the modes can be therefore useful to understand the usually complicated behavior of the total field.

Consider an infinite duct with an arbitrary cross section with inside a uniform mean flow with Mach number equal to M_0 and speed of sound equal to c_0. The duct is aligned with the z-axis and its cross section, $\mathcal{A}$, is on the $(x,\,y)$-plane (figure 14). Assuming a harmonic time dependency, the acoustic field can be written as

$$p'\left(\boldsymbol{x},\,t\right) \equiv \hat{p}\left(\boldsymbol{x}\right) \mathrm{e}^{\mathcal{I}\omega t}, \qquad \boldsymbol{v}'\left(\boldsymbol{x},\,t\right) \equiv \hat{\boldsymbol{v}}\left(\boldsymbol{x}\right) \mathrm{e}^{\mathcal{I}\omega t}, \tag{67}$$

where ω is the angular frequency and the Fourier modes $\hat{p}$ and $\hat{\boldsymbol{v}}$ satisfy continuity and momentum equations

$$\left(\mathcal{I}\frac{\omega}{c_0} + M_0\frac{\partial}{\partial z}\right)\hat{p} + c_0^2\rho_0\nabla\cdot\hat{\boldsymbol{v}} = 0 \tag{68a}$$

$$\left(\mathcal{I}\frac{\omega}{c_0} + M_0\frac{\partial}{\partial z}\right)\hat{\boldsymbol{v}} + \frac{1}{\rho_0 c_0}\nabla\hat{p} = 0\,. \tag{68b}$$

The assumption of time-harmonic dependence is not so restrictive. If the acoustic field is not harmonic, the linearity of sound waves allows to build up the acoustic field as a sum of time harmonic components via a Fourier analysis. Duct walls are considered acoustically rigid, i.e. they completely reflects the acoustic waves

$$\frac{\partial\hat{p}}{\partial n} = 0 \quad \text{for } \boldsymbol{x}\in\partial\mathcal{A}\,, \tag{69}$$

where n is the normal to the wall pointing into the surface. From Eqs. (68) it is possible to obtain the convected Helmholtz equation for a uniform mean flow. In cylindrical coordinates this equation reads

$$\left(\mathcal{I}\frac{\omega}{c_0} + \boldsymbol{M_0}\cdot\nabla\right)^2\hat{p} - \left(\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)\hat{p} = 0\,, \tag{70}$$

where $\boldsymbol{M_0} = [M_0,\,0,\,0]$. The solution of Eq. (70) may be given by

$$\begin{aligned}\hat{p}\,(r,\theta,z) &= \sum_{n=0}^{\infty}\left[\hat{p}_n^+\,(r,\theta,z) + \hat{p}_n^-\,(r,\theta,z)\right]\\ &= \sum_{n=0}^{\infty}\left[C_n^+\psi_n\,(r,\theta)\,\mathrm{e}^{-\mathcal{I}k_n^+ z} + C_n^-\psi_n\,(r,\theta)\,\mathrm{e}^{-\mathcal{I}k_n^- z}\right]\,,\end{aligned} \tag{71}$$

where the function $\psi_n\,(r,\theta)\,\mathrm{e}^{-\mathcal{I}k_n^\pm z}$ defines the shape of the n-th so-called duct mode, with a $+$ for right-running modes and a $-$ for left running ones. Each mode is characterized by an amplitude $C_n^\pm$ and by an axial wave number $k_n^\pm$. The functions ψ_n are the eigenfunctions of the Laplace operator reduced to the cross section $\mathcal{A}$, i.e., solutions of

$$\begin{aligned}-\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right]\psi &= \alpha^2\psi \qquad \text{for } (r,\theta)\in\mathcal{A}\\ \frac{\partial\psi}{\partial n} &= 0 \qquad \text{for } (r,\theta)\in\partial\mathcal{A}\,,\end{aligned} \tag{72}$$

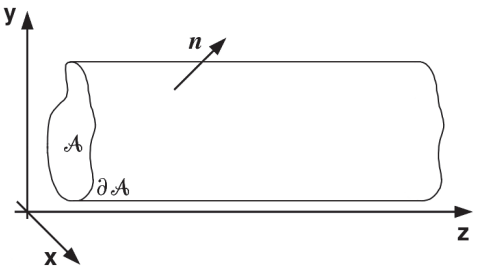


Fig. 14. Infinite duct with arbitrary cross section $\mathcal{A}$

where α^2 is the corresponding eigenvalue. If the duct cross section is annular and the boundary condition is uniform everywhere, the solution of the eigenvalue problem is relatively simple and may be obtained by separation of variables. Considering an annular duct with inner radius equal to h and outer radius equal to a, the eigensolutions consist of combinations of exponential and Bessel functions. In cylindrical coordinates this yields

$$\psi_n(r,\theta) = U_{m\mu}(r)\, \mathrm{e}^{\mathcal{I}m\theta}, \tag{73}$$

where $m = 0, \pm 1, \pm 2, \ldots$ is the azimuthal index of the mode and $\mu = 1, 2, \ldots$ is the radial one. Positive values of m correspond with clockwise rotating modes and negative m with counter-clockwise rotating modes. The radial function $U_{m\mu}$ is obtained from

$$U''_{m\mu} + \frac{1}{r}U'_{m\mu} + \left(\alpha^2_{m\mu} - \frac{m^2}{r^2}\right) U_{m\mu} = 0, \tag{74}$$

where $U''_{m\mu}$ and $U'_{m\mu}$ are, respectively, the second and the first order derivative of $U_{m\mu}(r)$. Solving Eq (74), the radial shape of the modes can be evaluated as

$$U_{m\mu}(r) = J_m(\alpha_{m\mu} r) - \frac{J'_m(\alpha_{m\mu} a)}{Y'_m(\alpha_{m\mu} a)} Y_m(\alpha_{m\mu} r), \tag{75}$$

where J_m and Y_m denote the Bessel functions of the first and second kind of order m, whereas J'_m and Y'_m are their derivatives. Eigenvalues $\alpha_{m\mu} = \alpha_n$ are evaluated via boundary conditions, imposing that $U'_{m\mu}(a) = U'_{m\mu}(h) = 0$, the value of $\alpha_{m\mu}$ corresponds to the μ-th zero of the equation

$$J'_m(\alpha_{m\mu} a)\, Y'_m(\alpha_{m\mu} h) - J'_m(\alpha_{m\mu} h)\, Y'_m(\alpha_{m\mu} a) = 0. \tag{76}$$

Although technically speaking $\{\alpha^2_{m\mu}\}$ are the eigenvalues of minus the cross-sectional Laplace operator, it is common practice to refer to $\alpha^2_{m\mu}$ as the radial eigenvalue or radial modal wave number, to m as the circumferential eigenvalue or circumferential wave number, and to $k^{\pm}_n = k^{\pm}_{m\mu}$ as the corresponding axial eigenvalue or axial wave number. The radial and axial wave numbers satisfy

$$\alpha^2_{m\mu} = \left(\omega/c_0 - M_0 k_{m\mu}\right)^2 - k^2_{m\mu}, \tag{77}$$

$$k^{\pm}_{m\mu} = \frac{-\omega M_0/c_0 \pm \sqrt{\omega^2/c_0^2 - \left(1 - M_0^2\right)\alpha^2_{m\mu}}}{1 - M_0^2}. \tag{78}$$

For an annular duct mode the general solution may therefore be written as

$$\hat{p}(z,r,\theta) = \sum_{m=-\infty}^{\infty} \sum_{\mu=1}^{\infty} \left(C^{+}_{m\mu} \mathrm{e}^{-\mathcal{I}k^{+}_{m\mu} z} + C^{-}_{m\mu} \mathrm{e}^{-\mathcal{I}k^{-}_{m\mu} z}\right) U_{m\mu}(r)\, \mathrm{e}^{\mathcal{I}m\theta}, \tag{79}$$

where $C^{\pm}_{m\mu}$ is the amplitude of the mode, $U_{m\mu}\left(r, \alpha_{m\mu}\right)$ is given by Eq. (75), the radial wave number $\alpha_{m\mu}$ is the μ-th zero of Eq. (76) and the axial wave number $k^{\pm}_{m\mu}$ is given by Eq. (78).

A.1 Cut-off frequency

The z-axis pressure dependence of a duct mode takes two completely different forms depending on whether the driving frequency is below or above a critical value. The reason of this radical change in transmission properties as the driving frequency is gradually swept through this critical value is due to the fact that the nature of the z-variation of the pressure field in a duct behaves, see Eq. (71), like the term $e^{-\mathrm{i}k_{m\mu}z}$. When $k_{m\mu}$ is purely imaginary, the amplitude of the pressure fluctuations are subjected to an exponential decay (because $\mathrm{Im}\left(k_{m\mu}\right) < 0$). Usually the rate of decay is large enough to reduce the pressure intensity to a negligible value in a short distance compared to the duct radius. When $k_{m\mu}$ is real, true wave motion propagates in the duct. From the definition of $k_{m\mu}$, Eq. (78), it is easy to see that $k_{m\mu}$ is purely imaginary if $\left(1-M_0^2\right)\alpha_{m\mu}^2 > \omega^2/c_0^2$. The frequency f_c which tells apart this two different behaviors is called *cut-off frequency*

$$f_c = \frac{\alpha_{m\mu} c_0 \sqrt{1-M_0^2}}{2\pi}\,. \tag{80}$$

Cut-off frequency depends on the geometry of the duct, on the mode of the incident wave and also on the mean flow properties. For any ω there are always a (finite) $\mu = \mu_0$ and $m = m_0$ beyond which $\left(1-M_0^2\right)\alpha_{m\mu}^2 > \omega^2/c_0^2$, so that $k_{m\mu}$ is purely imaginary, and the mode decays exponentially in z. Therefore there is always a finite number of modes with real $k_{m\mu}$. Since they are the only modes that propagate, they are called cut-on. The remaining infinite number of modes, with imaginary $k_{m\mu}$, are evanescent and therefore called cut-off.

7. Acknowledgments

This work was supported by the Regione Piemonte (Italy) through the GREAT 2020 project.

8. References

Agarwal, A., Morris, P. & Mani, R. (2004). The calculation of sound propagation in nonuniform flows: Suppression of instability waves, *AIAA Journal* 42(1): 80–88.

Amestoy, P., Guermouche, A., L'Excellent, J.-Y. & Pralet, S. (2006). Hybrid scheduling for the parallel solution of linear systems, *Parallel Computing* 32(2): 136–156.

Bayliss, A. & Turkel, E. (1980). Radiation boundary conditions for wave-like equations, *Commun. on Pure Appl. Math.* 33: 708–725.

Berenger, J. P. (1994). A perfectly matched layer for the absorption of electromagnetic waves, *Journal of Computational Physics* 114: 185–200.

Bernacki, M., Fezoui, L., Lanteri, S. & Piperno, S. (2006). Parallel discontinuous Galerkin unstructured mesh solvers for the calculation of three-dimensional wave propagation problems, *Applied Mathematical Modelling* 30: 744–763.

Bodony, D. J. (2006). Analysis of sponge zones for computational fluid mechanics, *Journal of Computational Physics* 212: 681–702.

Cockburn, B., Karniadakis, G. & Shu, C. (2000). *Discontinuous Galerkin Methods. Theory, Computation and Applications*, Springer-Verlag.

Crighton, D. G. (1975). Basic principles of aerodynamic noise generation, *Progress in Aerospace Sciences* 16(1): 31 – 96.

Demir, A. & Rienstra, S. (2006). Sound radiation from an annular duct with jet flow and a lined centerbody, *12th AIAA/CEAS Aeroacoustics Conference; Cambridge, Massachusetts;* AIAA Paper *2006-2718*.

Gabard, G. & Astley, R. (2006). Theoretical model for sound radiation from annular jet pipes: Far- and near-field solutions, *J. Fluid. Mech.* 549: 315–341.

Giles, M. B. (1990). Non-reflecing boundary conditions for Euler equation calculations, *AIAA Journal* 28(12): 2050–2058.

Goldstein, M. (1976). *Aeroacoustics*, McGraw-Hill.

Hesthaven, J. S. (1998). From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex, *SIAM Journal on Numerical Analysis* 35: 655–676.

Hu, F. Q. (1996). On absorbing boundary conditions of linearized Euler equations by a perfectly matched layer, *Journal of Computational Physics* 129: 201–219.

Hu, F. Q. (2001). A stable, perfectly matched layer for linearized Euler equations in unsplit physical variables, *Journal of Computational Physics* 173: 455–480.

Hu, F. Q. (2004). Absorbing boundary conditions (a review), *Int J Comput Fluid Dyn* 18(6): 513–522.

Iob, A., Arina, R. & Schipani, C. (2010). A frequency-domain linearized Euler model for turbomachinery noise radiation through engine exhaust, *AIAA Journal* 48(4): 848–858.

Karniadakis, G. E. & Sherwin, S. J. (2005). *Spectral/hp Element Methods for Computational Fluid Dynamics (Numerical Mathematics and Scientific Computation)*, Oxford University Press, USA.

Munt, R. (1977). The interaction of sound with a subsonic jet issuing from a semi-infinite cylindrical pipe, *J. Fluid Mechanics* 83(4): 609–640.

Özyörük, Y. (2009). Numerical prediction of aft radiation of turbofan tones through exhaust jets, *Journal of Sound and Vibration* 325(1-2): 122 – 144.

Rao, P. & Morris, P. (2006). Use of finite element methods in frequency domain aeroacoustics, *AIAA Journal* 44(7): 1643–1652.

Reed, W. & Hill, T. (1973). Triangular mesh methods for the neutron transport equation, *Technical Report LA-UR-73-479*, Los Alamos Scientific Laboratory.

Rienstra, S. W. (1984). Acoustic radiation form a semi-infinte annular duct in a uniform subsonic mean flow, *J. of Sound and Vibration* 94(2): 267–288.

Scott, E. & Sherer (2004). Acoustic scattering from multiple circular cylinders: Category 2, problems 1 and 2, analytic solution, *Proceedings of the 4th CAA Workshop on Benchmark Problems, NASA CP-2004-212954.*

Tam, C. K. W., Auriault, L. & Cambulli, F. (1998). Perfectly matched layers as an absorbing boundary condition for the linearized Euler equations in open and ducted domains, *Journal of Computational Physics* 144: 213–234.

Tam, C. & Webb, J. (1993). Dispersion-Relation-Preserving finite differnce schemes for computational acoustics, *Journal of Computational Physics* 107: 262–281.

Tester, B. J., Gabard, G. & Özyörük, Y. (2008). Influence of mean flow gradients on fan exhaust noise predictions, *14th AIAA/CEAS Aeroacoustics Conference 5-7 May 2008 Vancouver, British Columbia Canada;* AIAA Paper *2008-2825*.

Thompson, K. W. (1990). Time-dependent boundary conditions for hyperbolic systems II, *Journal of Computational Physics* 89: 439–461.

Toro, E. F. (1999). *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*, 2nd edn, Springer, Berlin.

Zhang, X., Chen, X. X., Morfey, C. L. & Tester, B. J. (2003). Computation of fan noise radiation through a realistic engine exhaust geometry with flow, *9th AIAA/CEAS Aeroacoustics Conference 12-14 May 2003 Hilton Head, South Carolina, United States;* AIAA Paper *2003-3267*.

7

High-Order Numerical Methods for BiGlobal Flow Instability Analysis and Control

Javier de Vicente[1], Daniel Rodríguez[2], Leo González[3] and Vassilis Theofilis[4]
[1]*Department of Applied Mathematics and Statistic. School of Aeronautics, Universidad Politécnica de Madrid*
[2]*School of Aeronautics, Universidad Politécnica de Madrid*
[3]*School of Naval Engineering, Universidad Politécnica de Madrid*
[4]*School of Aeronautics, Universidad Politécnica de Madrid*
Spain

1. Introduction

In recent years flow instability and flow control research has focused attention on two novel and promising areas, investigation of perturbations in the limit of short times after their introduction into the flow Schmid & Henningson (2001) and study of modal and non-modal perturbations in complex, essentially non–parallel flows. A rather complete, up to that time, account of the rapidly growing latter area was discussed in Collis et al. (2004); Theofilis (2003); the role of (especially global) instability analysis in flow control is discussed elsewhere Theofilis (2009b). The present article deals with the theory and numerical aspects underlying the recent rapid developments in global instability research, in an attempt to generate a self–contained account of the areas in which the authors have been working in the last decade, as opposed to producing a review-type article on the many developments which have recently taken place.

When discussing numerical methods for the solution of the partial–derivative eigenvalue problem it is instructive to remind the reader of the process by which one arrives at the various large–scale eigenvalue problems solved in a global instability context; pictorially, this process is shown in the charts presented in figure 1. In the first of these decision trees, one is confronted with the temporal derivative in the linearized Navier–Stokes equations (LNSE) and the possibility to either discretize this term (in a so–called "time-stepping" approach), thereby solving for arbitrary temporal development of the perturbations (including transient growth) or work in frequency space by introducing eigenmodes, a procedure which is permissible by the separability of the temporal and spatial derivatives in the LNSE. Once this decision has been made, the next step is to deal with how to treat the LNSE matrix; there too, two paths may be followed, one ("Time-steppers") along which the matrix describing the evolution of perturbations is not formed explicitly and tools analogous with those used in direct numerical simulations are employed, and another one (Generalized Eigenvalue Problem – GEVP) along which the matrix (of leading dimension potentially reaching 1 Tb Kitsios et al. (2008)) is stored in (shared or distributed) memory. From the outset the question

may be posed as to which is the more appropriate path to be followed. The answer depends on whether one is interested in recovering only a few or many members of the eigenspectrum, time–steppers (alongside high–quality numerics) being appropriate for the first task, while solution to the GEVP provides access to a potentially large window of eigenmodes.

The chart of figure 1 ,in the lower case, in principle describes spatial discretization approaches applicable to both time–steppers and GEVP solutions, although the present article will be devoted almost exclusively to the latter. As is known from classic linear instability analysis Mack (1984) accuracy is of prime concern also in global instability work. This need not translate to the necessity for high–order numerical methods for the spatial discretization of the LNSE operators, but such methods are increasingly more convenient to use as the Reynolds number increases, the structures to be resolved become increasingly tighter and the corresponding resolution requirements become more stringent. Nevertheless, successful global instability analyses have been reported in the literature, using both the low– and the high–order methods shown in this chart. To the best of the authors' knowledge, the first works featuring numerical methods shown in the lower case in figure 1 are discussed next.

In the context of methods of low formal order of accuracy, finite–difference methods were used by Crouch *et al.* Crouch et al. (2007) in a shock–capturing context and by Giannetti and Luchini Giannetti & Luchini (2007) in conjunction with an immersed–boundary approach. The respective methods dealt successfully with flow instability around a wing at realistic flight conditions and low–Reynolds number flow in the classic cylinder wake. Kuhlmann and co–workers (e.g. Albensoeder et al. (2001a;b)) and Sahin and Owens Sahin & Owens (2003) have used standard finite–volume methods in order to analyze the stability of lid–driven cavity flow Theofilis (AIAA-2000-1965). Marquet *et al.* Marquet et al. (2006) have employed low–order finite–element methods to the solution of the stability and control problems in an internal (S-shaped duct) flow, while González *et al.* González et al. (2007) have employed methods of the same class to the solution of global stability problems in both internal (rectangular duct and triangular cavity) and open (bluff–body wake) flows.

Turning to high–order methods, Henningson Henningson (1987) and Tatsumi and Yoshimura Tatsumi & Yoshimura (1990) were the first to introduce Fourier and Chebyshev (single–domain) spectral collocation methods, respectively, into inviscid and viscous global linear instability analysis. Spectral collocation has been the main workhorse of global instability research in the last decade; see Theofilis Theofilis (2011) for a review. De Vicente *et al.* de Vicente et al. (2006) and Robinet Robinet (2007) have discussed multidomain spectral collocation methodologies for the solution of incompressible (open cavity) and compressible (shock–induced laminar separation bubble) global flow stability problems, respectively. Barkley and Henderson Barkley & Henderson (1996) first employed Floquet theory and a structured spectral–element method to identify bifurcations in the wake of a circular cylinder, while Theofilis *et al.* Theofilis et al. (2002) were the first to introduce spectral/hp methods Karniadakis & Sherwin (2005) in order to address stability of flow in the wake of a NACA–0012 airfoil by solution of the BiGlobal eigenvalue problem. Abdessemed *et al.* Abdessemed et al. (2004; 2006) extended the spectral/hp methodology to the study of time–periodic flows in the wake of a low–pressure turbine. Finally, compact finite–difference methods have been introduced into the study of aeroacoustic and hydrodynamic global instability problems by Theofilis and Colonius Theofilis & Colonius (2003; 2004) and Bres and Colonius Bres & Colonius (2008), in the context of analysis of compressible laminar flow over an open cavity.

Short of describing in–depth the many successes of global instability theory, the main objective of the present article is to familiarize the reader with the concepts of this analysis methodology and present in detail techniques which the authors have employed in the course of their research in the past decade. The intention is not only to generate a self–contained manuscript, in which sufficient detail is presented for newcomers to the field to be able to construct their own solvers, but also to expose implementation aspects which may condition decisions on which method is to be chosen for the numerical solution of the eigenvalue problem. Section 2 presents separately the equations governing the incompressible and compressible limits, since the size of the associated problems makes inefficient the use of a single (compressible) code for the flow instability analysis at all Mach numbers. In addition, attention is paid to the adjoint eigenvalue problem, which underlies theoretically–founded flow control methodologies, and the the compressible Rayleigh equation, which may be used for computational aeroacoustics work. Section 3 discusses in–depth finite–element, spectral collocation (in both their single– and multi–domain flavor) and spectral/hp methods for the spatial discretization of the BiGlobal eigenvalue problem. A short section 4 presents the Arnoldi algorithm invariably used for the recovery of an arbitrarily–large window of eigenvalues, while a selection of results obtained by the authors is shown in section 5. A brief discussion in section 6 closes the presentation.

2. The linearized Navier-Stokes equations

The analysis of flow instability is based on the compressible equations of motion

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{1}$$

$$\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u}\mathbf{u}) = \frac{-1}{\gamma M^2} \nabla p + \frac{1}{Re} \nabla \cdot \boldsymbol{\text{œ}}, \tag{2}$$

$$\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p + \gamma p \nabla \cdot \mathbf{u} = \frac{\gamma}{RePr} \nabla \cdot (\kappa \nabla T) + \frac{\gamma(\gamma - 1) M^2}{Re} \Phi, \tag{3}$$

where,

$$\boldsymbol{\text{œ}} = \mu \left[(\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \frac{2}{3} (\nabla \cdot \mathbf{u}) \mathbf{I} \right]$$

and

$$\Phi = \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T) : \boldsymbol{\text{œ}}$$

are the viscous stress tensor and the dissipation function, respectively, and Stokes' hypothesis is invoked. The equation of state of perfect gases,

$$p = \rho T,$$

valid up to hypersonic Mach numbers, is used in order to close this system of equations; under these conditions $\gamma = 7/5$ and $Pr = 0.72$ is taken to be a constant. In line with the analogous assumption of classic linear theory Mack (1984), the viscosity is assumed to be a function of temperature alone,

$$\mu = \mu(T).$$

Central to linear flow instability research is the concept of decomposition of any flow quantity into an $O(1)$ steady or time–periodic laminar *basic* flow upon which small–amplitude three–dimensional disturbances are permitted to develop. The most general framework in which a linear instability analysis can be performed is one in which three inhomogeneous spatial directions are resolved and time–periodic small–amplitude disturbances, *inhomogeneous* in all three directions, are superimposed upon the underlying steady or time-periodic $O(1)$ basic state. The related three-dimensional global *TriGlobal* instability Ansatz (named according to the dimensionality of the basic stateTheofilis (2003)) yields a three-dimensional eigenvalue problem in which all three spatial directions must be resolved simultaneously in a coupled manner. Though this most general Ansatz is consistent with the separability in the governing equations of time on the one hand and the three spatial directions on the other, the size of the resulting EVP is such that currently available computing hardware and algorithms permit its solution in a very limited range of Reynolds numbers, of $Re \approx O(10^2)$. A time–stepping approach for the TriGlobal EVP is discussed in another contribution to the present volume.

2.1 BiGlobal linear theory: the compressible two-dimensional linear EVP

In order to proceed in the context of global linear theory, the basic state is assumed independent of one spatial coordinate, say z, in a Cartesian framework. Flow quantities are then decomposed according to

$$\mathbf{q}(x,y,z,t) = \bar{\mathbf{q}}(x,y) + \varepsilon\, \tilde{\mathbf{q}}(x,y,z,t), \tag{4}$$

with $\bar{\mathbf{q}} = (\bar{u},\bar{v},\bar{w},\bar{T},\bar{p})^{\mathrm{T}}$ and $\tilde{\mathbf{q}} = (\tilde{u},\tilde{v},\tilde{w},\tilde{T},\tilde{p})^{\mathrm{T}}$ representing the steady *two-dimensional* basic flow and the unsteady *three-dimensional* infinitesimal perturbations, respectively, the latter being inhomogeneous in x and y and periodic in z. Note also that, unlike the incompressible case which will be discussed shortly, pressure is a predictive variable in, rather than a constraint of, the equations of motion. On substituting (4) into the governing equations (1-3), taking $\varepsilon \ll 1$ and linearizing about $\bar{\mathbf{q}}$, one may write

$$\tilde{\mathbf{q}}(x,y,z,t) = \hat{\mathbf{q}}(x,y)\, \mathrm{e}^{\mathrm{i}\,\Theta_{2D}} + c.c., \tag{5}$$

with $\hat{\mathbf{q}} = (\hat{u},\hat{v},\hat{w},\hat{\theta},\hat{p})^{\mathrm{T}}$ representing the vector of *two–dimensional* complex amplitude functions of the infinitesimal three-dimensional perturbations, ω a complex eigenvalue and

$$\Theta_{2D} = \beta z - \omega t \tag{6}$$

a complex phase function.

The linear disturbance equations of BiGlobal stability analysis are obtained at $O(\varepsilon)$ by substituting the decomposition (4-6) into the equations of motion, subtracting out the $O(1)$ basic flow terms and neglecting terms at $O(\varepsilon^2)$. In the present temporal framework, β is

taken to be a real wavenumber parameter describing an eigenmode in the $z-$direction, while the complex eigenvalue ω, and the associated eigenvectors $\hat{\mathbf{q}}$ are sought. The real part of the eigenvalue, $\omega_r \equiv \Re\{\omega\}$, is related with the frequency of the global eigenmode while the imaginary part is its growth/damping rate; a positive value of $\omega_i \equiv \Im\{\omega\}$ indicates exponential growth of the instability mode $\tilde{\mathbf{q}} = \hat{\mathbf{q}} e^{i\,\Theta_{2D}}$ in time t while $\omega_i < 0$ denotes decay of $\tilde{\mathbf{q}}$ in time. The system for the determination of the eigenvalue ω and the associated eigenfunctions $\hat{\mathbf{q}}$ in its most general form can be written as the complex nonsymmetric generalized EVP

$$\mathcal{L}\,\hat{\mathbf{q}} = \omega \mathcal{R}\,\hat{\mathbf{q}}, \tag{7}$$

or, more explicitly,

$$\begin{pmatrix} \mathcal{L}_{x\hat{u}} & \mathcal{L}_{x\hat{v}} & \mathcal{L}_{x\hat{w}} & \mathcal{L}_{x\hat{\theta}} & \mathcal{I}\mathcal{L}_{x\hat{p}} \\ \mathcal{L}_{y\hat{u}} & \mathcal{L}_{y\hat{v}} & \mathcal{L}_{y\hat{w}} & \mathcal{L}_{y\hat{\theta}} & \mathcal{I}\mathcal{L}_{y\hat{p}} \\ \mathcal{L}_{z\hat{u}} & \mathcal{L}_{z\hat{v}} & \mathcal{L}_{z\hat{w}} & \mathcal{L}_{z\hat{\theta}} & \mathcal{I}\mathcal{L}_{z\hat{p}} \\ \mathcal{L}_{e\hat{u}} & \mathcal{L}_{e\hat{v}} & \mathcal{L}_{e\hat{w}} & \mathcal{L}_{e\hat{\theta}} & \mathcal{I}\mathcal{L}_{e\hat{p}} \\ \mathcal{J}\mathcal{L}_{c\hat{u}} & \mathcal{J}\mathcal{L}_{c\hat{v}} & \mathcal{J}\mathcal{L}_{c\hat{w}} & \mathcal{J}\mathcal{L}_{c\hat{\theta}} & \mathcal{L}^G_{c\hat{p}} \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{w} \\ \hat{\theta} \\ \hat{p} \end{pmatrix} = \omega \begin{pmatrix} \mathcal{R}_{x\hat{u}} & 0 & 0 & 0 & 0 \\ 0 & \mathcal{R}_{y\hat{v}} & 0 & 0 & 0 \\ 0 & 0 & \mathcal{R}_{z\hat{w}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathcal{I}\mathcal{R}_{e\hat{p}} \\ 0 & 0 & 0 & \mathcal{J}\mathcal{R}_{c\hat{\theta}} & \mathcal{R}^G_{c\hat{p}} \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{w} \\ \hat{\theta} \\ \hat{p} \end{pmatrix}, \tag{8}$$

subject to appropriate boundary conditions. Here the linearized equation of state

$$\hat{p} = \hat{\rho}/\bar{\rho} + \hat{\theta}/\bar{T}$$

has been used, viscosity and thermal conductivity of the medium have been taken as functions of temperature alone, resulting in

$$\hat{\mu} = \frac{d\bar{\mu}}{dT}\hat{\theta},\ \hat{\kappa} = \frac{d\bar{\kappa}}{dT}\hat{\theta}.$$

Moreover,

$$\mathcal{I} = I^G_{GL},\ \mathcal{J} = I^{GL}_G$$

are interpolation arrays transferring data from the Gauss–Lobatto to the Gauss and from the Gauss to the Gauss-Lobatto spectral collocation grids, respectively. Details on the spectral collocation spatial discretization will be provided in § 3.1. All submatrices of matrix $\mathcal{L}$ are defined on a two–dimensional Chebyshev Gauss-Lobatto (CGL) grid, except for $\mathcal{L}^G_{c\hat{p}}$ and $\mathcal{R}^G_{c\hat{p}}$, which are defined on a two–dimensional Chebyshev Gauss (CG) grid.

2.2 Classic linear theory: the one-dimensional compressible linear EVP

It is instructive at this point to compare the theory based on solution of (7) against results obtained by use of the established classic theory of linear instability of boundary- and shear-layer flows (cf. MackMack (1984), MalikMalik (1991)). The latter theory is based on the Ansatz

$$\mathbf{q}(x,y,z,t) = \bar{\mathbf{q}}(y) + \varepsilon\,\hat{\mathbf{q}}(y)\,e^{i\,\Theta_{1D}} + c.c. \tag{9}$$

In (9) $\hat{\mathbf{q}}$ is the vector of *one–dimensional* complex amplitude functions of the infinitesimal perturbations and ω is in general complex. The phase function, Θ_{1D}, is

$$\Theta_{1D} = \alpha x + \beta z - \omega t, \tag{10}$$

where α and β are wavenumber parameters in the spatial directions x and z, respectively, underlining the wave-like character of the linear perturbations in the context of the one-dimensional EVP.
Substitution of the decomposition (9-10) into the governing equations (1-3) linearization and consideration of terms at $O(\varepsilon)$ results in the eigenvalue problem governing linear stability of boundary- and shear-layer flows; the same system results directly from (7) if one makes the following ("parallel flow") assumptions:

- $\partial\bar{q}/\partial x \equiv 0$, $\partial\bar{q}/\partial x \equiv 0$
  (basic flow independent of x),
  i.e. $\partial\hat{q}/\partial x \equiv \mathrm{i}\,\alpha\,\hat{q}$, $\partial\hat{q}/\partial z \equiv \mathrm{i}\,\beta\,\hat{q}$
  (harmonic expansion of disturbances in x and z),
- $\bar{v} \equiv 0$, and
- $\bar{p} \equiv cnst.$,

then (7) takes the form of the system of equations governing linear stability of viscous compressible boundary- and shear-layer flows (cf. eqns. (8.9) of Mack Mack (1984)). None of these approximations are necessary in the context of BiGlobal theory, but invoking the parallel flow assumption in the latter context provides direct means for comparisons between the present (relatively) novel and the established methodologies. Such comparisons have been performed, e.g. by Theofilis and Colonius Theofilis & Colonius (2004). It should be noted that the crucial difference between the two–dimensional eigenvalue problem (7) and the limiting case of the one–dimensional EVP is that the eigenvector $\hat{q}$ in (7) comprises *two–dimensional* amplitude functions, while those in the limiting parallel–flow case are *one–dimensional*. Further, while $\bar{p}(y) = cnst.$ in taken to be a constant in one-dimensional basic states satisfying (9), $\bar{p}(x,y)$ appearing in (7) is, in general, a known function of the two resolved spatial coordinates.

2.3 The compressible BiGlobal Rayleigh equation

Linearizing the viscous compressible equations of motion neglecting the viscous terms in (8) and introducing the elliptic confocal coordinate system Morse & Feshbach (1953) for reasons which will become apparent later leads to the generalized Rayleigh equation on this coordinate system,

$$\begin{aligned}\hat{p}_{\xi\xi} + \hat{p}_{\eta\eta} - h^2\beta^2\hat{p} + \frac{j_1^2 + j_2^2}{h^2}\left[\left(\frac{\bar{p}_\xi}{\gamma\bar{p}} - \frac{\bar{\rho}_\xi}{\bar{\rho}}\right) - \frac{2\beta\bar{w}_\xi}{(\beta\bar{w}-\omega)}\right]\hat{p}_\xi \\ + \frac{j_1^2 + j_2^2}{h^2}\left[\left(\frac{\bar{p}_\eta}{\gamma\bar{p}} - \frac{\bar{\rho}_\eta}{\bar{\rho}}\right) - \frac{2\beta\bar{w}_\eta}{(\beta\bar{w}-\omega)}\right]\hat{p}_\eta \\ + \left[\frac{\bar{\rho}\,(\beta\bar{w}-\omega)^2}{\gamma\bar{p}}\right]\hat{p} = 0. \qquad (11)\end{aligned}$$

Since the metrics of the elliptic confocal coordinate system satisfy $j_1^2 + j_2^2 = h^2$, one finally has to solve

$$\mathcal{M}\hat{p} + \left[\left(\frac{\bar{p}_\xi}{\gamma\bar{p}} - \frac{\bar{\rho}_\xi}{\bar{\rho}}\right) - \frac{2\beta\bar{w}_\xi}{(\beta\bar{w}-\omega)}\right]\hat{p}_\xi + \left[\left(\frac{\bar{p}_\eta}{\gamma\bar{p}} - \frac{\bar{\rho}_\eta}{\bar{\rho}}\right) - \frac{2\beta\bar{w}_\eta}{(\beta\bar{w}-\omega)}\right]\hat{p}_\eta + \left[\frac{\bar{\rho}\,(\beta\bar{w}-\omega)^2}{\gamma\bar{p}}\right]\hat{p} = 0. \tag{12}$$

where the linear operator $\mathcal{M} \equiv \partial_{\xi\xi} + \partial_{\eta\eta} - h^2\beta^2$. In a manner analogous with classic one-dimensional linear theory Mack (1984); Malik (1991), (12) may be solved either iteratively or by direct means. In view of the lack of any prior physical insight into global linear disturbances in the application at hand, a direct method is preferable on account of the access to the full eigenvalue spectrum that it provides. Either the temporal or the spatial form of the eigenvalue problem may be solved at the same level of numerical effort using a direct method since, in both cases, a cubic eigenvalue problem must be solved. In its temporal form, the temporal global inviscid instability problem reads

$$T_1\,\hat{p}_{\xi\xi} + T_2\,\hat{p}_{\eta\eta} + T_3\,\hat{p}_\xi + T_4\,\hat{p}_\eta + T_5\,\hat{p} = \omega\left(T_6\,\hat{p}_{\xi\xi} + T_7\,\hat{p}_{\eta\eta} + T_8\,\hat{p}_\xi + T_9\,\hat{p}_\eta + T_{10}\,\hat{p}\right) + \omega^2\,T_{11}\,\hat{p} + \omega^3\,T_{12}\,\hat{p} \tag{13}$$

while the spatial generalized Rayleigh equation on the elliptic confocal coordinate system is

$$S_1\,\hat{p}_{\xi\xi} + S_2\,\hat{p}_{\eta\eta} + S_3\,\hat{p}_\xi + S_4\,\hat{p}_\eta + S_5\,\hat{p} = \beta\left(S_6\,\hat{p}_{\xi\xi} + S_7\,\hat{p}_{\eta\eta} + S_8\,\hat{p}_\xi + S_9\,\hat{p}_\eta + S_{10}\,\hat{p}\right) + \beta^2\,S_{11}\,\hat{p} + \beta^3\,S_{12}\,\hat{p}. \tag{14}$$

In the incompressible limit, equation (11) reduces to that solved by Henningson Henningson (1987), while in the absence of flow (and its derivatives) altogether, (11) simplifies in the classic two-dimensional Helmholtz problem

$$\left(\partial_{\xi\xi} + \partial_{\eta\eta} + \kappa^2\right)\hat{p} = 0, \tag{15}$$

which has been recently employed extensively to the solution of resonance problems Hein et al. (2007); Koch (2007).

2.4 The incompressible limit

Since most global instability analysis work performed to-date has been in an incompressible flow context, this limit will now be described in a little more detail. The equations governing incompressible flows may be directly deduced from (1-3) and are written in

primitive-variables formulation

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 u_i}{\partial x_j^2} \qquad \text{in} \qquad \Omega, \tag{16}$$

$$\frac{\partial u_i}{\partial x_i} = 0 \qquad \text{in} \qquad \Omega. \tag{17}$$

Here Ω is the computational domain, u_i represents the velocity field, p is the pressure field, t is the time and x_i represent the spatial coordinates. This domain is limited by a boundary Γ where different boundary conditions can be imposed depending on the problem and the numerical dicretization. The primitive variable formulation is preferred over the alternative velocity-vorticity form, simply because the resulting system comprises four- as opposed to six equations which need to be solved in a coupled manner.
The two-dimensional equations of motion are solved in the laminar regime at appropriate Re regions, in order to compute steady real basic flows $(\bar{u}_i, \bar{p})$ whose stability will subsequently be investigated. The basic flow equations read

$$\frac{\partial \bar{u}_i}{\partial t} + \bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 \bar{u}_i}{\partial x_j^2} \qquad \text{in} \qquad \Omega, \tag{18}$$

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \qquad \text{in} \qquad \Omega \tag{19}$$

The steady laminar basic flow is obtained by time-integration of the system (18-19) starting from rest until the steady state is obtained. The convergence in time of the steady basic flow must be $O(10^{-12})$ to make it adequate for the linear analysis. In case of unsteady laminar or turbulent flows, one may analyze time-averaged mean flows. Finally, in the particular case of addressing laminar flow over a symmetric body, steady flows can be obtained by forcing a symmetry condition along the line of geometric symmetry.
The basic flow is perturbed by small-amplitude velocity u_i^* and kinematic pressure p^* perturbations, as follows

$$u_i = \bar{u}_i + \varepsilon u_i^* + c.c. \qquad p = \bar{p} + \varepsilon p^* + c.c., \tag{20}$$

where $\varepsilon \ll 1$ and $c.c.$ denotes conjugate of the complex quantities (u_i^*, p^*). Substituting into equations (16-17), subtracting the basic flow equations (18-19), and linearizing, the incompressible Linearized Navier-Stokes Equations (LNSE) for the perturbation quantities are obtained

$$\frac{\partial u_i^*}{\partial t} + \bar{u}_j \frac{\partial u_i^*}{\partial x_j} + u_j^* \frac{\partial \bar{u}_i}{\partial x_j} = -\frac{\partial p^*}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 u_i^*}{\partial x_j^2}, \tag{21}$$

$$\frac{\partial u_i^*}{\partial x_i} = 0. \tag{22}$$

2.5 Time-marching

The initial condition for (21-22) must be inhomogeneous in order for a non-trivial solution to be obtained. In view of the homogeneity along one spatial direction, $x_3 \equiv z$, the most general form assumed by the small amplitude perturbations satisfies the following Ansatz

$$u_i^* = \hat{u}_i(x,y,t)e^{i\beta z} \tag{23}$$

$$p^* = \hat{p}(x,y,t)e^{i\beta z}, \tag{24}$$

where $i = \sqrt{-1}$, β is a wavenumber parameter, related with a periodicity length L_z along the homogeneous direction through $L_z = 2\pi/\beta$, $(\hat{u}_i, \hat{p})$ are the complex amplitude functions of the linear perturbations and *c.c.* denotes complex conjugates, introduced so that the LHS of equations (23-24) be real. Note that the amplitude functions may, at this stage, be arbitrary functions of time.

Substituting (23) and (24) into the equations (21) and (22), the equations may be reformulated as

$$\frac{\partial \hat{u}_1}{\partial t} + \bar{u}_j \frac{\partial \hat{u}_1}{\partial x_j} + \hat{u}_j \frac{\partial \bar{u}_1}{\partial x_j} = -\frac{\partial \hat{p}}{\partial x} + \frac{1}{Re}\left(\frac{\partial^2}{\partial x_j^2} - \beta^2\right)\hat{u}_1, \tag{25}$$

$$\frac{\partial \hat{u}_2}{\partial t} + \bar{u}_j \frac{\partial \hat{u}_2}{\partial x_j} + \hat{u}_j \frac{\partial \bar{u}_2}{\partial x_j} = -\frac{\partial \hat{p}}{\partial y} + \frac{1}{Re}\left(\frac{\partial^2}{\partial x_j^2} - \beta^2\right)\hat{u}_2, \tag{26}$$

$$\frac{\partial \hat{u}_3}{\partial t} + \bar{u}_j \frac{\partial \hat{u}_3}{\partial x_j} + \hat{u}_j \frac{\partial \bar{u}_3}{\partial x_j} = -i\beta\hat{p} + \frac{1}{Re}\left(\frac{\partial^2}{\partial x_j^2} - \beta^2\right)\hat{u}_3, \tag{27}$$

$$\frac{\partial \hat{u}_1}{\partial x} + \frac{\partial i\hat{u}_2}{\partial y} + i\beta\hat{u}_3 = 0. \tag{28}$$

This system may be integrated along time by numerical methods appropriate for the spatial discretization scheme utilized. The result of the time-integration at $t \to \infty$ is the leading eigenmode of the steady basic flow. In this respect, time-integration of the linearized disturbance equations is a form of power iteration for the leading eigenvalue of the system. Alternative, more sophisticated, time-integration approaches, well described by Karniadakis and Sherwin Karniadakis & Sherwin (2005) are also available for the recovery of both the leading and a relatively small number of additional eigenvalues. The key advantage of time-marching methods, over explicit formation of the matrix which describes linear instability, is that the matrix need never be formed. This enables the study of global linear stability problems on (relatively) small-main-memory machines at the expense of (relatively) long–time integrations. To–date this is the only viable approach to perform TriGlobal instability analysis. A potential pitfall of the time-integration approach is that results are sensitive to the quality of spatial integration of the linearized equations, such that this approach should preferably be used in conjunction with high-order spatial discretization methods; see Karniadakis & Sherwin (2005) for a discussion. The subsequent discussion will be exclusively focused on approaches in which the matrix is formed.

2.6 Matrix formation – the incompressible direct and adjoint BiGlobal EVPs

Starting from the (direct) LNSE (21-22) and assuming modal perturbations and homogeneity in the spanwise spatial direction, z, eigenmodes are introduced into the linearized direct Navier-Stokes and continuity equations according to

$$(\mathbf{q}^*, p^*) = (\hat{\mathbf{q}}(x,y), \hat{p}(x,y))\mathrm{e}^{+\mathrm{i}\,(\beta z - \omega t)}, \tag{29}$$

where $\mathbf{q}^* = (u^*, v^*, w^*)^{\mathrm{T}}$ and p^* are, respectively, the vector of amplitude functions of linear velocity and pressure perturbations, superimposed upon the steady two-dimensional, two- ($\bar{w} \equiv 0$) or three-component, $\bar{\mathbf{q}} = (\bar{u}, \bar{v}, \bar{w})^T$, steady basic states. The spanwise wavenumber β is associated with the spanwise periodicity length, L_z, through $L_z = 2\pi/L_z$. Substitution of (29) into (21-22) results in the complex *direct* BiGlobal eigenvalue problem Theofilis (2003)

$$\hat{u}_x + \hat{v}_y + i\beta\hat{w} = 0, \tag{30}$$

$$(\mathcal{L} - \bar{u}_x + \mathrm{i}\,\omega)\,\hat{u} - \bar{u}_y\hat{v} - \hat{p}_x = 0, \tag{31}$$

$$-\bar{v}_x\hat{u} + \left(\mathcal{L} - \bar{v}_y + \mathrm{i}\,\omega\right)\hat{v} - \hat{p}_y = 0, \tag{32}$$

$$-\bar{w}_x\hat{u} - \bar{w}_y\hat{v} + (\mathcal{L} + i\omega)\,\hat{w} - i\beta\hat{p} = 0, \tag{33}$$

where

$$\mathcal{L} = \frac{1}{Re}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \beta^2\right) - \bar{u}\frac{\partial}{\partial x} - \bar{v}\frac{\partial}{\partial y} - i\beta\bar{w}. \tag{34}$$

The concept of the adjoint eigenvalue problem has been introduced in the context of receptivity and flow control respectively by Zhigulev and Tumin Zhigulev & Tumin (1987) and Hill Hill (1992). The derivation of the complex BiGlobal eigenvalue problem governing adjoint perturbations is constructed using the Euler-Lagrange identity Bewley (2001); Dobrinsky & Collis (2000); Giannetti & Luchini (2007); Morse & Feshbach (1953); Pralits & Hanifi (2003),

$$\begin{aligned}\left[\left(\frac{\partial\hat{\mathbf{q}}^*}{\partial t} + \mathcal{N}\hat{\mathbf{q}}^* + \nabla\hat{p}^*\right)\cdot\tilde{\mathbf{q}}^* + \nabla\cdot\hat{\mathbf{q}}^*\tilde{p}^*\right] + \\ \left[\hat{\mathbf{q}}^*\cdot\left(\frac{\partial\tilde{\mathbf{q}}^*}{\partial t} + \mathcal{N}^\dagger\tilde{\mathbf{q}}^* + \nabla\tilde{p}^*\right) + \hat{p}^*\nabla\cdot\tilde{\mathbf{q}}^*\right] = \\ \frac{\partial}{\partial t}\left(\hat{\mathbf{q}}^*\cdot\tilde{\mathbf{q}}^*\right) + \nabla\cdot j(\hat{\mathbf{q}}^*, \tilde{\mathbf{q}}^*),\end{aligned} \tag{35}$$

as applied to the linearized incompressible Navier-Stokes and continuity equations. Here the operator $\mathcal{N}^\dagger(\bar{\mathbf{q}})$ results from linearization of the convective and viscous terms in the direct and adjoint Navier-Stokes equations and is explicitly stated elsewhere (e.g. Dobrinsky & Collis (2000)). The quantities $\tilde{\mathbf{q}}^* = (\tilde{u}^*, \tilde{v}^*, \tilde{w}^*)^{\mathrm{T}}$ and $\tilde{p}^*$ denote adjoint disturbance velocity components and adjoint disturbance pressure, and $j(\hat{\mathbf{q}}^*, \tilde{\mathbf{q}}^*)$ is the bilinear concomitant. Vanishing of the RHS term in the Euler-Lagrange identity (35) defines the adjoint linearized incompressible Navier-Stokes and continuity equations

$$\frac{\partial\tilde{\mathbf{q}}^*}{\partial t} + \mathcal{N}^\dagger\tilde{\mathbf{q}}^* + \nabla\tilde{p}^* = 0, \tag{36}$$

$$\nabla\cdot\tilde{\mathbf{q}}^* = 0, \tag{37}$$

Assuming modal perturbations and homogeneity in the spanwise spatial direction, z, eigenmodes are introduced into (36-37) according to

$$(\tilde{\mathbf{q}}^*, \tilde{p}^*) = (\tilde{\mathbf{q}}(x,y), \tilde{p}(x,y))\mathrm{e}^{-\mathrm{i}\,(\beta z - \omega t)}. \tag{38}$$

Note the opposite signs of the spatial direction z and time in (29) and (38), denoting propagation of $\tilde{\mathbf{q}}^*$ in the opposite directions compared with the respective one for $\hat{\mathbf{q}}^*$. Substitution of (38) into the adjoint linearized Navier-Stokes equations (36-37) results in the complex *adjoint* BiGlobal EVP

$$\tilde{u}_x + \tilde{v}_y - \mathrm{i}\,\beta\tilde{w} = 0, \tag{39}$$

$$\left(\mathcal{L}^\dagger - \bar{u}_x + \mathrm{i}\,\omega\right)\tilde{u} - \bar{v}_x\tilde{v} - \bar{w}_x\tilde{w} - \tilde{p}_x = 0, \tag{40}$$

$$-\bar{u}_y\tilde{u} + \left(\mathcal{L}^\dagger - \bar{v}_y + \mathrm{i}\,\omega\right)\tilde{v} - \bar{w}_y\tilde{w} - \tilde{p}_y = 0, \tag{41}$$

$$\left(\mathcal{L}^\dagger + \mathrm{i}\,\omega\right)\tilde{w} + \mathrm{i}\,\beta\tilde{p} = 0, \tag{42}$$

where

$$\mathcal{L}^\dagger = \frac{1}{Re}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \beta^2\right) + \bar{u}\frac{\partial}{\partial x} + \bar{v}\frac{\partial}{\partial y} - \mathrm{i}\,\beta\bar{w}. \tag{43}$$

Note also that, in the particular case of two–component two–dimensional basic states, i.e. $(\bar{u} \neq 0, \bar{v} \neq 0, \bar{w} \equiv 0)^{\mathrm{T}}$ such as encountered, f.e. in the lid–driven cavity Theofilis (AIAA-2000-1965) and the laminar separation bubble Theofilis et al. (2000), both the direct and adjoint EVP may be reformulated as *real EVPs* Theofilis (2003); Theofilis, Duck & Owen (2004), thus saving half of the otherwise necessary memory requirements for the coupled numerical solution of the EVPs (30-33) and (39-42).

Boundary conditions for the partial-derivative adjoint EVP in the case of a closed system are particularly simple, requiring vanishing of adjoint perturbations at solid walls, much like the case of their direct counterparts. In open systems containing boundary layers, adjoint boundary conditions may be devised following the general procedure of expanding the bilinear concomitant in order to capture traveling disturbances Dobrinsky & Collis (2000). When the focus is on global modes concentrated in certain regions of the flow, as the case is, for example, for the global mode of laminar separation bubble (Theofilis (2000); Theofilis et al. (2000)) the following procedure may be followed. For the direct problem, homogeneous Dirichlet boundary conditions are used at the inflow, $x = x_{IN}$, wall, $y = 0$, and far-field, $y = y_\infty$, boundaries, alongside linear extrapolation at the outflow boundary $x = x_{OUT}$. Consistently, homogeneous Dirichlet boundary conditions at $y = 0$, $y = y_\infty$ and $x = x_{OUT}$, alongside linear extrapolation from the interior of the computational domain at $x = x_{IN}$, are used in order to close the adjoint EVP.

Once the eigenvalue problem has been stated, the objective becomes its numerical solution in any of its compressible viscous (8), inviscid (11), or incompressible (30-33) direct or adjoint forms. Any of these eigenvalue problems is a system of coupled partial-differential equations for the determination of the eigenvalues, ω, and the associated sets of amplitude functions, $\hat{\mathbf{q}}$. Intuitively one sees that, when the matrix is formed, resolution/memory requirements will be the main concern of any numerical solution approach and this is indeed the case in all

but the smallest (and least interesting) Reynolds number values. The following discussion is devoted to this point and is divided in two parts, one devoted to the spatial discretization of the PDE-based EVP and one dealing with the subspace iteration method used for the determination of the eigenvalue.

3. Numerical discretization – weighted residual methods

The approximation of a function u as an expansion in terms of a sequence of orthogonal functions, is the starting point of many numerical methods of approximation. Spectral methods belong to the general class of weighted residuals methods (WRM). These methods assume that a solution of a differential equation can be approximated in terms of a truncated series expansion, such that the difference between the exact and approximated solution (residual), is minimized.
Depending on the set of base (trial) functions used in the expansion and the way the error is forced to be zero several methods are defined. But before starting with the classification of the different types of WRM it is instructive to present a brief introduction to vector spaces.
Define the set,

$$\mathcal{L}^2_w(I) = \{v : I \to \mathcal{R} |\ v \text{ is measurable and } \| v \|_{o,w} < \infty\}$$

where $w(x)$ denotes a weight function, i.e., a continuous, strictly positive and integrable function over the interval $I = (-1, 1)$ and

$$\| v \|_w = \left(\int_{-1}^{1} |v(x)|^2 w(x) dx \right)^{1/2}$$

is the norm induced by the scalar product

$$(u, v)_w = \int_{-1}^{1} u(x) v(x) w(x) dx$$

Let $\{\varphi_n\}_{n \geq 0} \in \mathcal{L}^2_w(I)$ denote a system of algebraic polynomials, which are mutually orthogonal under the scalar product defined before.

$$(\varphi_n, \varphi_m)_w = 0 \text{ whenever } m \neq n$$

Using the Weierstrass approximation theorem every continuous function included in $\mathcal{L}^2_w(-1, 1)$ can be uniformly approximated as closely as desired by a polynomial expansion, i.e. for any function u the following expansion holds

$$u(x) = \sum_{k=0}^{\infty} \hat{u}_k \varphi_k(x) \text{ with } \hat{u}_k = \frac{(u, \varphi_k)_w}{\| \varphi_k \|^2_{0,w}} \tag{44}$$

The $\hat{u}_k$ are the expansion coefficients associated with the basis $\{\varphi_n\}$, defined as

$$\hat{u}_k = \frac{1}{\| \varphi_k \|^2_{0,w}} \int_{-1}^{1} u(x) \varphi_k(x) w(x) dx \tag{45}$$

Consider now the truncated series of order N

$$u_N(x) = \sum_{k=0}^{N} \hat{u}_k \varphi_k(x)$$

$u_N(x)$ is the orthogonal projection of u upon the span of $\{\varphi_n\}$.
Due to the completeness of the system $\{\varphi_n\}$, the truncated series converges in the sense of $\forall u \in \mathcal{L}^2_w(I)$

$$\| u - u_N \|_w \to 0 \text{ as } N \to \infty$$

Now the residual could be defined as

$$\mathcal{R}_N(x) = u - u_N$$

In the weighted residuals methods the goal of annulling $\mathcal{R}_N$ is reached in an approximate sense by setting to zero the scalar product

$$(\mathcal{R}_N, \phi_i)_{\hat{w}} = \int_{-1}^{1} \mathcal{R}_N \phi_i(x) \hat{w}(x) dx$$

where ϕ_i are test functions and $\hat{w}$ is the weight associated with the trial function.

A first and main classification of the different WRM is done depending on the choice of the trial functions φ_i. Finite Difference and Finite Element methods use overlapping local polynomials as base functions.
In Spectral Methods, however, the trial functions are global functions, typically tensor products of the eigenfunctions of singular Sturm-Liouville problems. Some well–known examples of these functions are: Fourier trigonometric functions for periodic– and Chebyshev or Legendre polynomials for nonperiodic problems.

Focusing on the Spectral Methods and attending to the residual, a second distinction could be:

- Galerkin approach: This method is characterized by the choice $\phi_i = \varphi_i$ and $\hat{w} = w$. Therefore, the residual

  $$\mathcal{R}_N(x) = u - u_N = u - \sum_{k=0}^{N} \hat{u}_k \varphi_k(x)$$

  is forced to zero in the mean according to

  $$(\mathcal{R}_N, \varphi_i)_w = \int_{-1}^{1} \left(u - \sum_{k=0}^{N} \hat{u}_k \varphi_k \right) \varphi_i w dx = 0 \quad i = 0, ..., N.$$

These $N+1$ Galerkin equations determine the coefficients $\hat{u}_k$ of the expansion.

- Collocation approach: The test functions are Dirac delta-functions $\phi_i = \delta(x - x_i)$ and $\hat{w} = 1$.
  The collocation points x_i, are selected as will be discussed later.Now, the residual

$$\mathcal{R}_N(x) = u - u_N = u - \sum_{k=0}^{N} \hat{u}_k \varphi_k(x)$$

  is made equal zero at the $N+1$ collocation points, $u(x_i) - u_N(x_i)$, hence,

$$\sum_{k=0}^{N} \hat{u}_k \varphi_k(x) = u(x_i)$$

  This gives an algebraic system to determine the $N+1$ coefficients $\hat{u}_k$.

- Tau approach: It is a modification of the Galerkin approach allowing the use of trial functions not satisfying the boundary conditions; it will not be discussed in the present context.

3.1 Spectral collocation methods

In the general framework of Spectral Methods the approximation of a function u is done in terms of global polynomials. Appropriate choices for non-periodic functions are Chebyshev or Legendre polynomials, while periodic problems may be treated using the Fourier basis. The exposition that follows will be made on the basis of the Chebyshev expansion only.

3.1.1 Collocation approximation

The Chebyshev polynomials of the first kind $T_k(x)$ are the eigenfunctions of the singular Sturm-Liouville problem

$$\begin{cases} -(pu')' + qu = \lambda w u & \text{in the interval } (1, -1) \\ \text{plus boundary conditions for } u \end{cases}$$

where $p(x) = (1-x^2)^{1/2}$, $q(x) = 0$ and $w(x) = (1-x^2)^{-1/2}$. The problem is reduced to

$$(\sqrt{1-x^2}T_k'(x))' + \frac{k^2}{\sqrt{1-x^2}} T_k(x) = 0$$

For $x \in [-1, 1]$ an important characterization is given by

$$T_k(x) = \cos k\theta \quad \text{with } \theta = \arccos x$$

One of the main features of the Chebyshev polynomials is the orthogonality relationship, Chebyshev family is orthogonal in the Hilbert space $\mathcal{L}^2_w[-1,1]$, with the weight $w(x) = (1-x^2)^{-1/2}$.

$$(T_k, T_l)_w = \int_{-1}^{1} T_k(x)T_l(x)w(x)dx = \frac{\pi}{2}c_k\delta_{k,l}$$

where $\delta_{k,l}$ is the Kronecker delta and c_k is defined as:

$$c_k = \begin{cases} 2 & \text{if } k = 0 \\ 1 & \text{if } k \geq 1 \end{cases}$$

The spectral Chebyshev representation of any function $u(x)$ defined for $x \in [-1,1]$ is its orthogonal projection on the space of polynomials of degree $\leq N$:

$$u_N(x) = \sum_{k=0}^{N} \hat{u}_k T_k(x)$$

in the collocation method the expansion coefficients are calculated so the residual is setting to zero at the collocation points. The choice of such points is not arbitrary, depends on the quadrature formulas for integration used and the characteristics of the problem tackled:

Chebyshev-Gauss points,

$$x_i = \cos\frac{(2i+1)\pi}{2N+2}; \quad i = 0,..,N \tag{46}$$

are the roots of the Chebyshev polynomial T_{N+1}, and are related to the Gauss integration in $(-1,1)$.

Chebyshev-Gauss-Lobatto points

$$x_i = \cos\frac{i\pi}{N} \quad i = 0,..,N \tag{47}$$

are the points where T_N reaches its extremal values. Gauss-Lobatto nodes are related to the Gauss-Lobatto integration and include the end points ± 1, useful if there is a requirement of imposing boundary conditions.

As mentioned the technique consists of setting to zero the residual $R_N = u - u_N$ at the collocation points x_i, $i = 0,..,N$, so

$$\sum_{k=0}^{N} \hat{u}_k T_k(x_i) = u(x_i), i = 0,...N.$$

This gives an algebraic system to determine the $N+1$ coefficients $\hat{u}_k, k = 0,...,N$. The existence of a solution implies that $detT_k(x_i) \neq 0$.
As a matter of fact the expression for the coefficients can be found without solving the system, this is done using the discrete orthogonality property of the basis functions. From the Gauss quadrature formula,

$$\int_{-1}^{1} pwdx \cong \frac{\pi}{N}\sum_{i=0}^{N}\frac{p(x_i)}{\hat{c}_i}$$

where

$$\hat{c}_i = \begin{cases} 2 & \text{if } i = 0, N \\ 1 & \text{if } i = 1, ..., N-1 \end{cases}$$

This relation is exact when $p(x)$ is a polynomial of degree less than $2N$, so using the Chebyshev-Gauss-Lobatto collocation points and since $T_k T_l$ is a polynomial of degree at most $2N-1$ the discrete orthogonality property is deduced:

$$\int_{-1}^{1} T_k(x)T_l(x)w(x)dx = \frac{\pi}{N}\sum_{i=0}^{N}\frac{1}{\hat{c}_i}T_k(x_i)T_l(x_i)$$

therefore

$$\sum_{i=0}^{N}\frac{1}{\hat{c}_i}T_k(x_i)T_l(x_i) = \frac{\hat{c}_k}{2}N\delta_{k,l}$$

Now, multiplying each side of (the equation for residual), by $T_l(x_i)/\hat{c}_i$, summing from $i = 0$ to $i = N$, and using the discrete orthogonality relation, the next expression for the collocation coefficients is obtained:

$$\hat{u}_k = \frac{2}{\hat{c}_k N}\sum_{-1}^{1}\frac{1}{\hat{c}_i}u(x_i)T_k(x_i)\ , k = 0, ..., N. \tag{48}$$

It must be noted that such expression is nothing but the numerical approximation of the integral form. The grid values $u(x_i)$ and the expansion coefficients $\hat{u_k}$ are related by truncated discrete Fourier series in cosine, so it is possible to use the Fast Fourier Transform (FFT) to connect the physical space to the spectral space.

From other point of view the expression for the approximation of a function using the collocation technique at the Chebyshev-Gauss-Lobatto points,

$$u_N(x) = \sum_{k=0}^{N}\hat{u}_k T_k(x)$$

could be seen as the Lagrange interpolation polynomial of degree N based on the set x_i. Hence it can be written in the form

$$u_N(x) = \sum_{k=0}^{N} h_k u(x_j) \tag{49}$$

where the Lagrange functions $h_j \in \mathcal{P}_{\mathcal{N}}$ are such that $h_j(x_k) = \delta_{jk}$ and are defined by

$$h_j(x) = \frac{(-1)^{j+1}(1-x^2)T_N'(x)}{\hat{c}_j N^2 (x - x_j)}$$

This expression for the approximation does not involve the spectral coefficients, the unknowns are the grid values what makes it useful for expressing the derivatives at any collocation point in terms of the grid values of the function.

$$(\partial_N u)(x_j) = \sum_{k=0}^{N} h_k'(x_j) u(x_k),\ j = 0, ..., N. \tag{50}$$

The matrix $(D_N)_{ij} = h_j'(x_i)$ is named Chebyshev pseudo-spectral matrix and its entries can be computed explicitly,

$$(D_N)_{ij} = \begin{cases} \frac{\hat{c}_i}{\hat{c}_j} \frac{(-1)^{i+j}}{(x_i - x_j)} & \text{if } 0 \le i, j \le N, i \ne j \\ -\frac{x_i}{2(1-x_i^2)} & \text{if } 1 \le i = j \le N-, i \ne j \\ \frac{2N^2+1}{6} & \text{if } 0 = i = j \\ -\frac{2N^2+1}{6} & \text{if } i = j = N \end{cases} \tag{51}$$

In vector form the derivatives may be expressed as

$$U' = \mathcal{D}U \approx \begin{pmatrix} d_{0,0} & d_{0,1} & \dots & d_{0,N} \\ d_{1,0} & d_{1,1} & \dots & d_{1,N} \\ & & \ddots & \\ d_{N,0} & d_{N,1} & \dots & d_{N,N} \end{pmatrix} \begin{pmatrix} U_0 \\ U_1 \\ \\ U_N \end{pmatrix} = (U_0', U_1', \dots, U_N') \tag{52}$$

Second derivatives may be computed explicitly, although it is useful to recall that $U'' = \mathcal{D}(\mathcal{D}U) = \mathcal{D}^2 U$

3.1.2 Mappings

Expansion in Chebyshev polynomials of functions defined on other finite intervals from $[-1, 1]$ are required not only owing to geometric demands but also when the function has regions of rapid change, boundary layers, singularities and so on. Mappings can be useful in improving the accuracy of a Chebyshev expansion.

But not any choice of the collocation points x_i is appropriate, the polynomial approximation on them does not necessarily converge when $N \to \infty$.

If $x \in [-1, 1]$ the coordinate transformation $y = f(x)$ must meet some requirements. It must be one-to-one, easy to invert and at least C^1. So, let

$$A = [a, b] \text{ with } y \in A$$

the physical space, and f the mapping in the form

$$y = f(x)$$

The approximation of a function u in $A = [a,b]$ can be easily done assuming $u_N(y) = u_N(f(x)) = v_N(x)$. The Chebyshev expansion:

$$u_N(y) = \sum_{k=0}^{N} \hat{v}_k T_k(x)$$

Now the goal is to represent the derivative of u in terms of its values in y_N. From elementary calculus

$$\frac{d}{dx} = \frac{d}{dy}\frac{dy}{dx} = \frac{d}{dy}\frac{df(x)}{dx} \text{ so}$$

$$\frac{d}{dy} = \frac{1}{f'(x)}\frac{d}{dx}$$

where $f' = dy/dx$.

In vector form and using Chebyshev pseudo-spectral matrix the derivatives of a function $u(y)$ with $y \in A$ may be expressed as:

$$U'_y = D_y \, \mathcal{D}U(x)$$

where $D_y = dx/dy$.
The expression for second derivatives of u incorporating the metrics of the transformation can be computed explicitly:

$$\frac{d^2}{dy^2} = \frac{d}{dy}\left(\frac{d}{dx}\right)\frac{dx}{dy} + \frac{d}{dx}\left(\frac{d^2x}{dy^2}\right) =$$

$$= \frac{d^2}{dx^2}\left(\frac{dx}{dy}\right)^2 + \frac{d}{dx}\left(\frac{d^2x}{dy^2}\right)$$

3.1.3 Stretching

Frequently the situation arises where fine flow structures in boundary layers forming on complex bodies must be adequately resolved. While the natural distribution of the Chebyshev-Gauss-Lobatto points may be used to that end, it is detrimental for the quality of the results expected to apply the same distribution at the far field, where it is not needed, while the sparsity of the Chebyshev points in the center of the domain may result in inadequate resolution of this region. One possible solution is to use stretching, so that the nodes get concentrated around a desired target region. In this case the goal is to transform the initial domain $I = [-1,1]$ into $A = [a,b]$ with the special feature that the middle point, zero, turns into an arbitrary $y_{1/2} \in [a,b]$.

So let $x_i = \cos\frac{i\pi}{N}$, the stretching function,

$$f(x_i) = a + \frac{\frac{(b-a)(y_{1/2}-a)}{b+a-2y_{1/2}}(1+x_i)}{\frac{2(y_{1/2}-a)}{b+a-2y_{1/2}} + (1-x_i)} \tag{53}$$

transform every point in x_i into A, such that, $f(-1) = a$, $f(1) = b$ and $f(0) = y_{1/2}$.
This function could also be written as:

$$f(x_i) = a + \frac{c(b-a)(1+x_i)}{(1-2\cdot c)(1-x_i)}$$

where c is a stretching factor and represents the displacement ratio of the image of $x_i = 0$ in A. This function is not continuous when $y_{1/2} = (b+a)/2$; $c = 0.5$, i.e. when there is no stretching.

For representing a function u in the new set of points y_i in A, the procedure is the same than for any mapping, taking into account that,

$$f^{-1}(y_i) = \frac{(2\cdot c+1)(y_i-a) - (b-a)c}{(b-a)c + (y_i-a)}$$

and its derivatives could be easily calculated.

3.1.4 Two–dimensional expansions

All the one-dimensional results presented up to this point can be extended to the multidimensional approach.
The extension to two dimensional approximations is done using tensor products of one-dimensional expansions. Thus in the unit square $[-1,1]^2$ and for Chebyshev polynomials,

$$T_{ij}(x,y) \equiv T_i(x)T_j(y) \quad i = 1,\ldots,N_x; j = 1,\ldots,N_y \tag{54}$$

The truncated Chebyshev series of degree N_x in the x-direction and N_y in the y-direction is

$$\begin{aligned} u_{\mathcal{N}}(x,y) &= \sum_{i=0}^{N_x}\sum_{j=0}^{N_y} \hat{u}_{ij}T_i(x)T_j(y) \\ &= \sum_{i=0}^{N_x}\sum_{j=0}^{N_y} \hat{u}_{ij}T_{ij}(x,y) \end{aligned} \tag{55}$$

Orthogonality of each one-dimensional Chebyshev basis implies orthogonality of the tensor product two-dimensional basis.

$$(T_{km}, T_{ln})_w = \frac{\pi^2}{4} c_k c_m \delta_{k,l}\delta_{m,n}$$

The approximation of a function $u(x,y)$ with the collocation technique makes use of the Gauss or Gauss-Lobatto mesh:

$$ll\left(x_i, y_j\right) = \left(\cos\frac{i\pi}{N}, \cos\frac{j\pi}{N}\right) \quad \text{Gauss-Lobatto choice.} \tag{56}$$

$$\left(x_i, y_j\right) \in \Omega_N = [-1,1]\times[-1,1] \tag{57}$$

This matrix of nodes is arranged in an array fixing the y-value while x-value changes. This choice is the responsible for the characteristic form of differential Chebyshev pseudo-spectral matrix in each direction.
These matrices can be formed from the 1D differential operator placing every coefficient in its respective row and column or easily if Kronecker tensor product $(\otimes)$ is consider Trefethen (2000) .
The Kronecker product of two matrices A and B of dimension $p \times q$ and $r \times s$ respectively is denoted by $A \otimes B$ of dimension $pr \times qs$. For instance

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \otimes \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \left(\begin{array}{cc|cc} a & b & 2a & 2b \\ c & d & 2c & 2d \\ \hline 3a & 3b & 4a & 4b \\ 3c & 3d & 4c & 4d \end{array}\right) \tag{58}$$

Using this tensor product, the two dimensional derivatives matrices are computed. Let $[a, b] \times [c, d]$ the computational domain discretized using the same number of Gauss-Lobatto points in each direction (N), and let $\mathcal{D}$ be the one dimensional Chebyshev pseudo-spectral matrix for these number of nodes, then, $\mathcal{D}_x = I \otimes \mathcal{D}$ and $\mathcal{D}_y = \mathcal{D} \otimes I$ where I is the $N \times N$ identity matrix.
Second order derivative matrices are built using the same technique $\mathcal{D}_{xx} = I \otimes \mathcal{D}^2$, $\mathcal{D}_{yy} = \mathcal{D}^2 \otimes I$ and $\mathcal{D}_{xy} = (I \otimes \mathcal{D}) \times (\mathcal{D} \otimes I)$, now it is easy to translate any differential operator to its vector form.

3.1.5 Multidomain theory

Domain decomposition methods are based on dividing the computational region in several domains in which the solution is calculated independently but taking into account information from the neighboring domains. From now on boundary conditions coming from the interface between two domains were called interface conditions to distinguish from physical boundary conditions: inflow, outflow, wall,...
The advantages of this technique appear in several situations. The first one is related with the geometry of the problem to be solved. Chebyshev polynomials, without any metric transformation, require rectangular domains, using this multidomain technique it is possible to deal with problems which can be decomposed into rectangular subdomains. A second advantage of this method is the possibility of mapping specific areas of the computational domain with dense grids while in "less interesting" regions coarse grids could be used. This different resolution for different subdomains allows an accurate solution without wasting computational requirements where not needed.

3.1.6 One-dimensional multidomain method

The multidomain method applied to one dimensional problems means solving as many equations as domains. Due to the choice of Chebyshev-Gauss-Lobatto nodes the domains involved share one extremum point, this node will appear twice in the unknown vector $x_N^1 = x_0^2$.

In vector form for two domains the differential equation would be defined as:

$$\left(\begin{array}{c|c} \mathcal{L}_1 & \\ \hline & \mathcal{L}_2 \end{array}\right)\left(\begin{array}{c} U_1 \\ \\ U_2 \end{array}\right) \tag{59}$$

3.1.7 Two-dimensional multidomain method

It is in the extension to two dimensional problems when the features of the multidomain approach could be better exploited. There is no essential difference with the one–dimensional case but the complexity in the implementation of the technique warrants a detailed explanation.

First the domains are enumerated from bottom to top and from right to left. The connection among them is now not a single point but a row of nodes. In the simplest situation these nodes match between domains, i.e. two domains share a row of nodes. But dealing with problems with "more interesting" regions made necessary the possibility of meshing each domain with a different number of nodes. In this case non-conforming grids are built, which make use of the interpolation tool to be discussed shortly. The matrix form is built in a straightforward manner considering each domain independently.

$$\left(\begin{array}{c|c} \mathcal{L}^1 & 0 \\ \hline 0 & \mathcal{L}^2 \end{array}\right)\left(\begin{array}{c} U^1_{0,0} \\ \vdots \\ U^1_{N_x^1,N_y^1} \\ \hline U^2_{0,0} \\ \vdots \\ U^2_{N_x^2,N_y^2} \end{array}\right) \tag{60}$$

3.1.8 Boundary and interface conditions

Both in one– and two–dimensional problems once the differential matrix has been formed boundary conditions need to be imposed. In multidomain methods two kinds of conditions are present, true *boundary* conditions arising from physical considerations on the behaviour of the sought functions at physical domain boundaries, such as inflow, outflow or wall, and *interface* conditions, imposed in order to provide adequate connection between the subdomains.

3.1.9 One-dimensional boundary and interface conditions

Depending on what kind of boundary condition the problem has (Dirichlet, Neumann, Robin), the implementation is different. The nodes affected by this conditions are, in any type of boundary condition, U_0 and U_N. That is why only the first and last row in the matrix operator are changed, for instance, homogeneous Dirichlet boundary condition in U_0 means

replacing the first row with a row full of zeros but in the first position where will be an one. Neumann bc in U_N needs the substitution of the last row in $\mathcal{L}$ for the last row in $\mathcal{D}$.
Interface conditions in one dimensional problems reduce to imposing continuity equations in the shared node. Depending on the order of the problem the interface continuity conditions are imposed for higher order derivatives. In a second order differential equation, such like the BiGlobal EVPs treated presently, the interface conditions consist of imposing continuity in function and first derivative as follows:

$$\begin{aligned} U_N^1 &= U_0^2 \\ U_N^{1\prime} &= U_0^{2\prime} \end{aligned} \tag{61}$$

The effect on vector form is again the substitution of as much as rows in the matrix as number of conditions needed.

3.1.10 Two-dimensional boundary and interface conditions

After building the matrix which discretized the differential operator, the issue of imposing boundary and interface conditions must be addressed. Boundary conditions do not present additional complexity compared with the one–dimensional case apart from the precise positioning of the coefficient in the matrix; however, interface conditions deserve a more detailed discussion.
The equations for the interface conditions in a two dimensional second order differential problem are the same that the ones for one dimensional case except for the number of nodes involved. If the grids in the two domains are conforming (point to point matching) these equation are (supposing connection between domains in x_{max}^1 to x_{min}^2):

$$\begin{aligned} U_{N_x,i}^1 &= U_{0,i}^2 \\ \left.\frac{\partial U^1}{\partial x}\right|_{N_x,i} &= \left.\frac{\partial U^2}{\partial x}\right|_{0,i} \end{aligned} \tag{62}$$

If non-conforming grids are present an interpolation tool is necessary for imposing interface conditions. Hence supposing connection between domains in y_{max}^1 to y_{min}^2:

$$\begin{aligned} U_{j,N_y}^1 &= {}^{2\to1}I_{j,i}\, U_{i,0}^2 \\ {}^{1\to2}I_{j,i}\left.\frac{\partial U^1}{\partial y}\right|_{i,N_y} &= \left.\frac{\partial U^2}{\partial y}\right|_{j,0} \end{aligned} \tag{63}$$

3.2 Galerkin approximation method

Turning to the Galerkin approach, the approximate solution of the problem is sought in a function space consisting of sufficiently smooth functions satisfying the boundary conditions. This method is based on the projection of the approximate solution in a finite dimensional

space of the basis functions, ψ. If $(\hat{u}_i, \hat{p})$ is the approximate solution of the problem then:

$$A \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{p} \end{pmatrix} + \omega B \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{p} \end{pmatrix} = R. \tag{64}$$

where now R is the residual or error that results from taking the approximate numerical solution instead of the exact solution. The residual is projected on a finite basis $\psi_j \quad j = 1, ...N$ with dimension N and the objective of the methodology is to drive R to zero.

$$\int_\Omega R\psi_j d\Omega = \int_\Omega \left[A \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{p} \end{pmatrix} + \omega B \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{p} \end{pmatrix} \right] \psi_j d\Omega = 0. \tag{65}$$

The operator A contains second derivatives in the viscous term and also the pressure gradient term, for those terms integration by parts must be applied taking into account the boundary conditions. The application of boundary condition vanish boundary integrals where Dirichlet boundary conditions are fixed and also boundaries where natural boundary conditionsGonzález & Bermejo (2005) are imposed.
The approximate solution $(\hat{u}_1, \hat{u}_1, \hat{u}_1, p)$ can be expressed as linear expansion over the number of degrees of freedom of the system. Let us call N the number of velocity points or degrees of freedom and NL the number of pressure points, then the final solution can be expressed as:

$$\hat{u}_i = \psi_\alpha \hat{u}_i^\alpha \qquad (\alpha = 1, ..., N) \tag{66}$$

$$\hat{p} = \psi_\lambda \hat{p}^\lambda \qquad (\lambda = 1, ..., NL) \tag{67}$$

After the variational formulation, the operator A is represented by a $(3N + NL)^2$ matrix, becomes

$$A = \begin{pmatrix} F_{ij} + C_{ij}^{11} + i\beta E_{ij} & C_{ij}^{12} & 0 & -\lambda_{ij}^x \\ C_{ij}^{21} & F_{ij} + C_{ij}^{22} + i\beta E_{ij} & 0 & -\lambda_{ij}^y \\ C_{ij}^{31} & C_{j}^{32} & F_{ij} + i\beta E_{ij} & i\beta D_{ij} \\ \lambda_{ji}^x & \lambda_{ji}^y & i\beta D_{ji} & 0 \end{pmatrix}, \tag{68}$$

where $F_{ij} \equiv \gamma_{ij} + \left(R_{ij} + \beta^2 M_{ij} \right)$. The real symmetric operator B is also introduced by

$$B = \begin{pmatrix} M_{ij} & 0 & 0 & 0 \\ 0 & M_{ij} & 0 & 0 \\ 0 & 0 & M_{ij} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, i, j = 1, \cdots, N \tag{69}$$

where M represents the mass matrix; the elements of all matrices introduced in (68) and (69) are presented next.
Defining the quadratic velocity basis functions as ψ and the linear pressure basis functions as ϕ, the following entries of the matrices A and B of the generalized BiGlobal EVP appearing in equation (65) are obtained

$$\gamma_{ij} = \bar{u}_m^l \int_\Omega \psi_l \frac{\partial \psi_i}{\partial x_m} \psi_j d\Omega, \qquad l,i,j = 1,...,N. \tag{70}$$

$$C_{ij}^{mk} = \left(\frac{\partial \bar{u}_m}{\partial x_k}\right)^l \int_\Omega \psi_l \psi_i \psi_j d\Omega, \qquad l,i,j = 1,...,N \quad m = 1,2,3 \quad k = 1,2. \tag{71}$$

$$E_{ij} = \bar{u}_3^l \int_\Omega \psi_l \psi_i \psi_j d\Omega, \qquad l,i,j = 1,...,N. \tag{72}$$

$$R_{ij} = \int_\Omega \frac{\partial \psi_i}{\partial x_m} \frac{\partial \psi_j}{\partial x_m} d\Omega, \qquad i,j = 1,...,N \quad m = 1,2. \tag{73}$$

$$M_{ij} = \int_\Omega \psi_i \psi_j d\Omega, \qquad i,j = 1,...,N. \tag{74}$$

$$D_{ij} = \int_\Omega \phi_i \psi_j d\Omega, \qquad i = 1,...,NL \quad j = 1,...,N \tag{75}$$

$$\lambda_{ij}^x = \int_\Omega \phi_i \frac{\partial \psi_j}{\partial x}, \qquad i = 1,...,NL \quad j = 1,...,N d\Omega, \tag{76}$$

$$\lambda_{ij}^y = \int_\Omega \phi_i \frac{\partial \psi_j}{\partial y}, \qquad i = 1,...,NL \quad j = 1,...,N d\Omega. \tag{77}$$

3.2.1 Low-order Taylor-Hood finite elements

Once a general Galerkin formulation of the EVP has been constructed, a choice of a certain base for the basis functions to construct a final version of the operators A and B must be made. All the terms contained in those operators are defined by an integral over the computational domain Ω. To perform the calculation of those integrals we are going to divide the full computational domain into a finite number of sub-domains or elements. Let us call M the number of elements used for the domain decomposition, this implies that a mesh generation has been performed in such a way that:

$$\bigcup_{i=1}^{M} \Omega_i = \Omega, \tag{78}$$

$$\bigcap_{i=1}^{M} \Omega_i = \emptyset, \tag{79}$$

The two-dimensional computational domain will be divided in either triangular or quadrilateral elements. As the integral calculation has been reduced to the summation of integrals over single elements, then the basic functions must be defined in the different

elements. An alternative way is to map all triangular/quadrilateral elements into a standard triangular/element element known as the reference element, this way the definition of the basis functions is only performed once for this reference element.
Before discussing the benefits of different types of polynomial expansions, we first need to introduce the concepts of modal and nodal expansions. In a nodal expansion the value of the coefficients that are used in the linear expansion of the solution are representing the approximate solution in a certain discretization point, this means that the numerical value has a physical interpretation. On the other hand a modal expansion does not have that physical meaning and the physical point comes from the full linear combination of modes.
The numerical solution of the EVPs described in the previous sections may be accomplished by a nodal expansion of the unknowns on a set of nodes, x_q, using a set of basis functions, $\Phi_q(x)$. Linear- and quadratic Lagrange polynomials are the method of choice for $\Phi(x)$ in low-order FEM Cuvelier et al. (1986), and have also been used in our earlier work Gonzalez L et al. (2007). The associated nodal points, x_q, are chosen such that that $\Phi_p(x_q) = \delta_{pq}$, where δ_{pq} represents the Kronecker delta. This property implies that the discrete approximation, u^δ, of a function may be defined at x_q in terms of the expansion coefficients $\widehat{u}_p$ as

$$u^\delta(x_q) = \sum_{p=0}^{P} \widehat{u}_p \Phi_p(x_q) = \sum_{p=0}^{P} \widehat{u}_p \delta_{pq} = \widehat{u}_q; \tag{80}$$

in other words, the expansion coefficients approximate the function at the set of the nodal points.
One of the typical example of nodal basis functions are the Taylor-Hood elements, this is a P2/P1-quadratic polynomials for velocity and linear polynomials for pressure (or Q2/Q1ÂŮbiquadratic polynomials for velocity and bilinear polynomials for pressure); see Allievi & Bermejo (2000); González & Bermejo (2005) for further details. This configuration ensures stability of the finite-element discretization of all EVPs solved, the $inf - sup$ compatibility condition must be satisfied by the discrete spaces in which disturbance velocity components and disturbance pressure are respectively discretized.
However, experience with the low-order method has shown that the necessity to resolve structures associated with linear perturbations at moderate $Re-$numbers results in the need for rather fine grids, with all the consequent large memory and CPU time requirements Gonzalez L et al. (2007). It is then natural to seek an alternative high order discretization, based on a modal expansion.

3.2.2 High order spectral/hp elements

The first characteristic of such an expansion is that there is no physical interpretation of the associated expansion coefficients. Second, modal expansion is hierarchical, meaning that the expansion set of order $P-1$ is contained within the expansion set of order P; a modal expansion based on the Legendre polynomials $L_p(x)$ will be used in what follows. The key property of this expansion set is its orthogonality which, in addition to the hierarchical construction, leads to well-conditioned matrices Karniadakis & Sherwin (2005). Note that, for problems involving up to second-order differentiation, as those encountered herein, it is sufficient to guarantee that the approximate solution is in H^1. Typically, in the finite element methods this is solved imposing a C^0 continuity between elemental regions, that is the global

expansion modes are continuous everywhere in the solution domain while continuity in the derivatives is achieved at convergence Karniadakis & Sherwin (2005). Boundary and interior nodes are distinguished in this expansion: the former are equal to unity at one of the elemental boundaries and are zero at all other boundaries; the latter class of modes are non-zero only at the interior of each element and are zero along all boundaries. In the standard interval $\Omega = \{\xi | -1 \leq \xi \leq 1\}$ the $p-$type modal expansion is denoted by $\psi_p(x)$ and is defined as

$$\psi_p(\xi) = \begin{cases} \dfrac{1-\xi}{2} & p = 0 \\ \left[\dfrac{1-\xi}{2}\right]\left[\dfrac{1+\xi}{2}\right] L_{p-1}(\xi) & 0 < p < P \\ \dfrac{1+\xi}{2} & p = P \end{cases}$$

It may be seen that the lowest expansion modes, $\psi_0(x)$ and $\psi_P(x)$, are the same as the low-order finite element expansion, these boundary modes being the only modes that are nonzero at the ends of the interval. The remaining interior modes are zero at the ends of the interval and increase in polynomial order. As a consequence of the orthogonality of the Legendre polynomials $L_{p-1}(\xi)$ the stiffness and mass matrices in one-dimensional problems are tri- and penta-diagonal, respectively. The resulting discretization is denoted as spectral/hp element method.

In an one-dimensional spectral/hp decomposition the global expansion basis is decomposed into elemental subdomains that can then be mapped into the standard interval $[-1, 1]$. The polynomial basis is then defined in the standard region. To complete a Galerkin formulation it will be necessary to choose some form of numerical integration; for the purpose of the present work Gaussian quadrature is selected. The one-dimensional concept may be extended to multiple dimensions in a straightforward manner. In two dimensions two standard regions, a quadrilateral or a triangle, may be used. All bases used in what follows can be expressed in terms of modified principal functions. In the quadrilateral expansion:

$$\psi_{pq}(\xi_1, \xi_2) = \psi_p(\xi_1)\psi_q(\xi_2). \tag{81}$$

In the triangular expansion:

$$\psi_{pq}(\xi_1, \xi_2) = \psi_p(\eta_1)\psi_{pq}(\eta_2). \tag{82}$$

In the standard quadrilateral region, the Cartesian coordinates (ξ_1, ξ_2) are bounded by constant limits

$$\mathcal{Q}^2 = \{(\xi_1, \xi_2) | -1 \leq \xi_1, \xi_2 \leq 1\}. \tag{83}$$

This is not the case in the standard triangular region, where the bounds of the Cartesian coordinates (ξ_1, ξ_2) depend on each other, that is,

$$\mathcal{T}^2 = \{(\xi_1, \xi_2) | -1 \leq \xi_1, \xi_2, \xi_1 + \xi_2 \leq 0\}. \tag{84}$$

A means of developing a suitable tensorial-type basis within unstructured regions, such as the triangle, is suggested by Karniadakis and Sherwin Karniadakis & Sherwin (2005),

in terms of a coordinate system in which the local coordinates have independent bounds. The advantage of such a system is that one-dimensional functions may be defined, upon which a multi-domain tensorial basis may be constructed. A suitable coordinate system, which describes the triangular region between constant independent limits, is defined by the transformation Karniadakis & Sherwin (2005)

$$\eta_1 = 2\frac{1+\xi_1}{1-\xi_2} - 1, \tag{85}$$

$$\eta_2 = \xi_2. \tag{86}$$

The final step in constructing a multi-dimensional spectral/hp approximation is a mapping of every subdomain (element) into the corresponding standard region $[-1,1]$ for the 1-D case, $\mathcal{Q}^2$ for the 2-D quadrilateral elements or $\mathcal{T}^2$ for the 2-D triangular elements.
In order to ensure stability of the finite-element discretization of all EVPs solved, the $inf-sup$ compatibility condition must be satisfied by the discrete spaces in which disturbance velocity components and disturbance pressure are respectively discretized. In the present spectral/hp solution the number of disturbance pressure modes has been kept one less than that of the disturbance velocities.

4. EVP solution – Krylov subspace iteration

Having dealt with the spatial discretization of the EVP, we now turn our attention to obtaining the eigenvalues of the LNSE direct or adjoint matrix. From a linear stability analysis point of view, the most important eigenvalues are those closest to the axis $\omega_r = 0$ and here an iterative method has been used for their determination. Specifically, the well-established in BiGlobal linear instability problemsTheofilis (2003) Arnoldi algorithm Nayar & Ortega (1993) has been used.
The Arnoldi method is a subspace iteration method, the computation time of which depends linearly on the subspace dimension. As experienced in earlier analogous studiesDing & Kawahara (1998) only eigenvalues with large modules can be obtained by straightforward application of the algorithm. Since the eigenvalues closest to the imaginary axis are sought, a simple transformation is used in order to convert the original problem into one where the desired values have large modules. Note that the eigenvectors are not affected by this transformation. Specifically, defining

$$\mu = -\omega^{-1}, \tag{87}$$

it follows that

$$A^{-1}B\hat{\mathbf{q}} = \mu\hat{\mathbf{q}}, \quad A^{-1}B = C, \quad C\hat{\mathbf{q}} = \mu\hat{\mathbf{q}}. \tag{88}$$

This transformation converts the original generalized into the standard EVP. A finite but small (compared with the leading dimension of A, B) number of eigenvalues (equal to the Krylov subspace dimension) m is sought, which is obtained by application of the Arnoldi algorithm as follows

1. CHOOSE an initial random vector v_1 and NORMALIZE it.
2. FOR j=1,2,....m DO:
   (a) Calculate w_j as $Cv_j = w_j$. Which is equivalent to solve the problem $Aw_j = Bv_j$(A non-symmetric).
   (b) FOR i=1,2,...j DO:

$$h_{ij} = (Cv_j, v_i), \tag{89}$$

$$a = \sum_{i=1}^{j} h_{ij} v_i, \tag{90}$$

$$\hat{v}_{j+1} = w_j - a, \tag{91}$$

$$h_{j+1,j} = \|\hat{v}_{j+1}\|, \tag{92}$$

$$v_{j+1} = \frac{\hat{v}_{j+1}}{h_{j+1,j}} \tag{93}$$

   END DO
END DO

This algorithm delivers an orthonormal basis $\mathbf{V}_m = [v_1, v_2, ..., v_m]$ of the Krylov subspace $\mathbf{K}_m = span\{v_1, Cv_1, ..., C^{m-1}v_1\}$. The restriction from $\mathbf{C}$ to $\mathbf{K}_m$ is represented by the matrix $\mathbf{H}_m = \{h_{ij}\}$. The eigenvalues of the latter matrix are an approximation of the m largest eigenvalues of the original problem (65). The eigenvectors associated with these eigenvalues may be obtained from

$$\hat{\mathbf{q}}_i = \mathbf{V}_m \tilde{y}_i \tag{94}$$

where $\tilde{y}_i$ is an eigenvector of $\mathbf{H}_m$ associated with the μ_i-th eigenvalue.
Note that, since the matrix C is unknown a-priori, a non-symmetric linear system $Cv_j = A^{-1}Bv_j = q_j$ or, equivalently, $Aq_j = Bv_j$ must be solved at each iteration, q_j being an unknown auxiliary vector. It is important to remark that the invert process needs the inversion of the A operator which means that at least one LU or Incomplete-LU decomposition must be done.
The total time needed for a complete Arnoldi analysis depends mostly on the efficiency of the linear solver described above, as well as on the Krylov space dimension m used to approximate the most important eigenvalues.
In order to find the leading eigenvalue with maximum real part, one can use the shift and invert strategy. If ω_0 is an approximation to the complex eigenvalue of interest, the shifted and inverted problem is

$$(C - \omega_0)^{-1} \hat{\mathbf{q}} = \mu \hat{\mathbf{q}} \tag{95}$$

where $\mu = -(\omega - \omega_0)^{-1}$.

5. Results

5.1 Single–domain spectral collocation computations

Selected results, representative of the applications to which the tools discussed earlier have been applied, are briefly exposed next; full discussion of the respective applications may

be found in the references provided. Starting with the single–domain spectral collocation method on a rectangular domain, which has been the main approach applied to the solution of the BiGlobal EVP by several authors Lin & Malik (1996a;b); Tatsumi & Yoshimura (1990); Theofilis et al. (2003), attention is focused on the global instability of laminar separation bubbles (LSBs) in the incompressible regime Theofilis et al. (2000), and on attachment–line instability in compressible flow Duc et al. (2006); Theofilis, Fedorov & Collis (2004).

In the context of LSB flows, two approaches have been followed for the construction of the basic flows, firstly an analytical one, making use of the attached and the separated Falkner–Skan branches, and secondly an inverse boundary–layer method, in which the wall–shear is prescribed. In the context of the first methodology, the instability analysis was performed for bubbles with different Reynolds numbers, but the same value of $m_{\rm lim} = -0.025$. All the bubbles analyzed have a peak reversed-flow higher than 10% of the reference velocity, chosen in order to ensure proximity to conditions of amplification of the global mode Theofilis et al. (2000). The amplitude functions of the global modes discovered in our previous work were centered on the bubble and extended for only a short distance downstream. Analogous boundary conditions were imposed here but substantially higher resolutions were required in order to converge results at (displacement–thickness–based) Reynolds number $Re = 50$. Concretely, the system (30-33) been solved for the determination of the (direct) linear global perturbations and convergence of the results has been attained using $(N_x, N_y) = (120, 40)$ spectral collocation points and a constant Krylov subspace dimension of 200. At these parameters the discretized eigenvalue problem for each pair of the parameters (Re,β) required the storage and inversion of a matrix with leading dimension 1.98×10^4, which translates on 6 Gbytes of memory. Though hardware was available in order to perform this work serially in a shared–memory manner, all computations shown were performed in parallel.

The spanwise wavenumber parameter range $0 \approx \beta \leq 35$ has been examined and two global modes have been identified. The first one is known from previous work Theofilis et al. (2000), is steady ($\omega_r = 0$) and its periodicity length in the spanwise direction is approximately the same of the bubble; at relatively high β values this mode is transformed into two travelling modes ($\omega_r \neq 0$). The newly discovered mode is always a traveling instability with spanwise periodicity length about a third of that of the bubble Rodríguez (2008). Both modes have been encountered for Reynolds numbers, based on the displacement-thickness at inflow, below 100. On the other hand, the solution of the adjoint BiGlobal EVP provides the spatial regions where the receptivity of the basic flow to small perturbations is higher: the adjoint eigensolution field defines the efficiency by which a particular forcing excites the direct eigensolution Hill (1992). The adjoint BiGlobal EVP (39-42) has been solved next; the boundary conditions applied to the direct problem are inverted here, as discussed in section 2. Dirichlet homogeneous boundary conditions are imposed at the right-hand boundary while linear extrapolation from the interior of the domain is imposed at the left. The amplitude functions recovered are again centered in the bubble, but this time they are extend upstream. Some perturbation velocity components seem to vanish at the left boundary of the domain.

Performing the same analysis for $Re_{\delta^*} = 500$ at inflow, the eigenspectrum is symmetric, and some branches can be identified corresponding to the global modes. The least stable eigenmode founded here is the same steady mode present at lower Reynolds numbers, and is depicted on figure 2 for the direct and adjoint eigenmodes . Much like the lower–Reynolds number case, the spatial structure of the global modes is centered on the bubble, and extends downstream for the direct eigenfunction and upstream for the adjoint Rodríguez & Theofilis

(2008). Finally, relaxing the homogeneous Dirichlet boundary conditions imposed at the inflow boundary, TS–like disturbances may be obtained.Rodríguez & Theofilis (2010)
Single–domain solutions of the compressible BiGlobal EVP (8) are also slowly finding themselves in the literature Robinet (2007); Theofilis & Colonius (2004); Theofilis, Fedorov & Collis (2004). A particularly interesting example is the compressible analogue of the classic swept Hiemenz flow, the global instability of which has been addressed by Lin and Malik Lin & Malik (1996a;b) and further analyzed by Theofilis *et al.* Theofilis et al. (2003). From a numerical point of view, an interesting aspect of this flow is the polynomial structure of its linear global eigenmodes, which is exact in the incompressible limit Theofilis et al. (2003). Recovery of the polynomial structure in the global eigenspectrum is a rather stringent test for the accuracy of any BiGlobal EVP solver. Linear extrapolation has been imposed as a boundary condition on the eigenvector–solution of (8) at $x \rightarrow \infty$, along the chordwise direction normal to the attachment–line, *for all* eigenmodes. Although such a boundary condition is exact only for the polynomials depending linearly on the chordwise coordinate x, this boundary condition has been found to perform well for all flow eigenmodes. It was found Theofilis, Fedorov & Collis (2004) that the polynomial structure of the global flow eigenmodes persists in the compressible flow regime in an approximate sense, the error of the approximation being of $O(Ma^2/Re^2)$. The Göertler-Häemmerlin structure of the leading eigenmode, Theofilis, Fedorov & Collis (2004), has been confirmed independently by direct numerical simulations of Le Duc *et al.* Duc et al. (2006). The interested reader is referred to Duc et al. (2006); Lin & Malik (1996a;b); Theofilis, Fedorov & Collis (2004); Theofilis et al. (2003) for more details of global instability analysis of this technologically important flow.
Remaining within the realm of single–domain compressible global linear flow instability, attention is turned to the numerical solution of the inviscid equation (11). The first configuration to be addressed in this context has been the elliptic cone at an angle of attack Theofilis (2001); here, validation results in the Poisson limit of (11) and instability analysis results of an aspect ratio 4 elliptic cone, schematically depicted in figure 3, are briefly discussed. A cut plane parallel to the cone base at a distance $z_0 = 0.5$ from the cone base results in an ellipse with geometric parameters

$$a = 0.5,\ b = 0.125,\ \xi_w \approx 0.255, \tag{96}$$

while $\xi_\infty = 1.5$ has been taken to truncate the integration domain. Several validation and verification problems have been solved in the Poisson and Helmholtz limits of (11), in order to demonstrate exponential convergence of the discretization algorithms on the elliptic confocal grid Theofilis (2001). The standard Chebyshev spectral collocation domain discussed in section 3.1 has been used in the wall–normal direction, while Fourier collocation has been used as the natural choice to discretize the azimuthal spatial direction; both collocation discretizations have been modified in order to introduce the elliptic confocal coordinate system. The convergence history of the results for the wall value $f(x = a, y = 0) = f(\xi = \xi_w, \eta = 0)$, obtained from numerical solution of the Neumann problems defined in Theofilis (2001; 2009a), is presented in table 1 to within six significant digits, while the convergence history of the numerical solution of problems S and A is presented in figure 4; the spatial distribution of $f(x, y)$ for both Dirichlet and Neumann boundary data can be found in the same figure (b).

$N\xi$	$N\eta$	$f_{\mathrm{H}}(x=a,y=0)$	$f_{\mathrm{S}}(x=a,y=0)$	$f_{\mathrm{A}}(x=a,y=0)$
10	10	-0.901970	-1.360926	-0.863940
20	20	-0.901970	-1.367825	-0.867146
40	40	-0.901970	-1.375666	-0.868566
80	80	-0.901970	-1.375823	-0.868586

Table 1. Convergence history of numerical solution of the Poisson problems H, S and A, subject to homogeneous Neumann boundary conditions. $N\xi$ and $N\eta$ respectively denote the number of collocation points along the ξ and η coordinate directions.

A key conclusion based on these results is that in all cases studied exponential convergence of the solutions sought has been obtained. The symmetries of the solutions are recovered, rather than being imposed, by the numerical approach taken. Furthermore, when the function sought has a simple, essentially one-dimensional (in ξ) structure, as the case is in problem H, a small number of collocation points suffices to solve the problem monitored and further increases in resolution are unnecessary. This result is attributed to the convergence properties of the Fourier expansion in η and underlines the significance of the choice of the natural elliptic confocal coordinate system for the problem at hand. The low resolution requirements in the η direction in this class of problems, effectively suggesting that the available computing power can be almost exclusively devoted to resolution of the ξ spatial direction, is significant in terms of the ability of the two-dimensional eigenvalue problem to recover results of classic one-dimensional linear theory; in this case the instability mode can be described with a small number of Fourier modes accounting for resolution of the geometry in the lateral (η) direction, while its principal variation is along the wall-normal ξ direction. As the structure of the sought function becomes increasingly more complicated higher resolution is necessary compared with that needed for problem H, although in both problems S and A also exponential convergence has been obtained. However, to the degree that conclusions on the spatial structure of the sought BiGlobal instability modes on the elliptic cone can be drawn upon evidence provided by the numerical solutions of problems A and S, a resolution issue for the inviscid compressible eigenvalue problem may arise if accuracy of the eigenvalue problem results beyond three to four significant digits is necessary, for instance in the case of near-neutral modes.

Turning to hydrodynamic and aeroacoustic instability analyses, the basic state around the elliptic cone has been computed using a standard, second–order accurate aerodynamic solver on a rather fine mesh. The (dominant) basic flow component along the axis generator has then been extracted and equation (11) has been solved at several subsonic and supersonic Mach numbers. The amplitude functions of the disturbance pressure pertaining to the two leading eigenmodes at $Ma = 0.5$ and $Ma = 4$ are shown also in figure 4. Of interest here is the capability of the global instability analysis to recover perturbations of hydrodynamic nature, such as the subsonic modes shown in figure 4, as well as aeroacoustic disturbances, such as the global eigenmodes of supersonic flow also shown in this figure. A full discussion of the physics of these instabilities may be found in Theofilis (2009a).

5.2 Multi–domain spectral collocation computations

In many situations of practical interest, single–domain computations are inappropriate for the analysis of flow instability. One possible solution in such cases is to employ the tools discussed in section 3.1.

Attention is focused on the solution of the incompressible BiGlobal EVP in a simple geometry, namely stable flow in the lid-driven cavity at $Re = 200, \beta = 1$, the instability of which is now well-documented Theofilis (AIAA-2000-1965); Theofilis, Duck & Owen (2004). The minimum nontrivial discretization involves partition of the square into two sub-domains, each of which is discretized independently. In addition, single-domain results were obtained with the reference codede Vicente et al. (2011); Theofilis (AIAA-2000-1965).

Based on the spanwise disturbance amplitude function $\hat{w}(x, y)$ results of the leading eigenmode at $Re = 200, \beta = 1$, shown in the first row of figure 5, and the associated eigenvalue (not presented here), several observations are worthy of mentioning. First, it can be seen that all combinations of domain decomposition examined, at convergence result in the same (stable) mode. Interestingly, the multidomain approach is more efficient than the single-domain computation on both counts of memory requirements and CPU execution time de Vicente et al. (2006); the highest resolved single-domain computation comprises less points than any of the moderately-resolved multidomain computations, and still requires substantially more computing time in order to deliver results of comparable accuracy.

Once the multi–domain BiGlobal solver has been developed, questions regarding flow instability in relatively complex configurations may be addressed. Such a situation arises in flow over an open cavity which contains store models and the influence of the latter on instability characteristics of an empty cavity. The (non–conforming) spectral multi–domain algorithm developed permits defining the simplest model store configuration possible, namely a rectangular object placed inside an open cavity. The stability of this flow has been considered elsewhere de Vicente et al. (2006); here, for brevity, the geometry, mesh and leading eigenmodes of the aforementioned model problem at parameters $Re = 400, \beta = 2\pi$ are shown in figure 6.

5.3 Finite–element global instability analyses

5.3.1 Low–order FEM

Although a great deal of geometric complexity can be dealt with by combinations of the spectral multi–domain methodology and appropriate choices of mapping functions, there exist situations in which this approach becomes cumbersome, on account of its utilizing regular meshes. An unstructured–mesh finite element or spectral element approach then becomes of interest. Such methods have been developed in our group and in what follows the path to building an unstructured–mesh BiGlobal EVP solver is highlighted via a selection of model applications.

Low–order FEM instability analyses, employing the tools of section 3.2.1, have been performed in several applications, discussed in González et al. (2007); they include closed and open systems and rather complex geometries, such as the triangular lid–driven cavity. Here, an additional paradigm is shown, namely instability in a system of two counter-rotating vortices, in which azimuthal spatial periodicity has not been invoked. The basic flow in this case has been obtained by integrating in time the vorticity-streamfunction equations of motion until certain predetermined criteria have been met. The vortex aspect ratio settles to a linear growth after a short initial transient. A BiGlobal instability analysis of the dipole has been

performed, details of which may be found elsewhere González, Gómez-Blanco & Theofilis (2008); the full flowfield composed of a linear superposition of the leading eigenmode upon the basic state at $Re = 3180$ is shown in figure 7. Such structures, qualitatively known from DNS and experiment, may well be confused with nonlinear phenomena; the present analysis is the first of its kind to demonstrate that their origin is a linear (albeit BiGlobal) modal mechanism.

5.3.2 High–order spectral/hp elements Karniadakis & Sherwin (2005)

The main experience obtained from low–order FEM BiGlobal instability analyses is that rather fine meshes are necessary for converged results to be obtained at Reynolds numbers $Re \geq 10^3$. In turn, these translate in large memory and computing time requirements, making the algorithms exposed in section 3.2 the obvious candidate in order to increase efficiency of the computations. Spectral/$hp-$element BiGlobal EVP analyses have been performed initially for several model problems and subsequently for a range of more realistic applications.

First, an isolated Batchelor vortex, the linear instability of which (studied by classic theory Mayer & Powell (1992)) is well understood. On the other hand, low–order FEM analyses of this flow González, Gómez-Blanco & Theofilis (2008) have demonstrated that very fine meshes are necessary for reliable results to be obtained, even for very low $Re = O(10^2)$. In the present spectral/hp context, a large domain compared with the vortex radius was used to make possible the use of Dirichlet boundary conditions in the outer boundary. The mesh has been concentrated in the neighborhood of the basic flow vortex, while very few elements have been found to be necessary in the outer part of the domain. Provided a large Krylov subspace dimension was used, very good agreement with the results presented by Broadhurst *et al.* Broadhurst et al. (2006) could be obtained at $Re = 667$ and $\beta = 2.0$; in figure 8 the axial component of the disturbance velocity of one of the unstable modes (four lobes) can be seen.

Several confined flows have been analyzed next, starting with the well–studied Hagen Poiseuille flow (HPF) in a circular pipe. While from a physical point of view this is probably the most prominent example of failure of modal linear theory to predict transition, the corresponding one-dimensional eigenvalue problem (of the Orr-Sommerfeld class) has been studied exhaustively over the years Lessen et al. (1968); Salwen et al. (1980), thus serving for the present validation work. The analytically-known basic flow, $\bar{u} = 1 - r^2$, has been recovered on the same unstructured mesh as that on which the BiGlobal instability analysis has been performed. The eigenvalues presented by Lessen Lessen et al. (1968) and Salwen Salwen et al. (1980) at four Reynolds numbers have been calculated and excellent agreement with the literature has been obtained; some results are shown in table 2. The corresponding amplitude functions at $Re = 100$ are shown in figure 9; The (rather coarse, but adequate for convergence) mesh utilized is superimposed in these figures, further underlining the power of the spectral/hp method.

While the HPF is a flow with a single basic flow velocity component, along the homogeneous flow direction, z, it is interesting to test the spectral/$hp-$element BiGlobal EVP solver on a problem, which possesses basic flow velocity components only on the plane normal to the $z-$direction and the instability results of which are well–known. The square lid–driven cavity (LDC) is such a problem, having a two-component basic velocity vector, $(\bar{u}(x,y), \bar{v}(x,y), 0)^{\mathrm{T}}$. In order to eliminate potential sources of sub–optimal convergence, the lid velocity is regularized according to Leriche *et al.* Leriche et al. (1998)

Re	frequency	damping rate	frequency	damping rate	
100	0.57256	-0.14714	0.55198	-0.37446	Salwen Salwen et al. (1980)
100	0.57256	-0.14714	0.55198	-0.37446	Lessen Lessen et al. (1968)
100	0.57256	-0.14714	0.55198	-0.37446	present(h=1,p=18)
200	0.64427	-0.12921	0.51116	-0.20266	Salwen Salwen et al. (1980)
200	0.64426	-0.12920	0.51117	-0.20265	Lessen Lessen et al. (1968)
200	0.64526	-0.12920	0.51117	-0.20265	present(h=1,p=20)
300	0.71295	-0.12900	0.56173	-0.16498	Salwen Salwen et al. (1980)
300	0.71295	-0.12907	0.56171	-0.16497	Lessen Lessen et al. (1968)
300	0.71295	-0.12901	0.56172	-0.16497	present(h=1,p=22)
1000	0.84675	-0.07086	0.46916	-0.09117	Salwen Salwen et al. (1980)
1000	0.84675	-0.07086	0.46924	-0.09090	Lessen Lessen et al. (1968)
1000	0.84682	-0.07090	0.46803	-0.09033	present(h=1,p=22)

Table 2. Eigenmodes of Hagen-Poiseuille flow (HPF) at $\beta = 1$ for different Reynolds numbers, obtained on an single element mesh $h = 1$ and p polynomial degree.

$$\bar{u} = (1 - (2x)^{18})^2 \qquad \text{at} \quad y = 1, \tag{97}$$

while $\bar{u} = 0$ on the other three and $\bar{v} = 0$ on all four cavity boundaries.

It is worth noting that regularization of the boundary conditions in the cavity problem is essential in order to obtain a well-posed problem and avoid having to enter into the somewhat artificial debate found in the literature on the critical conditions for instability (c.f. Poliashenko and AidunPoliashenko & Aidun (1995) and related subsequent work) of a singular basic flow. In order to ensure spatial convergence, a collocation Chebyshev spectral code has been used for the baseflow calculations, see details in Theofilis (2003). The mesh considered for this particular case has $[64x64]$ Gauss-Lobato collocation points. The instability analysis mesh comprises five elements $h = 5$ distributed as four trapezoidal elements, around a central square element. Very good agreement has been obtained with the results of the single–domain code Theofilis (AIAA-2000-1965), while the spatial distribution of the amplitude functions of the leading eigenmode in the regularized LDC flow at $(Re, \beta) = (1000, 17)$ is shown in figure 5.

Having demonstrated the high–order spectral/$hp-$element tools on classic academic problems, instability analyses of incompressible flow through two intakes, both having realistic cross-sectional profiles have been performed next. Without reference to concrete applications, representative duct geometries, inspired from motor racing and fighter jets has been selected González, Rodríguez & Theofilis (2008). In order for the analysis to proceed two key assumptions are made. First, flow is taken to be homogeneous along the streamwise spatial direction; second, flow is taken to be laminar and incompressible, driven by a constant pressure gradient along the streamwise spatial direction. First BiGlobal instability analyses have been performed for intake in_1. The Reynolds number in both cases has then been increased up to $Re = 1000$. Figures 10(upper) show the amplitude functions of the disturbance velocity components pertaining to the leading (least damped) eigenmode at $Re = 1000$ and $\alpha = 1$. While the first intake in_1 analyzed had a combination of straight

lines and rounded corners, the second intake in_2 has a completely curved perimeter. In order to perform an accurate calculation the analytical equation that describes the perimeter must be included as input information for the code. The least–stable mode of in_1 can be found in figure 10(lower). The symmetries in these results, existing on account of those of the corresponding basic state, are clearly visible. In line with the Hagen-Poiseuille flow analyzed earlier, these symmetries have not been imposed in the solution of the BiGlobal eigenvalue problem. However, exploitation of the symmetries is one obvious means of reducing the rather demanding memory requirements for the spectral/hp element method solution of the BiGlobal EVP as the Reynolds number increases.

Two more application are highlighted, namely instability of flow around the T106-300 low pressure turbine (LPT) blade Abdessemed et al. (2004; 2006) and that around an aspect ratio 8 ellipse at an angle of attack of 18^o Kitsios et al. (2008), both in incompressible flow. A main difference of the two analyses is that in the former, unlike the tools discussed up to now, in which the matrix has been formed and stored, a time–stepping approach based on the Nektar code Karniadakis & Sherwin (2005) has been used in order to analyze (primary) modal instability of steady basic flow at low Reynolds numbers, $Re \leq 900$, analyze by Floquet theory with respect to three–dimensional instability the flowfield ensuing amplification of the leading eigenmode at $1000 Re 5000$ and perform transient–growth analysis of the flow. On the other hand, instability around the ellipse has been analyzed by constructing an analytic mapping, based on the Joukowski transformation which, in the particular case of the ellipse reduces to the elliptic confocal mapping discussed in section 2.3. In this case the matrix is formed and stored; the analytical mapping makes it rather dense, while at convergence its dimension approaches 1 Tb Kitsios et al. (2008)), such that massive parallelization is the only viable option in order to recover the matrix eigensystem.

Results of the LPT application are discussed in detail by Abdessemed *et al.* Abdessemed et al. (2004; 2006; 2009). Figure 11 (upper-left) shows the leading traveling eigenmode of this flow at chord-based $Re = 890$ Abdessemed et al. (2004), while in the same figure (upper-right) that of steady, $Re = 200$ flow around the ellipse is shown. The qualitative analogies are striking, although it must be borne in mind that the former computations were serially performed on PCs, while for the latter computations $O(1000)$ processors on the Mare Nostrum supercomputer were required; both computations take about one day's computing to converge. Although the time–stepping approach is clearly orders–of–magnitude more efficient that the matrix storage, the latter approach offers access to two– to three orders–of–magnitude more members of the eigenspectrum than the time–stepper does, alongside added flexibility, especially with respect to interchanging various forms of the (compressible or incompressible, viscous or inviscid) LNSE operators; the latter task would imply major changes in the algorithms underlying the time–stepping approach. In the opinion of the authors, when designing a new BiGlobal instability analysis from scratch, provided one has access to supercomputing facilities either the time–stepping or the matrix formation/storage are valid alternatives to be followed. However, when time–periodic states are to be analyzed, the time–stepping approach is the only candidate to be considered. Figure 11 also shows in the lower case the structure of the leading Floquet eigenmode of the time–periodic flow which ensues in the wake of the LPT, at $Re = 2000$, when the eigenmode shown in the upper case has been linearly amplified. The nature of the secondary instability Floquet (BiGlobal) eigenvalue problem, in which a large number of snapshots in time of the time–periodic basic state are treated in a coupled

manner, make storage of the resulting matrix impossible on both present and near–future computing hardware, such that the time–stepping approach first introduced by Barkley and Henderson Barkley & Henderson (1996) and employed to the LPT problem by Abdessemed *et al.* Abdessemed et al. (2006; 2009) are the only viable alternatives in order to pursue this type of analysis.

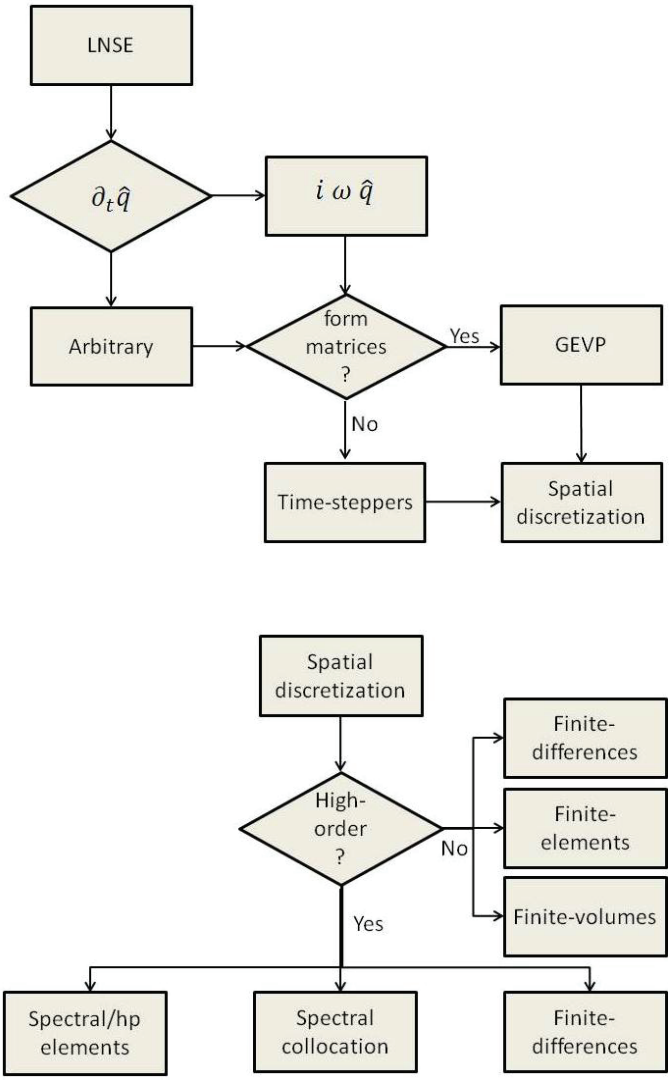


Fig. 1. *Upper:* Treatment of the temporal derivative in a global instability analysis context. *Lower:* Numerical methods which have been used for the spatial discretization of global instability eigenvalue problems.

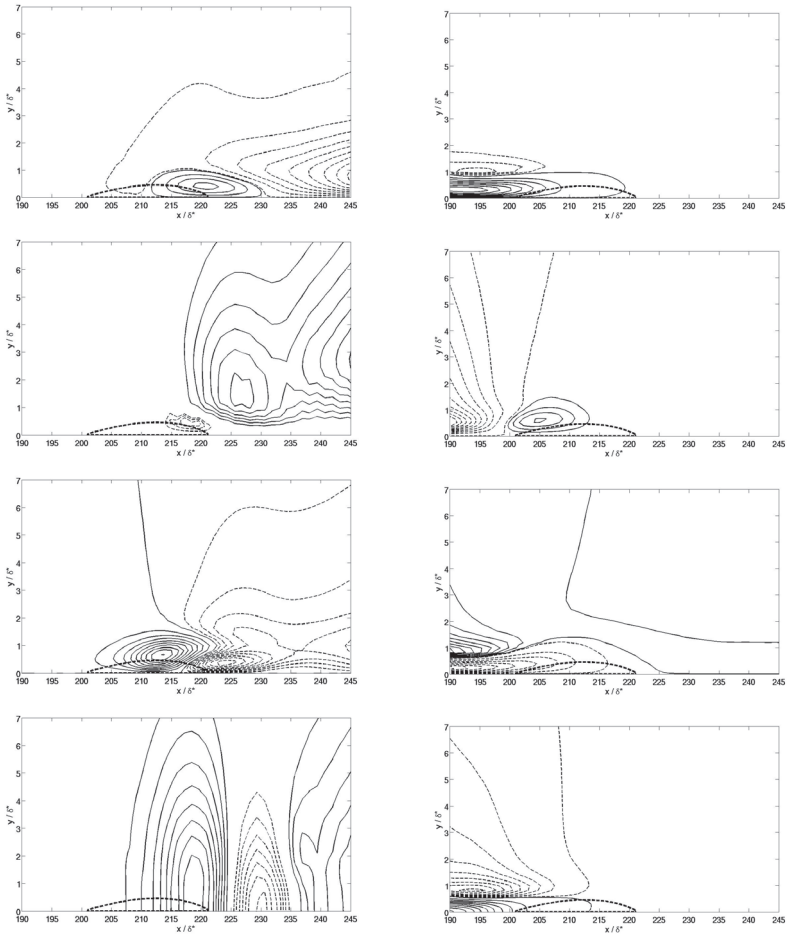


Fig. 2. Most unstable steady global mode for the direct and adjoint problem at $Re_{\delta^*} = 500$ and $\beta = 0.2$. The real streamwise (*first row*), wall-normal (*second row*) and spanwise (*third row*) velocities and pressure (*forth row*) components of the direct (*left column*) and of the adjoint (*right column*) eigenfunctions are shown. The thick dashed line separates the reversed flow region. Only a part of the domain is shown.

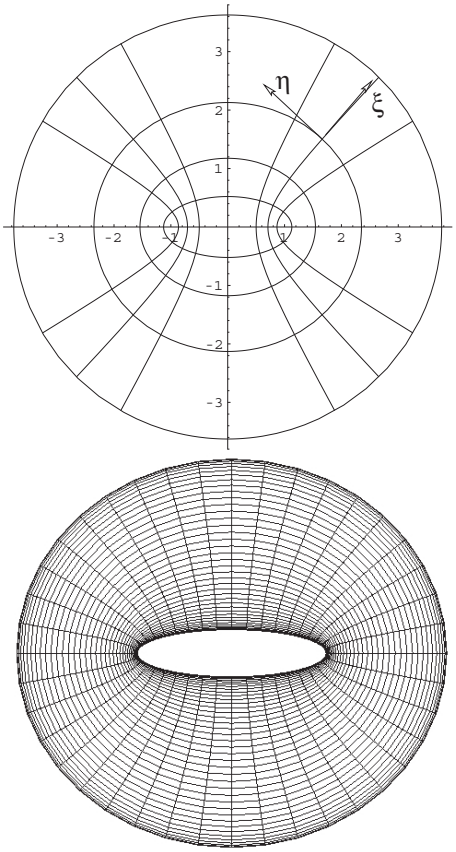
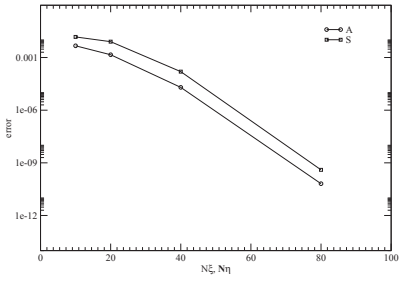
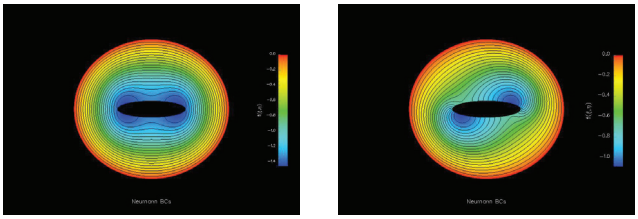


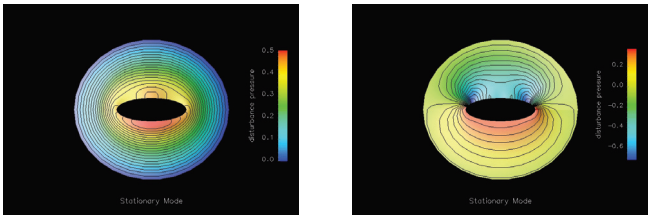
Fig. 3. The elliptic confocal coordinate systemMorse & Feshbach (1953) $O\xi\eta$ (left) and one of the actual grids on which validation work has been performed (right).



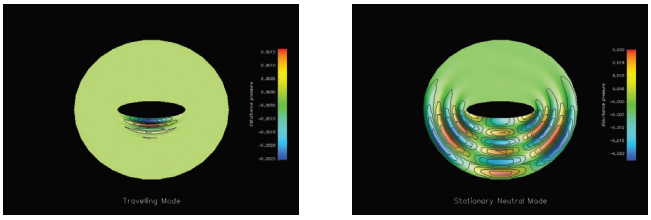
(a)



(b)



(c)



(d)

Fig. 4. (a): Convergence history of the Helmholtz equation on the elliptic confocal grid. (b): Solutions of the Symmetric (left) and Antisymmetric (right) validation problems. (c):Leading eigenmodes of compressible flow over an elliptic cone at $Ma = 0.5$ and (d):$Ma = 4.0$. In (c) and (d) are shown the amplitude functions of the disturbance pressure – solution of (11); first mode in the left– and second mode in the right column.

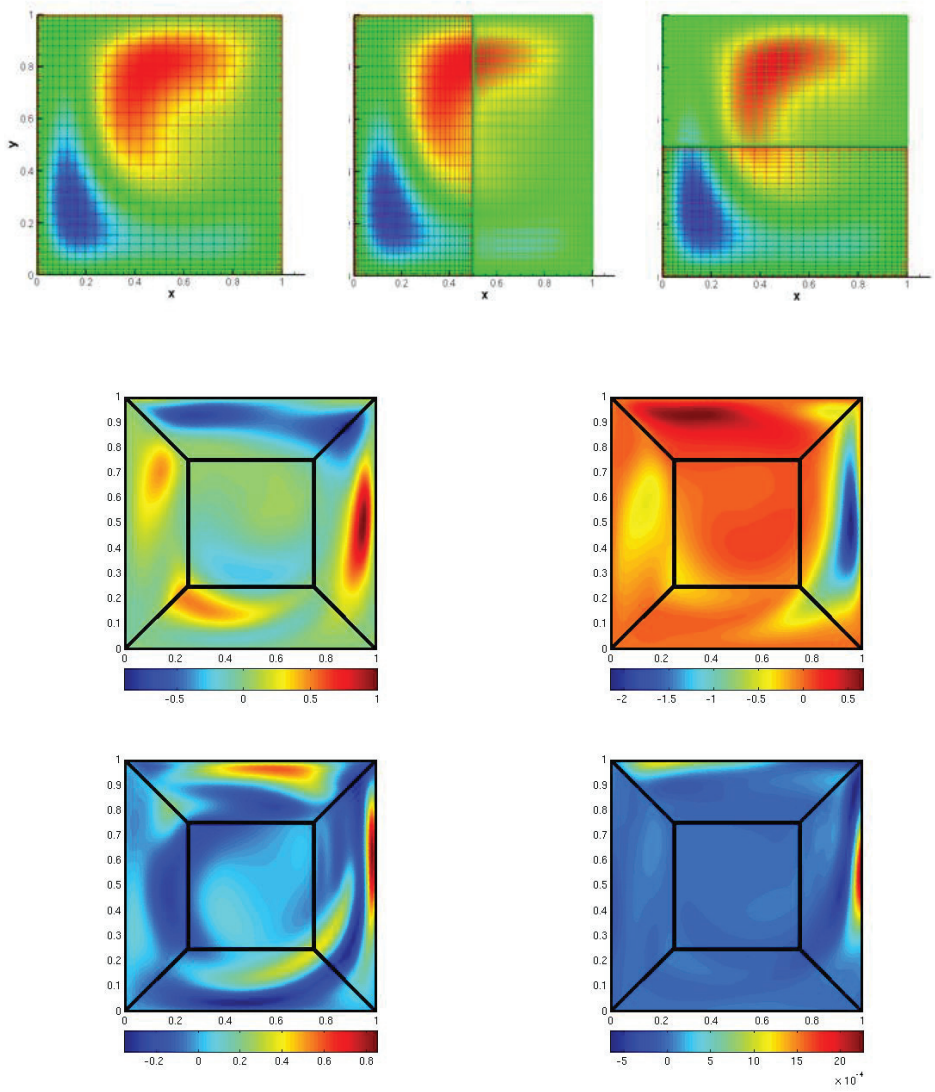


Fig. 5. *Upper:* Amplitude function of the spanwise disturbance velocity component of the least damped eigenmode of lid–driven cavity flow at $Re = 200$ Theofilis (AIAA-2000-1965); Theofilis, Duck & Owen (2004), obtained in single–domain and two different multi–domain configurations with the spectral collocation method. *Lower:* Four components of the most unstable mode of the square cavity flow at Re=1000 $\beta = 17$ obtained using the spectral/$hp-$element method. The mesh used is superposed over the figure.

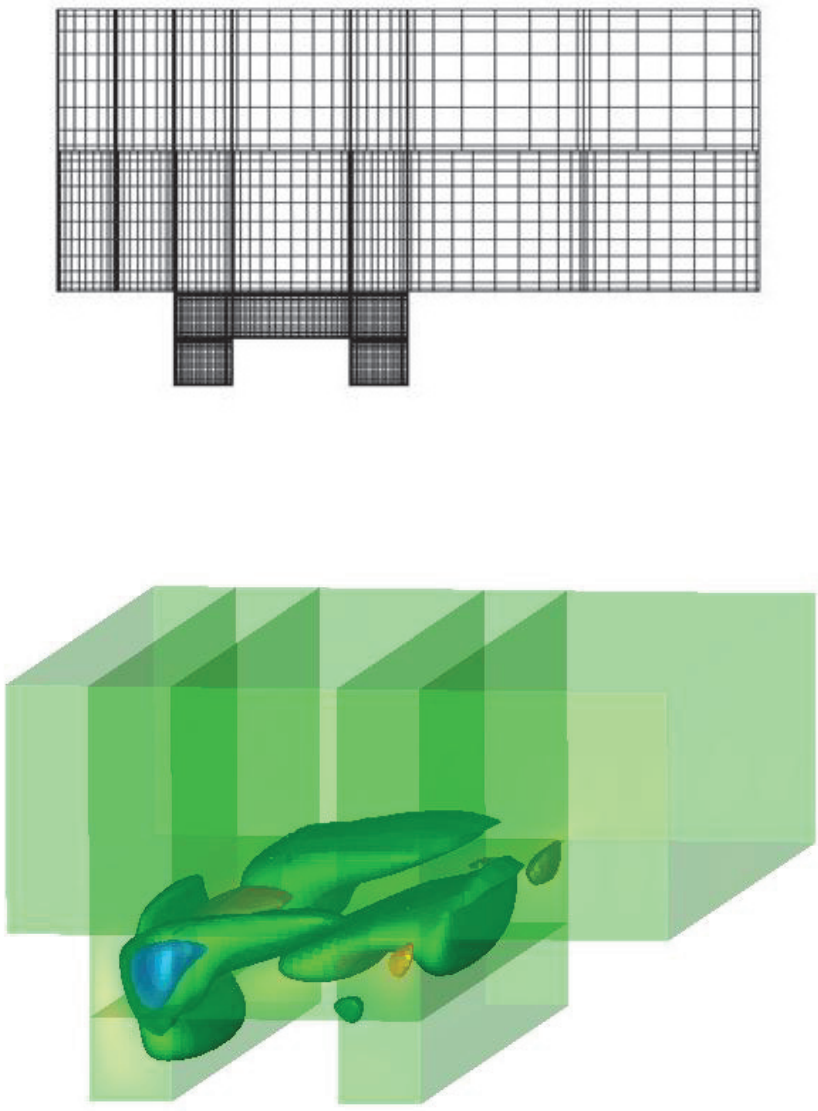


Fig. 6. *Upper:* Non-conforming spectral multidomain discretization of an open–cavity configuration, containing a rectangular object. *Lower:* Perspective view of the least–damped eigenmode of incompressible flow at $Re = 400$ de Vicente et al. (2006).

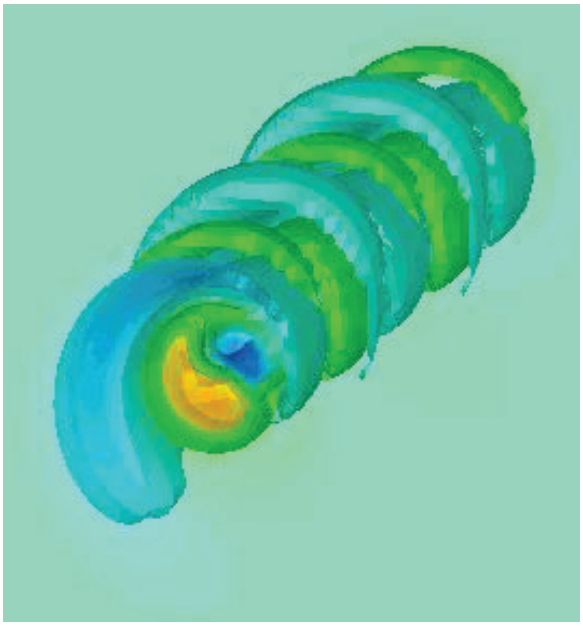


Fig. 7. Superposition upon the steady laminar basic state at $Re = 3180, U_0 = \frac{0.5}{\pi}$, of its most amplified BiGlobal eigenmode at amplitude 0.1%. Axial spatial direction reconstructed using $L_x = 2\pi/\beta$, with $\beta = 3.0$; eigenvalue $\omega_i = 0.04603, \omega_r = 0.06072$ González, Gómez-Blanco & Theofilis (2008).

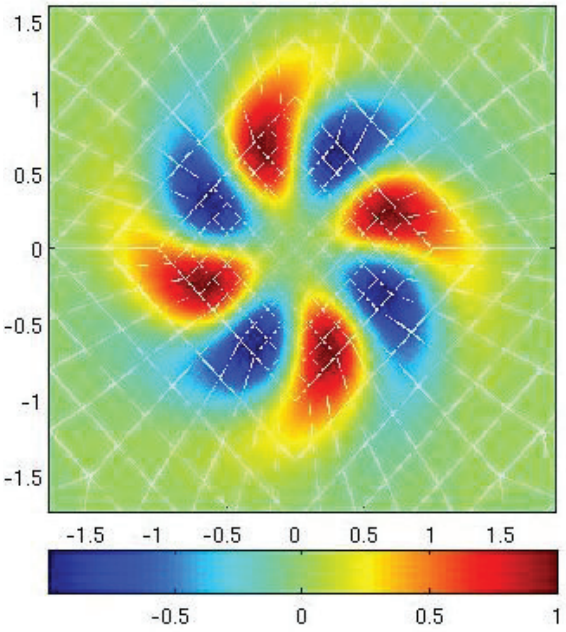


Fig. 8. Axial component of the disturbance velocity of one of the unstable modes(four lobes) of a isolated Batchelor vortex at $Re = 667$ and $\beta = 2.0$.

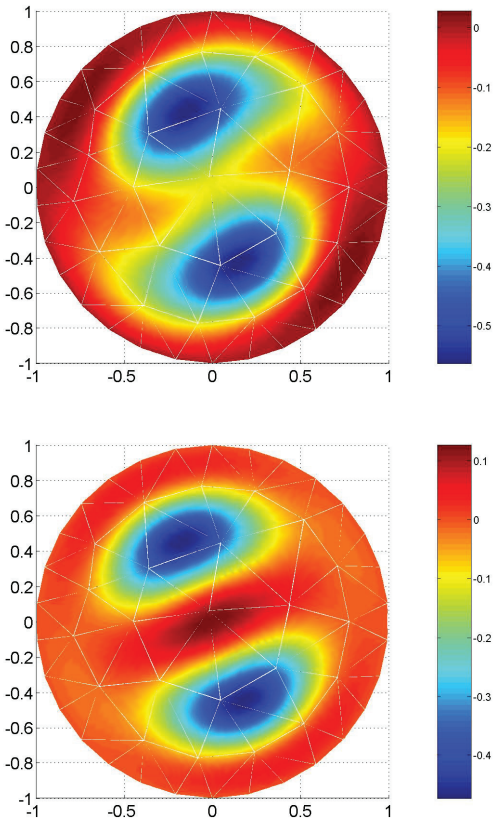


Fig. 9. Leading eigenmodes of HPF flow at $Re = 100, \beta = 1$: shown are the axial disturbance amplitude functions.

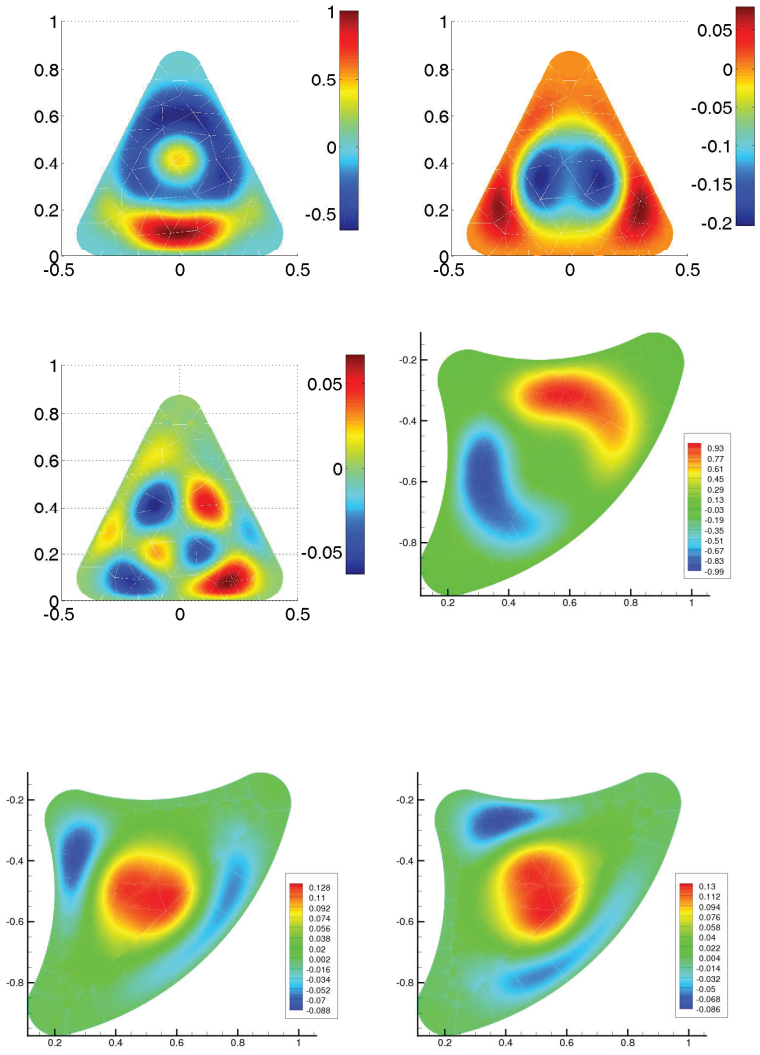


Fig. 10. *Upper:* Amplitude functions of the least-damped eigenmode of geometry "$\mathbf{in}_1$" at $Re = 1000,\ \alpha = 1$ González, Rodríguez & Theofilis (2008).*Lower:* Amplitude functions of the least-damped eigenmode of geometry "$\mathbf{in}_2$" at $Re = 1000,\ \alpha = 1$ González, Rodríguez & Theofilis (2008). Left to right column: $\hat{u}_1,\ \hat{u}_2,\ \hat{u}_3$.

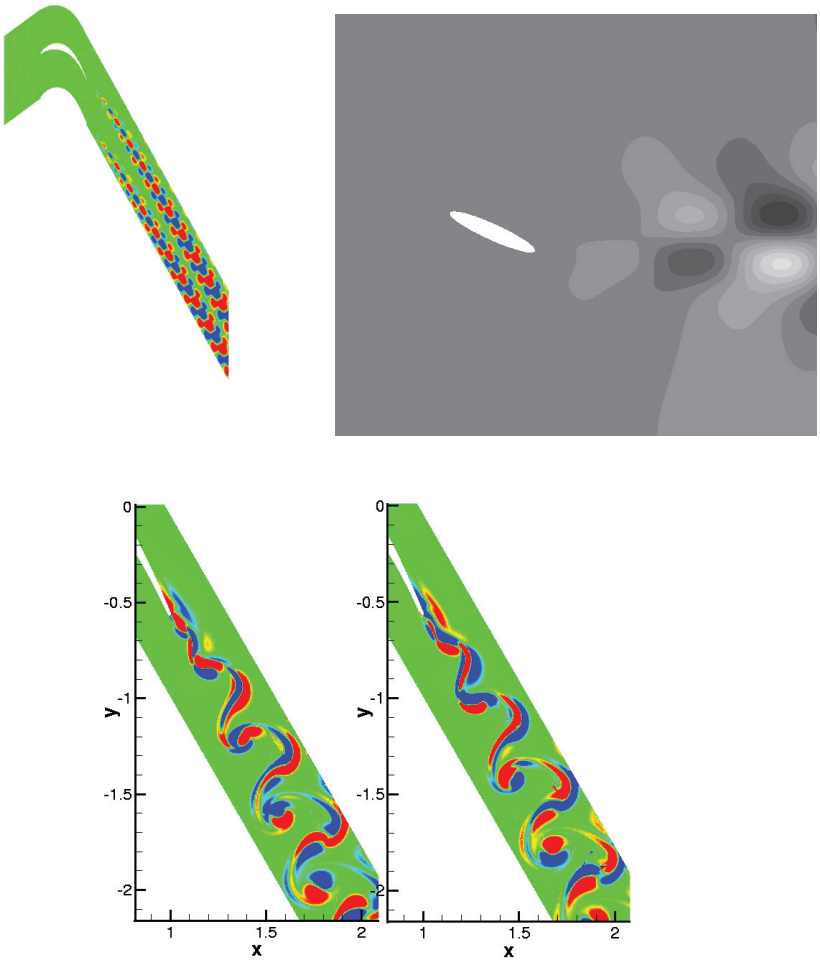


Fig. 11. *Upper-Left:* Leading eigenmode in the wake of the T106-300 LPT flow at $Re = 890$. *Upper-Right:* Leading (wake) eigenmode in flow over an aspect ratio 8 ellipse at $Re = 200$ Kitsios et al. (2008). *Lower:* Leading LPT Floquet mode at $Re = 2000$ Abdessemed et al. (2004).

6. Discussion

Numerical methods for the accurate and efficient solution of incompressible and compressible BiGlobal eigenvalue problems on regular and complex geometries have been discussed. The size of the respective problems warrants particular formulations for each problem intended to be solved: the compressible BiGlobal EVP is only to be addressed when essential compressible flow instability phenomena are expected, e.g. in the cases of shock–induced or supersonic instabilities of hydrodynamic origin or in aeroacoustics research. In all other problems the substantially more efficient incompressible formulation suffices for the analysis. Regarding the issue of time–stepping versus matrix formation approaches, there exist distinct advantages and disadvantages in either methodology; the present article highlights both, in the hope that it will assist newcomers in the field to make educated choices. No strong views on the issue of oder–of–accuracy of the methods utilized are offered, on the one hand because both low– and high–order methods have been successfully employed to the solution of problems of this class and on the other hand no systematic comparisons of the characteristics of the two types of methods have been made to–date. Intentionally, no further conclusions are offered, other than urging the interested reader to keep abreast with the rapidly expanding body of literature on global linear instability analysis.

7. Acknowledgments

The material is based upon work sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under Grants monitored by Dr. J. D. Schmisseur of AFOSR and Dr. Surya Surampudi of EOARD. The views and conclusions contained herein are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

8. References

Abdessemed, N., Sherwin, S. J. & Theofilis, V. (2004). On unstable 2d basic states in low pressure turbine flows at moderate reynolds numbers, number Paper 2004-2541 in *34th Fluid Dynamics Conference and Exhibit*, AIAA, Portland, Oregon.

Abdessemed, N., Sherwin, S. J. & Theofilis, V. (2006). Linear stability of the flow past a low pressure turbine blade, number Paper 2004-3530 in *36th Fluid Dynamics Conference and Exhibit*, AIAA, San Francisco, CA.

Abdessemed, N., Sherwin, S. J. & Theofilis, V. (2009). Linear instability analysis of low pressure turbine flows, *J. Fluid Mech.* 628: 57 – 83.

Albensoeder, S., Kuhlmann, H. C. & Rath, H. J. (2001a). Multiplicity of steady two-dimensional flows in two–sided lid–driven cavities, *Theor. Comp. Fluid. Dyn.* 14: 223–241.

Albensoeder, S., Kuhlmann, H. C. & Rath, H. J. (2001b). Three-dimensional centrifugal-flow instabilities in the lid-driven-cavity problem, *Phys. Fluids* 13(1): 121–136.

Allievi, A. & Bermejo, R. (2000). A characteristic-finite element conjugate gradient algorithm for the navier-sokes equations, *Int. J. Numer. Meth. Fluids* 32(*): 439–464.

Barkley, D. & Henderson, R. D. (1996). Three-dimensional floquet stability analysis of the wake of a circular cylinder, *J. Fluid Mech.* 322: 215–241.

Bewley, T. R. (2001). Flow control: New challenges for a new renaissance, *Progress in Aerospace Sciences* 37: 21–58.

Bres, G. & Colonius, T. (2008). Three-dimensional instabilities in compressible flow over open cavities, *J. Fluid Mech.* 599: 309–339.

Broadhurst, M., Theofilis, V. & Sherwin, S. J. (2006). Spectral element stability analysis of vortical flows, number Paper 2006-2877 in *6th IUTAM Laminar-Turbulent Transition Symposium*, Bangalore, India, Dec 13–17, 2004, pp. 153–158.

Collis, S. S., Joslin, R. D., Seifert, A. & Theofilis, V. (2004). Issues in active flow control: theory, control, simulation and experiment, *Prog. Aero. Sciences* 40: 237–289.

Crouch, J. D., Garbaruk, A. & Magidov, D. (2007). Predicting the onset of flow unsteadiness based on global instability, *J. Comp. Phys.* 224: 924–940.

Cuvelier, C., Segal, A. & van Steenhoven, A. A. (1986). *Finite Element Methods and Navier-Stokes Equations*, D. Reidel Publishing Company.

de Vicente, J., Rodríguez, D., Theofilis, V. & Valero, E. (2011). Stability analysis in spanwise-periodic double-sided lid-driven cavity flows with complex cross-sectional profiles, *Computers and Fluids* 43: 143–153.

de Vicente, J., Valero, E., González, L. & V.Theofilis (2006). Spectral multi-domain methods for biglobal instability analysis of complex flows over open cavity configurations, number Paper 2006-2877 in *36th AIAA Fluid Dynamics Conference and Exhibit*, AIAA, San Francisco, California.

Ding, Y. & Kawahara, M. (1998). Linear stability of incompressible flow using a mixed finite element method, *J. Comp. Phys.* 139: 243–273.

Dobrinsky, A. & Collis, S. S. (2000). Adjoint parabolized stability equations for receptivity prediction, *AIAA Paper 2000–2651*.

Duc, A. L., Sesterhenn, J. & Friedrich, R. (2006). Instabilities in compressible attachment–line boundary layers, *Phys. Fluids* 18: 044102–1–16.

Giannetti, F. & Luchini, P. (2007). Structural sensitivity of the first instability of the cylinder wake, *J. Fluid Mech.* 581: 167 – 197.

González, L., Gómez-Blanco, R. & Theofilis, V. (2008). Eigenmodes of a counter-rotating vortex dipole, *AIAA J.* DOI: 10.2514/1.35511: to appear.

González, L. M. & Bermejo, R. (2005). A semi-langrangian level set method for incompressible navier-stokes equations with free surface, *Int. J. Numer. Meth. Fluids* 49(*): 1111–1146.

González, L., Rodríguez, D. & Theofilis, V. (2008). On instability analysis of realistic intake flows, number Paper 2008-4380 in *38th Fluid Dynamics Conference and Exhibit*, AIAA, Seattle, WA.

González, L., Theofilis, V. & Gómez-Blanco, R. (2007). Finite element numerical methods for viscous incompressible biglobal linear instability analysis on unstructured meshes, *AIAA J.* 45: 840–855.

Gonzalez L, Theofilis V & Gomez-Blanco R (2007). Finite-element numerical methods for viscous incompressible biglobal linear instability analysis on unstructured meshes, *AIAA Journal* 45(4): 840–855.

Hein, S., Hohage, T., Koch, W. & Schoberl, J. (2007). Acoustic resonances in a high-lift configuration, *J. Fluid Mech.* 582: 179–202.

Henningson, D. S. (1987). Stability of parallel inviscid schear flow with mean spanwise variation, *Tech. Rep. FFA-TN-1987-57*, Bromma, Sweden.

Hill, D. C. (1992). A theoretical approach for the restabilization of wakes, *AIAA Paper 92–0067*.

Karniadakis, G. E. & Sherwin, S. J. (2005). *Spectral/hp element methods for CFD*, OUP.

Kitsios, V., Rodríguez, D., Theofilis, V., Ooi, A. & Soria, J. (2008). Biglobal instability analysis of turbulent flow over an airfoil at an angle of attack, number Paper 2008-4384 in *38th Fluid Dynamics Conference and Exhibit*, AIAA, San Francisco, CA.

Koch, W. (2007). Acoustic resonances in rectangular open cavities, *AIAA J.* 43: 2342–2349.

Leriche, E., Gavrilakis, S. & Deville, M. O. (1998). Direct simulation of the lid-driven cavity flow with chebyshev polynomials, *in* K. D. Papailiou (ed.), *Proc. 4th European Computational Fluid Dynamics Conference*, Vol. 1(1), ECCOMAS, pp. 220–225.

Lessen, M., Sadler, S. G. & Liu, T. Y. (1968). Stability of pipe poiseuille flow, *Phys. Fluids* 11: 1404–1409.

Lin, R.-S. & Malik, M. R. (1996a). On the stability of attachment-line boundary layers. Part 1. the incompressible swept hiemenz flow, *J. Fluid Mech.* 311: 239–255.

Lin, R.-S. & Malik, M. R. (1996b). On the stability of attachment-line boundary layers. Part 2. the effect of leading-edge curvature, *J. Fluid Mech.* 333: 125 – 137.

Mack, L. M. (1984). Boundary layer linear stability theory, *AGARD–R–709 Special course on stability and transition of laminar flow*, pp. 3.1–3.81.

Malik, M. R. (1991). Numerical methods for hypersonic boundary layer stability, *J. Comp. Phys.* 86: 376–413.

Marquet, O., Sipp, D. & Jacquin, L. (2006). Global optimal perturbations in separated flow over a backward-rounded -step, number Paper 2006-2879 in *3rd Flow Control Conference*, AIAA, San Francisco, CA.

Mayer, E. W. & Powell, K. G. (1992). Viscous and inviscid instabilities of a trailing vortex, *Journal of Fluid Mechanics* 245: 91–114.

Morse, P. M. & Feshbach, H. (1953). *Methods of Theoretical Physics, Parts I, II*, McGraw-Hill.

Nayar, N. & Ortega, J. M. (1993). Computation of selected eigenvalues of generalized eigenvalue problems, *J. Comput. Physics* 108: 8–14.

Poliashenko, M. & Aidun, C. K. (1995). A direct method for computation of simple bifurcations, *Journal of Computational Physics* 121: 246–260.

Pralits, J. O. & Hanifi, A. (2003). Optimization of steady suction for disturbance control on infinite swept wings, *Phys. Fluids* 15(9): 2756 – 2772.

Robinet, J.-C. (2007). Bifurcations in shock-wave/laminar-boundary-layer interaction: global instability approach, *J. Fluid Mech.* 579: 85–112.

Rodríguez, D. (2008). On instability and active control of laminar separation bubbles, *Proceedings of the 4th PEGASUS-AIAA Student Conference, Prague* .

Rodríguez, D. & Theofilis, V. (2008). On instability and structural sensitivity of incompressible laminar separation bubbles in a flat–plate boundary layer, number Paper 2008-4148 in *38th AIAA Fluid Dynamics conference, Jan. 23-26, 2008*, AIAA, Seattle, WA.

Rodríguez, D. & Theofilis, V. (2010). Structural changes of laminar separation bubbles induced by global linear instability, *J. Fluid Mech.* p. (to appear).

Sahin, M. & Owens, R. G. (2003). A novel fully-implicit finite volume method applied to the lid-driven cavity problem. part ii. linear stability analysis, *Int. J. Numer. Meth. Fluids* 42: 79–88.

Salwen, H., Cotton, F. W. & Grosch, C. E. (1980). Linear stability of poiseuille flow in a circular pipe, *J. Fluid Mech.* 98: 273–284.

Schmid, P. J. & Henningson, D. S. (2001). *Stability and transition in shear flows*, Springer-Verlag, New York.

Tatsumi, T. & Yoshimura, T. (1990). Stability of the laminar flow in a rectangular duct, *J. Fluid Mech.* 212: 437–449.

Theofilis, V. (2000). Global linear instability in laminar separated boundary layer flow., *in* H. Fasel & W. Saric (eds), *Proc. of the IUTAM Laminar-Turbulent Symposium V*, Sedona, AZ, USA, pp. 663 – 668.

Theofilis, V. (2001). Inviscid gglobal linear instability of compressible flow on an elliptic cone: algorithmic developments, *Technical Report F61775-00-WE069*, European Office of Aerospace Research and Development.

Theofilis, V. (2003). Advances in global linear instability of nonparallel and three-dimensional flows, *Prog. Aero. Sciences* 39 (4): 249–315.

Theofilis, V. (2009a). On multidimensional global eigenvalue problems for hydrodynamic and aeroacoustic instabilities, number Paper 2009-0007 in *47th Aerospace Sciences Meeting 5–8 Jan. 2009*, AIAA, Orlando, FL.

Theofilis, V. (2009b). The role of instability theory in flow control, *in* R. D. Joslin & D. Miller (eds), *Active Flow Control*, AIAA Progress in Aeronautics and Astronautics Series, AIAA.

Theofilis, V. (2011). Global linear instability, *Annu. Rev. Fluid Mech.* 43: 319–352.

Theofilis, V. (AIAA-2000-1965). Globally unstable basic flows in open cavities, *6th AIAA Aeroacoustics Conference and Exhibit* .

Theofilis, V., Barkley, D. & Sherwin, S. J. (2002). Spectral/hp element technology for flow instability and control, *Aero. J.* 106(619-625).

Theofilis, V. & Colonius, T. (2003). An algorithm for the recovery of 2- and 3-d biglobal instabilities of compressible flow over 2-d open cavities, AIAA Paper 2003-4143.

Theofilis, V. & Colonius, T. (2004). Three-dimensional instablities of compressible flow over open cavities: direct solution of the biglobal eigenvalue problem, 34th Fluid Dynamics Conference and Exhibit, AIAA Paper 2004-2544, Portland, Oregon.

Theofilis, V., Duck, P. W. & Owen, J. (2004). Viscous linear stability analysis of rectangular duct and cavity flows, *J. Fluid. Mech.* 505: 249–286.

Theofilis, V., Fedorov, A. & Collis, S. S. (2004). Leading-edge boundary layer flow: PrandtlÂŧ's vision, current developments and future perspectives, *IUTAM Symposium "One Hundred Years of Boundary Layer Research", Göttingen, Germany, August 12-14, 2004*, Springer, pp. 73–82.

Theofilis, V., Fedorov, A., Obrist, D. & Dallmann, U. C. (2003). The extended görtler-hämmerlin model for linear instability of three-dimensional incompressible swept attachment-line boundary layer flow, *J. Fluid Mech.* 487: 271–313.

Theofilis, V., Hein, S. & Dallmann, U. C. (2000). On the origins of unsteadiness and three-dimensionality in a laminar separation bubble, *Phil. Trans. Roy. Soc. (London) A* 358: 3229–3246.

Trefethen, L. N. (2000). *Spectral Methods in Matlab*, SIAM.

Zhigulev, V. N. & Tumin, A. (1987). Origin of turbulence, *Nauka* .

Part 2

Flight Performance, Propulsion, and Design

8

Rotorcraft Design for Maximized Performance at Minimized Vibratory Loads

Marilena D. Pavel
Faculty of Aerospace Engineering, Delft University of Technology
The Netherlands

1. Introduction

Rotorcraft (helicopters and tiltrotors) are generally reliable flying machines capable of fulfilling missions impossible with fixed-wing aircraft, most notably rescue operations. These missions, however, often lead to high and sometimes excessive pilot workload. Although high standards in terms of safety are imposed in helicopter design, studies show that "*it is ten times more likely to be involved in an accident in a helicopter than in a fixed-wing aircraft*" (Iseler et. al. 2001). According to World Aircraft Accident Summary (WAAS, 2002), nearly 45 percent of all accidents of single-piston helicopters is attributed to pilot loss of control, where - because of various causes, often involving vibrations, high workload, and bad weather - a pilot loses control of the helicopter and crashes, sometimes with fatal consequences. Quoting the Royal Netherlands Air Force ('Veilig Vliegen' magazine, 2003), "*for helicopters there is a considerable number of inexplicable incidents (…) which involved piloting loss of control*". The situation is likely to get worse, as rotorcraft missions are becoming more difficult, demanding high agility and rapid manoeuvring, and producing more violent vibrations. (Kufeld & Bousman, 1995; Hansford & Vorwald, 1996; Datta & Chopra, 2002)
The primary cause of pilot control difficulties and high-workload situations is that even modern helicopters often have poor Handling Qualities (HQs) (Padfield, 1998). Cooper and Harper (Cooper & Harper, 1969), pioneers in this subject, defined these as: "*those qualities or characteristics of an aircraft that govern the ease and precision with which a pilot is able to perform a mission*". Below, the current practice in rotorcraft handling qualities assessment will be discussed, introducing the key problem addressed in this chapter.

1.1 State-of-the-art in rotorcraft handling qualities – The aeronautical design standard ADS-33

Helicopter handling qualities used to be assessed with requirements defined for fixed-wing aircraft, as stated in the FAR (civil) and MIL (military) standards. In the 1960's, however, it became clear that these standards were not sufficient (Key, 1982). Helicopters have *strong cross-coupling effects* between longitudinal and directional controls, their behaviour is *highly non-linear* and requires *more degrees of freedom* in modelling than the rigid-body models used for aircraft. Therefore, the MIL-H-8501A standard (MIL-H-8501A, 1962) was developed. This standard was used up until mid 1980's. From a safety perspective, these requirements were merely 'good minimums', and a new standard was developed in the 1970's, that is used up until today, the Aeronautical Design Standard ADS-33 (ADS-33, 2000)

The crucial point, understood by ADS-33, is that helicopter HQ requirements need to be related to the mission executed, as this will determine the needed pilot effort. E.g., a shipboard landing at night and in high sea with strong ship motions demands more precision of control from the pilot than when flying in daytime and good weather.2 ADS-33 introduced handling qualities *metrics* (HQM), a *combination* of flight parameters such as rate of climb, turn rate, etc., that reflect how much manoeuvre-capability the pilot has *per specific mission*. These metrics are then mapped into handling qualities criteria (HQC) that yield boundaries between 'good' (Level 1), 'satisfactory' (Level 2) and 'poor' (Level 3) HQs.[1]

Despite their importance for the helicopter safety, its operators and, above all, the helicopter pilots, achieving good handling qualities is still mainly a secondary goal in helicopter design. The first phase in helicopter design is the 'conceptual design' phase in which the main rotor and fuselage parameters are established, based on desired performance and, to some extent, vibration criteria.20 Only in the following phase, that of preliminary design, are 'high-fidelity' simulation models developed and the handling qualities considered. The high fidelity models allow an analysis of helicopter behaviour for various flight conditions. Applying the ADS-33 metrics/criteria to these models results in predicted levels of HQs. When these are known, the *experimental* HQ assessment begins, illustrated in Fig. 1.

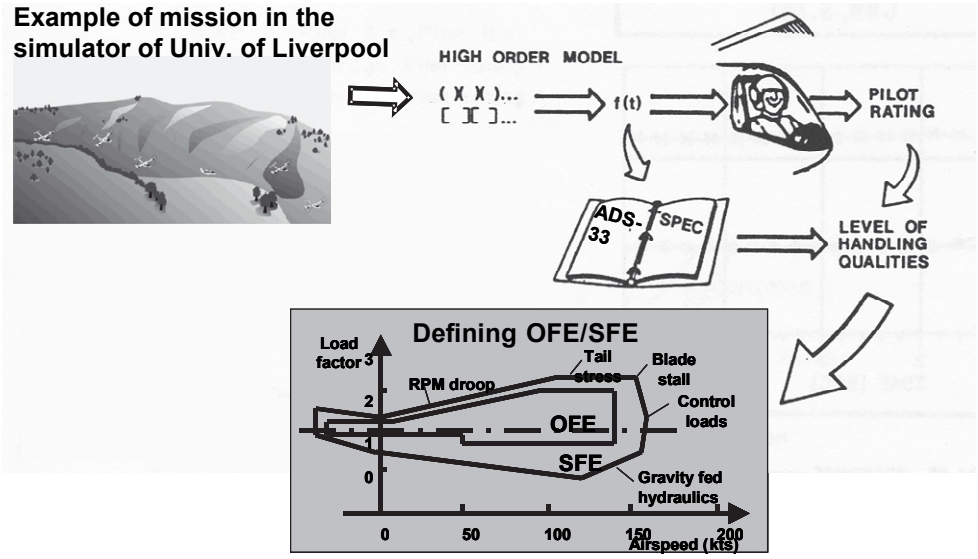


Fig. 1. Experimental assessment of helicopter handling qualities

[1] Level 1 HQs means that the rotorcraft is satisfactory without improvements required from the pilot's perspective. Level 2 HQs means that pilots can achieve adequate performance, but with compensation; at the extreme of Level 2, the mission is flyable, but pilots have little capacity for other duties and can not sustain flying for longer periods without the danger of pilot error. Level 3 HQs is unacceptable, it describes rotorcraft behaviour in extreme situations, like the loss of critical flight control systems.

In this process, first databases of missions and environments are defined, according to customer requirements. Manoeuvrability, i.e., how easy can pilots guide the helicopter, and agility, i.e., how quickly can they change its flight direction, are critical performance demands. Second, experienced test pilots are asked to fly the missions in a flight simulator. Through insisting pilots to execute the mission very thoroughly, one aims to expose deficient handling qualities. Pilots rate the HQs they experience in each manoeuvre flown, generally using the Cooper-Harper rating scale.[2] Third, Operational Flight Envelopes (OFEs) and Service Flight Envelopes (SFEs) are defined, based on a mapping of the HQ ratings. OFEs represent the limits within which the helicopter must be able to operate in order to accomplish the operational missions. SFEs stem from the helicopter limits and are expressed in terms of any parameters believed necessary to ensure safety, see Fig. 1.

In this first experimental assessment of the helicopter's handling qualities, the first problems arise, as more often than not, large differences arise between the theoretical predictions and the experimentally-determined pilot judgments. The gaps that occur are bridged by applying optimisation techniques using the simulation models developed in preliminary design, to improve the designs of helicopter, load alleviation system and flight control system (Celi, 1991; Celi, 1999; Celi, 2000; Fusato & Celi, 2001, Fusato & Celi, 2002a,2002b; Ganduli, 2004; Sahasrabudhe & Celi, 1997). Computational approaches have three important disadvantages, however. First, many designs that "roll out" of the procedure are unfeasible, the optimisation "pushing" the solution along the boundaries of the problem and not inside of the feasible region. Second, optimising for ADS-33 requires calculations of the helicopter time-domain responses, and the numerical methods become computationally very intensive (Sahasrabudhe & Celi, 1997; Tischler et. al., 1997) Third, and most important, the lack of quantitative, validated helicopter pilot models, capable of accurately predicting the effects of helicopter vibrations on pilot control behaviour, prevents the proper inclusion of pilot-centred considerations in mathematical optimization techniques (Mitchell et. al., 2004; Tischler et. al., 1996)

Whereas the first and second disadvantages are common in multi-dimensional design problems, the lack of knowledge on how helicopter vibrations affect pilot performance is a typical and fundamental problem of modern helicopter design (Mitchell et. al., 2004, Padfield, 1998). Although ADS-33 proposes criteria and missions regarding helicopter limits, these only characterize a helicopter's performance, and do not require an adequate knowledge of helicopter vibratory loads (Kolwey, 1996; Tischler et. al., 1996). This shortcoming stems from the fact that, when the ADS-33 criteria were defined, helicopter missions were not so demanding, and the vibratory loads associated with them were low. In the last twenty years, however, ever-increasing performance requirements and extended flight envelopes were defined, for reasons of heavy competition, demanding manoeuvres that impose heavy vibrations on both structure and pilot. These vibrations, combined with cross-coupling effects, rapidly lead to pilot overload and degradation in performance (Padfield, 2007).

Key Problem: The current practice of assessing rotorcraft handling qualities reveals significant gaps in the ADS-33 HQ criteria, especially regarding the effects of vibrations.

[2] The Cooper-Harper rating scale runs from 1 to 10. Rates from 1 to 3 1/2 correspond to Level 1 HQs; rates from 31/2 to 61/2 correspond to Level 2 HQs, and rates from 61/2 up to 10 correspond to Level 3 HQs.

1.2 Goal

The design of high-performance rotorcraft has become an arduous process, regularly leading to surprises, demanding 'patches' to safety-critical systems, and needing more iterations than expected, all contributing in very high costs. Unmistakably, the helicopter and flight control system design teams do not have up-to-date criteria to adequately assess the effects of helicopter vibrations on its handling qualities. There is an urgent need for a much more fundamental understanding of how helicopter vibrations affect pilot control behaviour (Mitchell et. al., 2004) and for new tools to incorporate this knowledge as early as possible in the design of both helicopter and flight control system. The goal of this chapter is to portray a novel approach to rotorcraft handling qualities (HQs) assessment by defining a set of consistent, complementary metrics for agility and structural loads pertaining to vertical manoeuvres in forward flight. These metrics can be used by the designer for making trade—offs between agility and vibrational/load suppression. The emphasis of the chapter will be on agility characteristics in the pitch axis applied to helicopter and tiltrotor. Especially in such new configurations, the proposed approach could be particularly useful as the performance tools for fixed-wing mode and helicopter mode must merge together within new criteria (Padfield, 2008).

The chapter is structured as follows: The second section will present an overview of traditional metrics for measuring pitch agility; The third section will present some alternative metrics proposed in the 90's for better capturing the transient characteristics of the agility; Then, based on the rational developments of the metrics from the previous two sections, fourth section will propose the new approach that can better quantify the agility from the designer point of view. Finally, general conclusions and potential extension of this work will be discussed.

2. Traditionally design of aircraft for pitch agility

One of the most important flying quality concepts defining the upper limits of performance is the so-called "agility". Generally, it is well known that the level of performance achieved by the pilot depends on the task complexity. Fig. 2 presents generically this situation, showing that there is a line of saturation up to which the pilot is able to perform optimally the specified mission; increasing the task difficulty above this line leads quickly to stress, panic and even incapacity to cope anymore with the task complexity and blocking, sometimes with fatal consequences.

It is difficult to point precisely to the origins of the concept of agility but probably these go back to the moment when it was realized that, in a combat, a "medium performance" fighter could win over its superior opponent if the first aircraft possesses the potential for faster transient motions, i.e. superior agility. In its most general sense, the concept of agility is defined with respect to the overall combat effectiveness in the so-called "Operational agility". Operational agility according to measures the '*ability to adapt and respond rapidly and precisely, with safety and poise, to maximize mission effectiveness*'(McKay, 1994). In the mid 80's a strong wave of interest arose in seeking metrics and criteria that could quantify the aircraft agility (Mazza, 1990; McKay, 1994). However, there have been developed almost as many criteria of agility as there were investigators in the field. The problem was partially due to the lack of coordination in the research studies performed but also due to a disagreement on the most fundamental level: there simply was very little agreement on what agility was.

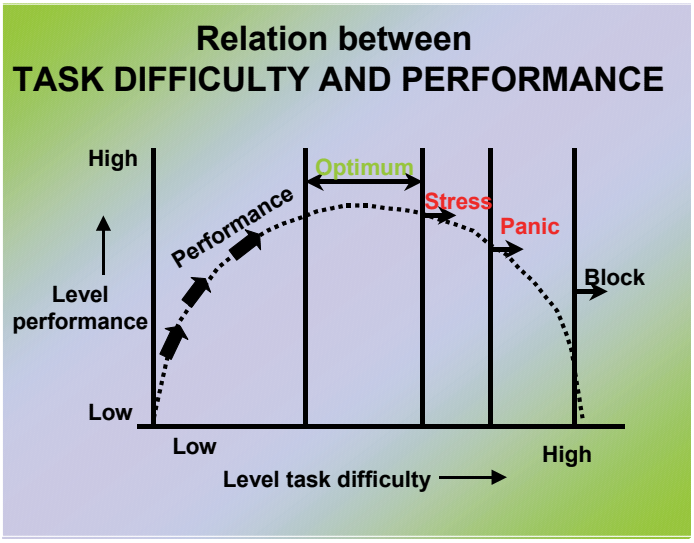


Fig. 2. Correlation between task difficulty and performance

Within the framework of operational agility one can see agility as a function of the airframe, avionics, weapons and pilot. Airframe agility is probably the most crucial component in the operational agility as it is designed in from the onset and cannot be added later. The present chapter focuses on airframe agility and within this, the chapter will relate to the airframe agility in the pitch axis.

A large number of agility metrics have been proposed during the years to determine the aircraft realm of agility. The AGARD Working Group 19 on Operational Agility (McKay, 1994) put together all the different metrics and criteria existing on agility and fit them into a generalizable framework for further agility evaluations. The present section presents the traditional approach on pitch agility using as example a tiltrotor aircraft. This specific aircraft combines the properties of both fixed and rotary-wing aircraft and can be used to define a unified approach in the agility requirements at both fixed and rotary-wings.

The tiltrotor example to be investigated in this study is the Bel XV-15 aircraft. Next, as vehicle model, the FLIGHTLAB model of the Bell XV-15 aircraft as developed by the University of Liverpool (this model is designated as FXV-15) will be used. For a complete description of this model and the assumptions made the reader is referred to (Manimala et. al., 2003). For the tiltrotor in helicopter mode, the pilot's controls command pitch through longitudinal cyclic, roll through differential collective (lateral cyclic is also provided for trimming), yaw through differential longitudinal cyclic and heave through combined collective. In airplane mode, the pilot controls command conventional elevator, aileron and rudder (a small proportion of differential collective is also included).

Pitch agility is the ability to move, rapidly and precisely, the aircraft nose in the longitudinal plane and complete with easiness that movement. This implies that in the agility analysis one has to search for sample manoeuvres to be carried out by the flight vehicle dominated by high flight path changes and high rate of change of longitudinal acceleration which can give a good picture of the agility characteristics. As starting point in this discussion on pitch

agility we will consider the kinematics of a sharp pitch manoeuvre – a simple example of this type is the tiltrotor trying to fly over an obstacle (see Fig. 3). Assume that the manoeuvre is executing starting from different forward speeds (helicopter mode 60 kts and 120kts; airplane mode 120 kts and 300 kts) and the manoeuvre aggressiveness is varied by varying the pulse duration (from 1 to 5 sec). The pilot flies the manoeuvre by giving a pulse input in the longitudinal cyclic stick of 1 inch amplitude.

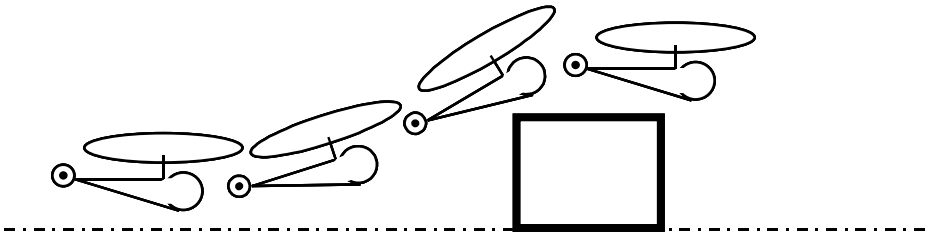


Fig. 3. Executing an obstacle-avoid manoeuvre in the pitch axis

2.1 Transient metrics

The first class of metrics developed to quantify the agility corresponds to the so-called "transient metrics". The transient class contains metrics which can be calculated at any moment for any manoeuvre. For pitch agility these metrics are pitch rate (entitled attitude manoeuvrability metric) and accelerations along the axes a_x, a_y, a_z (entitled manoeuvrability of the flight path). These metrics are next studied for the pull up manoeuvres flown with the FXV-15 in a 1 second pulse given from the initial trim at 120kts in helicopter mode and 300 kts in airplane mode. The presentation of the transient metric information is best achieved through a time history plot. Fig. 4 presents the transient metrics parameters of pitch rate q and vertical acceleration n_z (in the form of normal load factor). Looking at Fig. 4 one may see local maxima in the metric parameters q and n_z illustrating peak events in the agility characteristics. This clearly demonstrates that in a "real" manoeuvre sequence, the agility characteristics occur at key moments, depending on the manoeuvre.

2.2 Experimental metrics

The above conclusion gave the idea to develop a new class of agility metrics, the so-called "experimental metrics" formulated as discrete parameters during a real manoeuvre sequence. These metrics are actually the basic building blocks for understanding the agility and can be related to flying qualities and aircraft design. The metrics describing pitch agility during aggressive manoeuvring in vertical plane were defined by (Murphy et. al., 1991) and are described in the next section. They referred to the ability of an aircraft to pint the nose in at an opponent and commented that what was not clear in such manoeuvres was the behaviour of the flight path. Was the nose pointing w.r.t. the velocity vector or did it include the flight path bending or perhaps both? The authors noted that longitudinal stick displacements would be expected to command the flight path in addition to the aircraft nose pointing pitch angle for agile aircraft. The study pointed out that current aircraft behave differently in the high speeds and slow speed regimes. In the high speed case the flight path displaced as per the nose pointing displacement. The low speed case exhibited no flight path or even opposite flight path displacements.

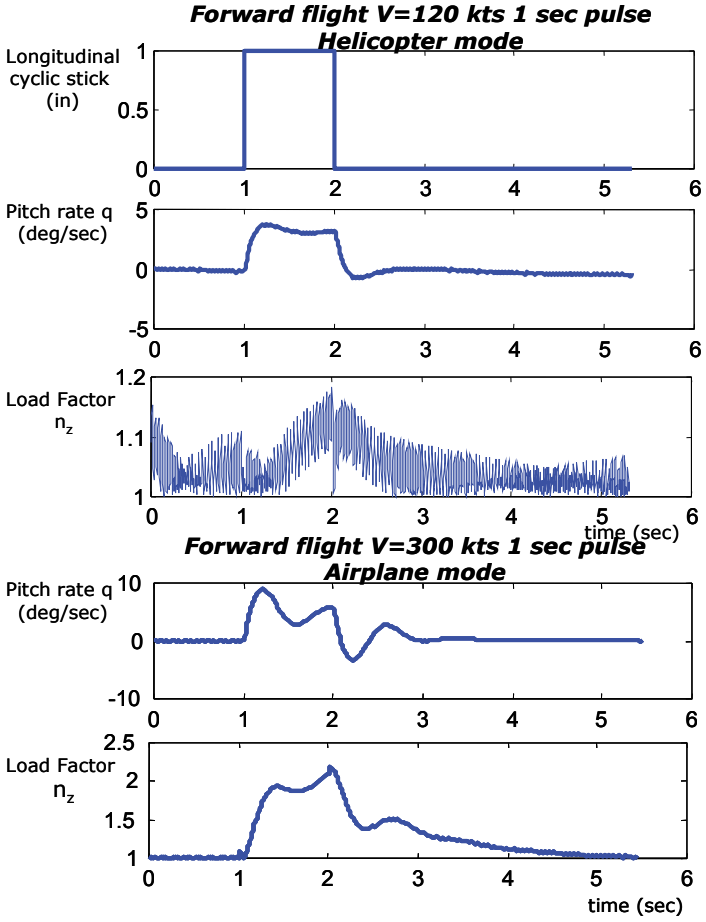


Fig. 4. Transient agility metrics for pull-up manoeuvres with the tiltrotor

2.2.1 Peak and time to peak pitch rates

The peak and time to peak pitch rates metrics were proposed by (Murphy et. al., 1991) for fixed wing aircraft. These metrics measure the time to reach peak pitch rate and the corresponding pitch rate. Fig. 5 presents charts of peak pitch rate and time to reach this peak as a function of the velocity for the tiltrotor flying pull-up manoeuvres of increasing pulse duration. The pull-up manoeuvres are executed gradually increasing the velocity and the nacelle angle from the helicopter mode (90deg nacelle, hover and 60 kts) to conversion (60deg nacelle 120 kts) and ending in airplane mode (0deg nacelle 200 kts). Looking at these figures one can see that as the velocity increases the pilot is able to achieve higher pitch rates, the time to achieve these peaks being and faster especially if the pulse duration is short. As attributes, the peak and time to peak pitch rates metrics have the advantage that can be related to design.

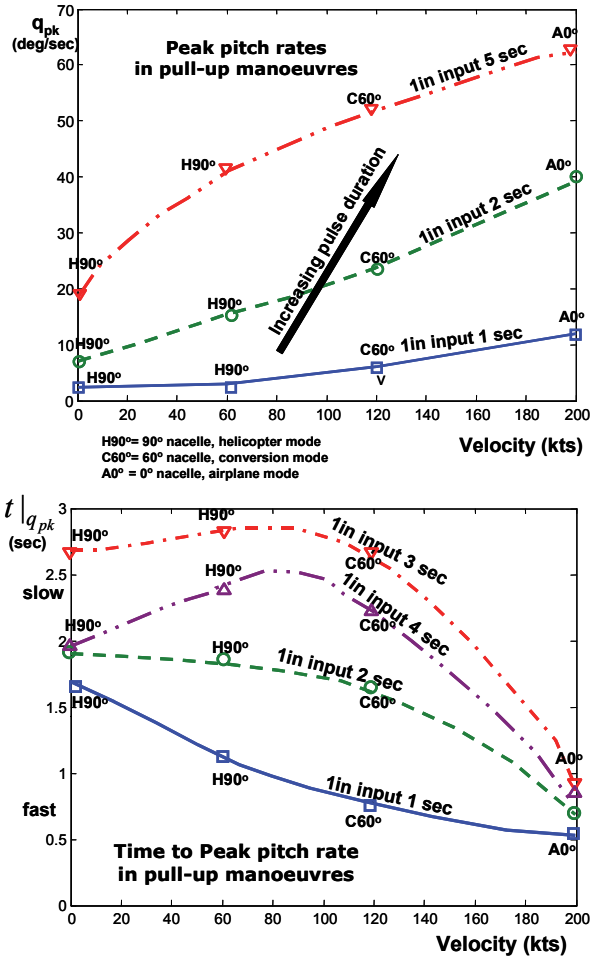


Fig. 5. Peak and time to peak pitch rates in pull-up manoeuvres

2.2.2 Peak and time to peak pitch accelerations

(Murphy et. al., 1991) considered the so-called peak and timer to peak pitch accelerations as the primary metrics for pitch motion agility. The time to peak acceleration provides insight into the jerk characteristics of pitch motion: if it is too slow, then the pilot may complain that the aircraft is too sluggish for tracking-type tasks; if it is too fast, then the pilot may complain of jerkiness or over-sensitivity. Fig. 6 presents charts of peak and time to leak pitch acceleration as a function of velocity when flying pull-ups manoeuvres. One can see that as the velocity increases the pilot is able to obtain higher pitch accelerations but as is passing from the helicopter to aircraft mode this capability diminishes. For fixed wing aircraft, (Murphy et. al., 1991) commented on the differences in the data for the peak accelerations in the body and wind axes. This effect has implications on the pilot selection of flight path or nose pointing control during manoeuvring.

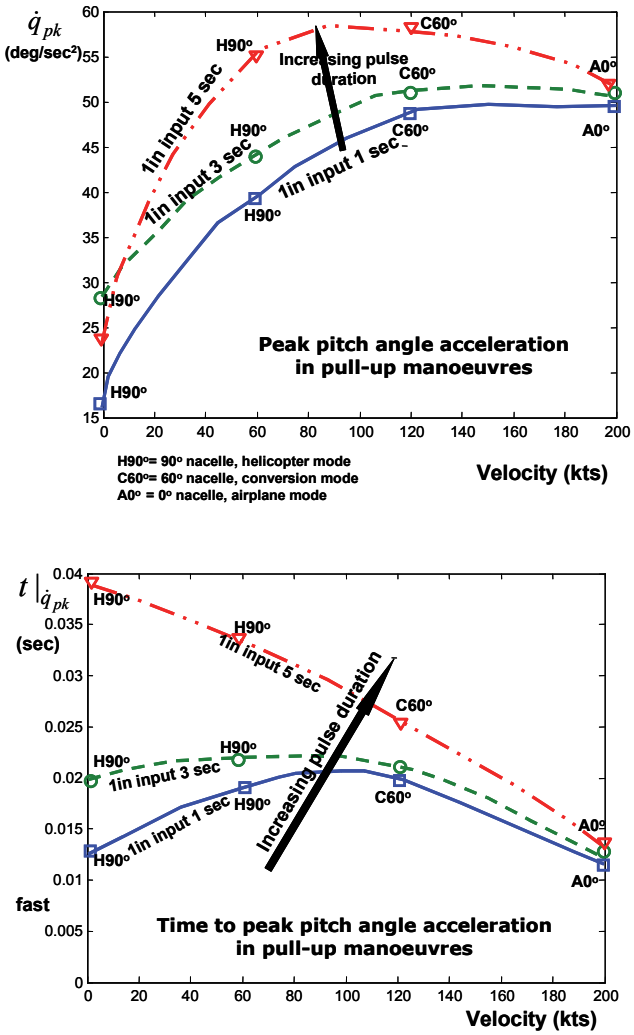


Fig. 6. Peak and time to peak pitch angle acceleration in pull-ups with the tiltrotor

2.2.3 Peak and time to peak load factor

Peak and time to peak load factor metrics describe the peak and the transition time to the peak normal load factor during a manoeuvre in pitch axis. They can be used at best to determine the flight path bending capability of an aircraft. Fig. 7 presents these two metrics as a function of the velocity for the tiltrotor example. One may see that as the velocity increases the pilot is able to pull more g's as going from the airplane to helicopter mode.

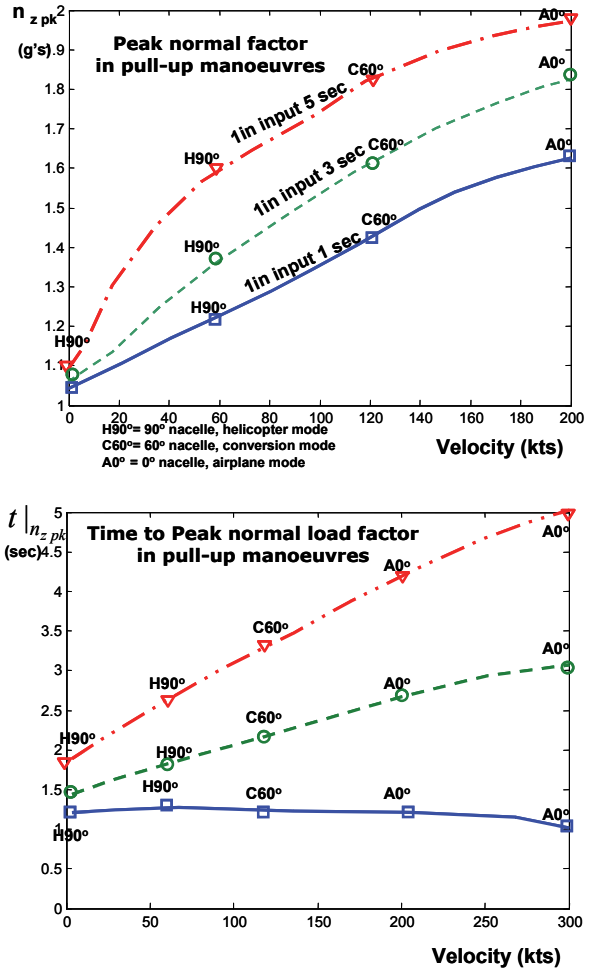


Fig. 7. Peak and time to peak normal load factor

2.2.4 Pitch attitude quickness parameter

One of the most important pitch agility metrics introduced by ADS-33 helicopter standard (ADS-33, 2000) is the so-called "pitch attitude quickness" parameter and is defined as the ratio of the peak pitch rate to the pitch angle change:

$$Q_\theta \overset{def}{=} \frac{q_{pk}}{\Delta\theta}\left(\sec^{-1}\right) \tag{1}$$

The advantage of this parameter is that it was linked to handling qualities so that potential bounds for agility could be identified. In this sense, ADS-33 presents HQs boundaries for the pitch quickness parameter as a function of the minimum pitch angular change $\Delta\theta_{min}$

(considered as the pitch angle corresponding to a 10% decay from q_{pk}). These boundaries are defined to separate different quality levels, but because they relate too to an agility metric, they become now boundaries of available agility. Fig. 8 illustrates the attitude quickness charts for the tiltrotor executing pull-up maneuvers of 1 to 5 sec 1in amplitude input at 60, 120 and 300 kts in helicopter and airplane mode. The figure shows also the Level 1/2 boundaries as defined by 1) ADS-33 for a general mission task element, low speed helicopter flight (<45kts) and 2) MIL STD 1797A for fixed wing aircraft. One may see that whereas in helicopter mode FXV-15 hardly meets Level 1 performance in ADS-33 standard, being mostly at Level 2 performance, in airplane mode FXV-15 meets Level 1 performance in AHS-33 but exhibits Level 2 performance according to the MIL standard for airplanes (MIL HDBK-1797, 1997).

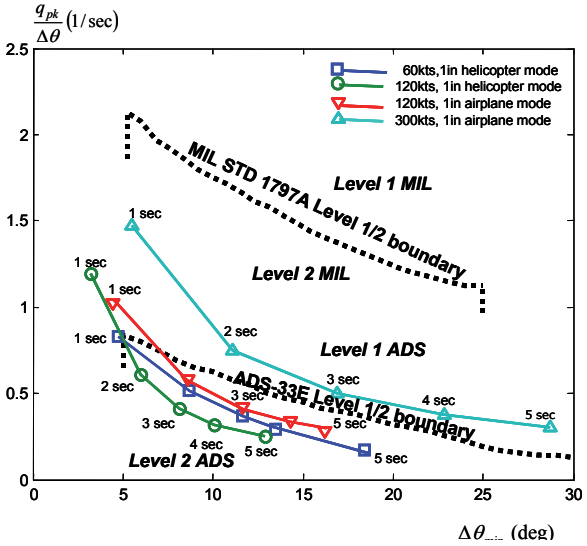


Fig. 8. Pitch quickness for the tiltrotor

3. Flying qualities metrics for agility designing

Linking the agility to flying qualities raised up a new question: is agility limited by pilot handling parameters or in other words what are the upper limits to agility set by flying qualities considerations? Flying qualities considerations do limit agility according to (Padfield, 1998). In this sense, in a series of flight and simulation trials research conducted at DERA (now Qinetics) the pilots were asked to fly maneuvers with increasing tempo until either performance or safety limit was reached. The results showed that in all cases the safety limit came first, thus the agility was constraint by safety.

3.1 Agility factor

A new metric was therefore introduced as a measure of performance margin (Padfield & Hodkinson, 1993), the so-called agility factor A_f, defined as the ratio of used to usable performance. For the simple case of the pull-up maneuver this metric can be easily calculated as the ratio of ideal task time T_i to actual task time T_a.

$$A_f \overset{def}{=} \frac{T_i}{T_a} = \frac{\omega_m \Delta t}{\omega_m \Delta t - \ln(0.1)} \tag{2}$$

where $T_i = \Delta t$ is the control pulse duration (1 to 5 sec), Ta is the time to reduce the pitch angle to 10% of the peak value achieved and ωm is the fundamental first-order break frequency or pitch damping which for this simple case represents the maximum achievable value of quickness. Fig. 9 illustrates the variation of Af with $\omega_m \Delta t$ -thus the quickness. The values considered for ω_m were: ωm=1.81 rad/s in hover helicopter mode, ω_m=2.6 rad/s at 60kts helicopter mode, ω_m=3.6 rad/s at 120kts 60deg conversion mode. Fig. 9 underlines an important aspect of the link between handling and agility: the higher the quickness, the higher the agility but when this agility is connected to Fig. 8 one may see that at the highest agility poor Level 2 ratings are awarded, i.e. the performance degrades rather than improves. This shows that actually, in practice, the closer the pilot flies to the performance boundary the more difficult it becomes to control the maneuver and thus the higher the agility the worse the HQs. In conclusion, handling qualities considerations do limit the agility.

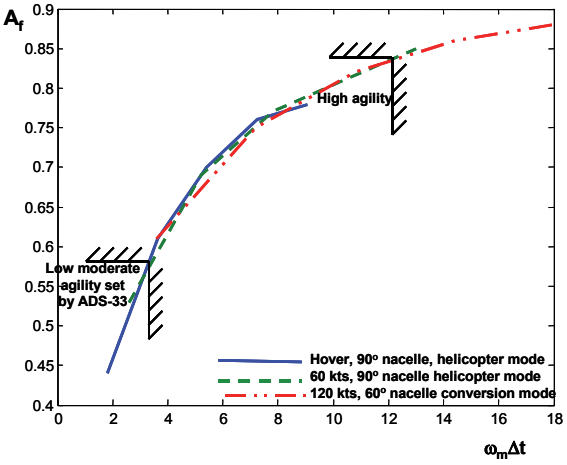


Fig. 9. Agility factor as a function of quickness

3.2 Control anticipation parameter

The discussion on the experimental metrics suggests that on the one side the best metric for pitch motion agility is the peak pitch acceleration and on the other side the best metric for determining the aircraft flight path bending capability is the peak load factor. In order to capture both the transients of the maneuver and the precision achieved in flight path control, MIL standard on fixed-wing aircraft (MIL-HDBK-1797, 1997) introduced as metric a combination between these two metrics, the so-called 'control anticipation parameter CAP'. CAP is defined as the ratio of the initial pitch acceleration to the steady state load factor (effectively pitch rate) after a step-type control input:

$$CAP \overset{def}{=} \frac{\dot{q}(0)}{n_z^{qs}} \tag{3}$$

MIL standard defines CAP boundaries for fixed-wing aircraft. Fig. 10 presents the agility of the FXV-15 CAP as a function of speed (60 kts, 120 kts and 200 kts) in the MIL boundaries. Looking at this figure one can see the tiltrotor meets Level 1 MIL performance and some degradation to Level 2 is seen when flying at high speeds in airplane mode.

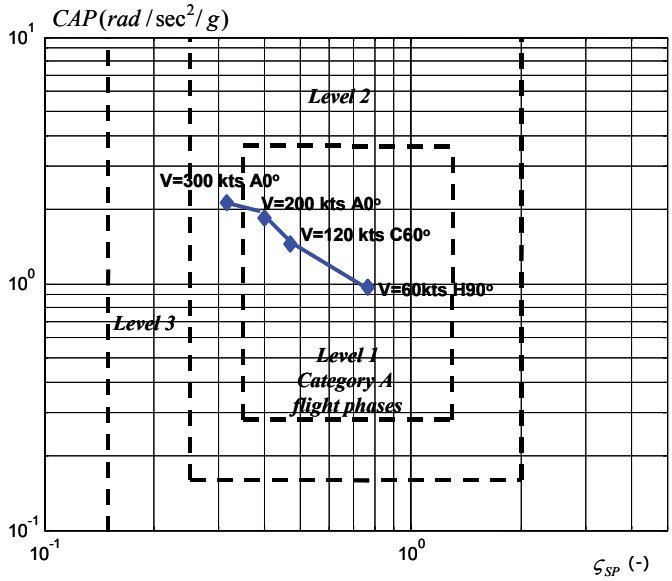


Fig. 10. CAP boundaries for the tiltrotor

3.3 Rate pitch quickness

For helicopters a similar metric to CAP was introduced by (Padfield & Hodkinson, 1993). This metric was called 'rate pitch quickness' and was defined as the ratio of pitch acceleration to the pitch angle change:

$$\dot{Q}_\theta \overset{def}{=} \frac{\dot{q}_{pk}}{\Delta\theta}\left(\sec^{-2}\right) \tag{4}$$

and can be used to determine upper limits to agility based on maneuver acceleration. Fig. 11 plotted the rate quickness in the normalized form as a function of acceleration time constant $\omega_m t_{pk}$ (where t_{pk} is the time to peak acceleration).

One can see that as the rate quickness increases the time to peak that rate is decreasing, so the agility is increasing. However, (Padfield & Hodkinson, 1993) commented that simply increasing the agility in terms of acceleration rates would lead to over-responsiveness and thus decreasing in operational capability since an over-responsive vehicle would not be controllable. In this sense, also CAP was quoted as an example of a criterion defining over-responsiveness. Unfortunately, there were no boundaries defined in this chart, although it was mentioned that intuitively there are likely to be upper and lower bounds for this metric. '*Hard and fast may be as unacceptable as soft and slow, both leading to low agility factors*' (Padfield & Hodkinson, 1993).

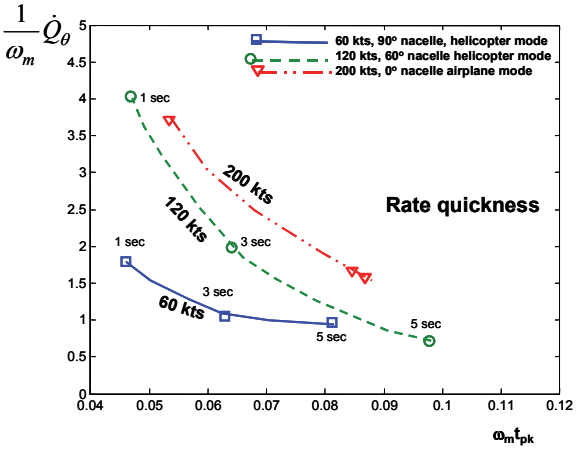


Fig. 11. Rate quickness as a function of time peak acceleration

4. A rational development of a multi-disciplinary approach to agility

Combining equation (3) for CAP with equation (4) for rate quickness it follows that:

$$\dot{Q}_\theta = CAP \cdot \frac{n_z}{\Delta\theta} \tag{5}$$

Equation (5) gives the idea that rate quickness and CAP can be related to each other through a new metric which will be presented in the next paragraph.

4.1 Agility quickness metric as a measurement of performance

As a potential successful metric for agility, (Pavel & Padfield, 2002) proposed a new metric for characterizing agility, the so-called 'agility quickness' defined as the ratio of peak quasi-steady normal acceleration n_{zpk}^{qs} in g units corresponding to a step change in flight path angle $\Delta\gamma$:

$$Q_\gamma \overset{def}{=} \frac{n_{zpk}^{qs}}{\Delta\gamma} \quad \left(\frac{g's}{\deg}\right) \tag{6}$$

Observe that the pitch angle from (5) was substituted by the flight path angle, this has been done because actually during vertical axis maneuvering agility is more related to how quickly the flight path can be changed, the pilot being in reality more interested in the flight path angle change than in the pitch change. Furthermore, (Pavel & Padfield, 2003) proposed a Level 1/2 performance boundary for agility quickness by flying yo-yo maneuvers in the full motion simulator at the University of Liverpool the UH-60A model. Fig. 12 presents the example of tiltrotor on the agility quickness charts as determined in (Pavel & Padfield, 2003).

One can see that the tiltrotor is mostly at Level 2 performance in helicopter and airplane modes. (Pavel & Padfield, 2003) derived a relation between CAP and Q_γ and (Padfield & Meyer, 2003; Cameron & Padfield, 2010) connected CAP to other flying qualities parameters. One of the reasons the attitude quickness criterion has gained large acceptance was due to its physical interpretation (in the limiting case gives the time constant of the aircraft as a function

of the time constant of the maneuver). It can be demonstrated that agility quickness has also a physical interpretation, in the limiting case for small-amplitude maneuvers giving the heave damping, for large amplitudes giving the attitude quickness (Pavel & Padfield, 2002).

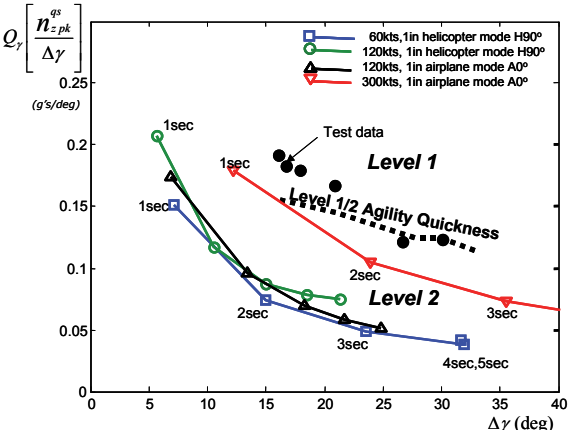


Fig. 12. Tiltrotor on Agility quickness chart

4.2 Vibratory quickness metric as a measurement of vibratory activity

The advantage of the agility quickness metric as above defined is that it could be linked to a complementary vibratory for structural alleviation. In this way, the designer is able to optimize in parallel both the performance and the vibratory loads in maneuvering flight. This novel approach may reform the methods presently used to accomplish agility of a new design as it is well known that, in practice, performance is itself often compromised by the need to control and minimize vibration.

(Pavel and Padfield, 2002) define an parallel metric, so-called 'vibratory load quickness' quantifying the buildup of loads in the rotor during maneuvering flight:

$$Q_l \overset{def}{=} \frac{F_{pk}^{vib}}{W\Delta\gamma}\left(\frac{1}{\deg}\right);\quad Q_l \overset{def}{=} \frac{M_{pk}^{vib}}{\Delta\gamma}\left(\frac{lbf\cdot ft}{\deg}\right) \tag{7}$$

where F_{pk}^{vib}, M_{pk}^{vib} represent the peak amplitudes in the critical vibratory components for respectively hub shears and hub moments corresponding to a change $\Delta\gamma$ in flight path angle. The peak load amplitude can be calculated by using the FFT and time representations of the hub shears ($F_{x\,hub}$, $F_{y\,hub}$ or $F_{z\,hub}$) and/or moments ($M_{x\,hub}$, $M_{y\,hub}$, $M_{z\,hub}$) during a maneuver flown and determining the critical loads (i.e. the loads achieving the highest peaks) during the maneuver.

For example, for the pull-up maneuver flown with the tiltrotor it was found that when flying in helicopter mode at 60 and 120 kts the critical loads developed were the 3/rev vibratory component of the hub vertical shear, the 1/rev and 2/rev components of the blade inplane moment and the 1/rev component of the blade flapping moment. When flying in the airplane mode at 120 and 300 kts, the critical loads measured by the FXV-15 were the 2/rev and 3/rev components of the vertical shear, the 1/rev and 2/rev components of the blade inplane moment.

Fig. 13 presents the equivalent vibratory quickness charts for the critical 3/rev component of the hub vertical shear in helicopter mode and 2/rev and 3/rev components of hub vertical shear in airplane mode when flying respectively at 60 and 120 kts and 120 and 300 kts, giving an 1 in input in longitudinal cyclic and varying the pulse duration (1 to 5 seconds). Each of these vibratory chart can be associated with an equivalent agility quickness chart as plotted in Fig. 12.

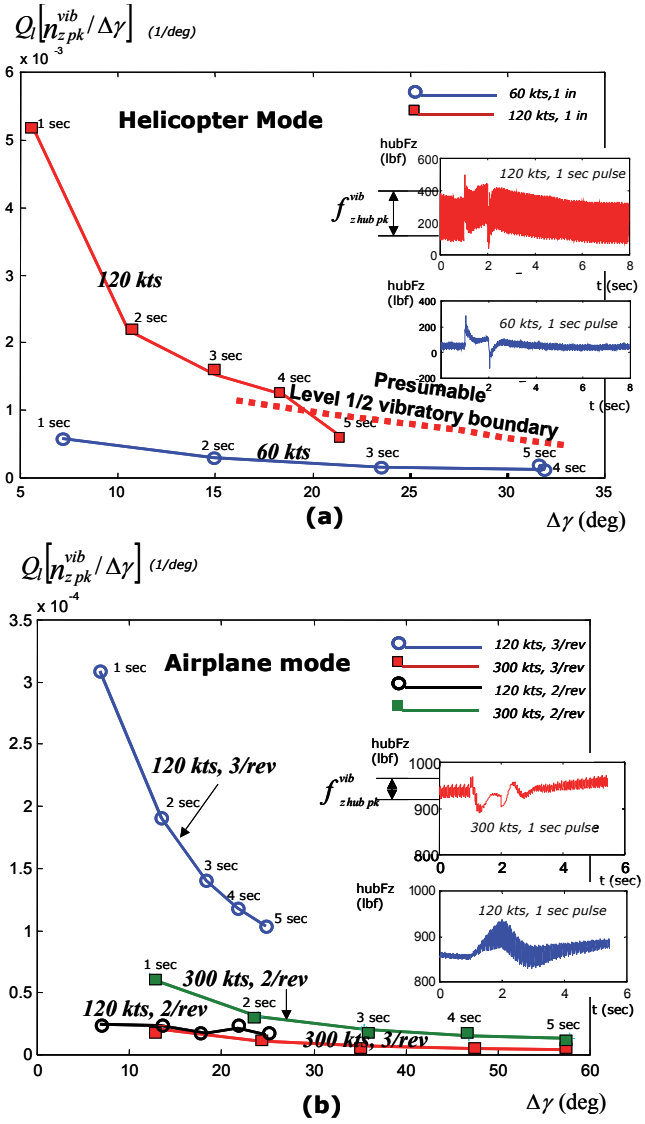


Fig. 13. Vibratory quickness envelopes for the critical components of the hub vertical shear during a pull-up maneuver

For helicopter mode it was plotted a presumable vibratory quickness boundary as derived in (Pavel & Padfield, 2003) when flying piloted yo-yo's in the simulator with the UH-60A helicopter. It was there showed that actually, increasing the flight path change enables the pilot to pull more g's of course but also increases the vibratory activity in the rotor. The presumable vibratory boundary would mean then that the structural designer would aim for a

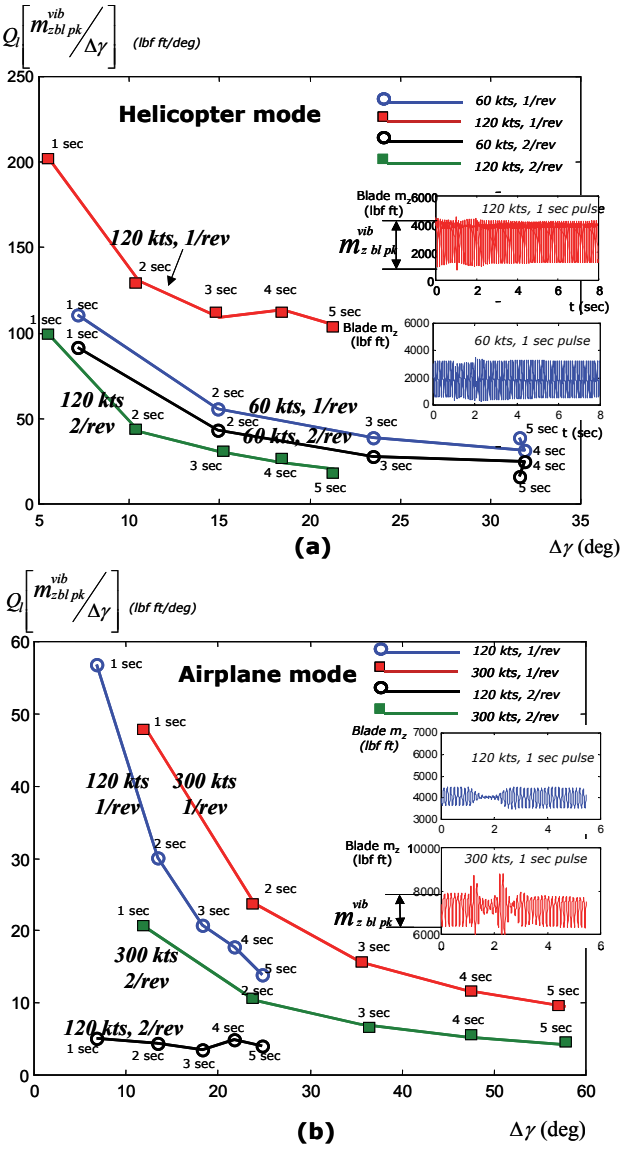


Fig. 14. Vibratory quickness envelopes for the critical components of the blade inplane moment during a pull-up maneuver

boundary that slope in a similar direction to the agility quickness boundary as plotted in Fig. 13. Looking at this figure one may see that as the pulse duration is increasing the vibratory quickness is decreasing. This is because the vibratory activity in the hub reaches its absolute peak rather quickly, depending mainly on the initial velocity, the input amplitude (which is a measure of the level of aggressiveness in executing the maneuver) and not on the pulse duration. However, for very aggressive maneuvering, (Pavel & Padfield, 2002) showed that it might appear the situation in which, increasing the flight path change enabled the pilot to pull more g's (so, increased performance) but also increased the vibratory activity in the rotor. The goal of the structural designer would be then to alleviate these high peak loads to lower levels and reduce the sensitivity of the vibratory loads to flight path angle.
Fig. 14 presents other critical load for the tiltrotor, namely the blade inplane moment. Both, for the helicopter and airplane mode, the critical components measured during the simulation of the pull-up maneuvers with the FXV-15 were corresponding to the 1/rev and 2/rev vibratory components.
Looking at Fig. 14 one may see again that the vibratory quickness parameter Q_I varies approximatively inversely with the flight path change. This means that the vibratory activity in the blade in the inplane direction reaches its absolute peak rather quickly, depending on the aggressiveness of the pulse (pulse amplitude) and not on pulse duration.

5. Conclusion

This chapter has presented a rational development of key metrics and criteria used to design for airframe agility. Concentrating on the agility in the pitch axis (vertical-plane maneuvers) and taking as case study the unique example of a tiltrotor aircraft, the chapter demonstrated how, starting from the more traditional way of quantifying the agility, the designer can develop new agility metrics that do a better job of capturing the aircraft transient motion characteristics. This chapter discussed on the many correspondences that exist between the study of agility and the study of flying qualities, emphasizing the fact that flying qualities do limit agility. In this sense, providing the pilot with a high level of maneuverability, without a high level of controllability, will reduce agility. However, especially for the tiltrotor, higher agility cannot be achieved without increasing the vibratory loads on the rotor, which means also an increasing in pilot workload. The chapter proposed therefore a unique approach by presenting a first set of complementary metrics capable of being applied to both agility and structural load analysis.
Subsequent phases of this study will include the expansion of this new approach and studying it in axial, turning (horizontal-plane maneuvers) and roll (torsion) axes. It is hoped that in this way a more unified set of design criteria will be developed enhancing multi-disciplinary design optimization.

6. Acknowledgment

The subject of this research was initiated as collaboration between The University of Liverpool and Delft University of Technology. Sincere thanks for all support given by the University of Liverpool and Prof. Padfield for using FLIGHTLAB FXV-15 model.

7. References

ADS-33 (2000) *Aeronautical Design Standard-33E-PRF, Performance Specification, Handling Qualities Requirements for Military Rotorcraft*, US Army AMCOM, Redstone, Alabama

Cameron, Neil & Padfield, Gareth D. (2010). Tilt Rotor Pitch/Flight-Path Handling Qualities, *J. Of the American Helicopter Society*, 55(4), pp. 042008-1 to 042008-13

Celi, R., (1991). Optimum Aeroelastic Design of Helicopter Rotors for Longitudinal Handling Qualities Improvement, *J. of Aircraft*, 28(1), pp. 49-57.

Celi, R, (1999). Recent Applications of Design Optimization to Rotorcraft – A Survey, *J. of Aircraft*, 36 (1), pp.

Celi, R, (2000). Optimization Based Inverse Simulation of a Helicopter Slalom Maneuver, *J. of Guidance, Control and Dynamics*, 23(2), pp. 289-297

Cooper, G.E., and Harper, R.P. Jr., (1969). The use of Pilot Ratings in the Evaluation of Aircraft Handling Qualities, NASA TM D-5133

Datta, A. & Chopra, I., (2002), Validation and Understanding of UH-60A Vibration Loads in Steady Flight, *58th American Helicopter Society Conference*, 11-14 June, Montreal

Fusato, D. & Celi, R., (2001). Formulation of a Design-Oriented Helicopter Flight Dynamic Simulation Program, *57th American Helicopter Society Forum*, Alexandria, VA

Fusato, D. & Celi, R., (2002). Design Sensitivity Analysis for ADS-33 Quickness Criteria and Maneuver Loads, 58th American Helicopter Society Forum, 11-14 June, Montreal

Fusato, D. & Celi, R., (2002). Multidisciplinary Design Optimization for Aeromechanics and Handling Qualities", *28th European Rotorcraft Forum*, 17-20 September, Bristol, UK

Ganduli, R., (2004). Survey of Recent Developments in Rotorcraft Design Optimization", *J. of Aircraft*, 41(3), pp. 493-510

Hansford, R.E. and Vorwald, J., (1996). Dynamic Workshop on Rotor Vibratory Loads. *52nd Forum of the American Helicopter Society*, June 4-6, Washington D.C.

Iseler, Laura; De Maio, Joe & Rutkowski, Michael (2002). *Analysis of US Civil Rotorcraft Accidents by Cost & Injury (1990-1996)*, Army/NASA Rotorcraft Division Ames Research Center, USA, NASA-TM-2002-209615

Key, D.L., (1982). The status of Military Handling Qualities Criteria, AGARD CP-333

Kolwey, H.G., (1996). The New ADS-33 Process: Cautions for Implementation, Technical Note, *J. of the American Helicopter Society*, 41(1), pp. 3-6.

Kufeld, R.M. & Bousman, W.G., (1995). High Load Conditions measured on a UH-60A in Manoeuvring Flight, *51st Forum of the American Helicopter Society*, May 9-11, Texas

McKay, K. (1994). Operational Agility: An Overview of AGARD Flight Mechanics Working Group 19, In: *Technologies for Highly Manoervrable Aircraft*, AGARD-CP-548

Manimala, Binoy; Naddei, Mauro &Rollet, Philippe (2003). Load Alleviation in Tiltrotor Aircraft through Active Control; Modelling and Control Concepts, *59th American Helicopter Society Conference*, Phoenix, Arizona, May 6-8, 2003

Mazza, C.J. (1990). Agility: A Rational Development of Fundamental metrics and their relationship to Flying Qualities. *Proceedings of Flight Mechanics Panel Symposium*, Quebec City, Canada, October 15–18, 1990, Published as AGARD-CP-508

MIL-H-8501A (1962). *Military Specification, General Requirements for Helicopter Flying and Ground Handling Qualities MIL-H-8501A*, Sept. 1961, through Amendment 1, April 1962 (superseding MIL-H-8501, Nov. 1952).

MIL-HDBK-1797 (1997). *Flying Qualities of Piloted Aircraft, MIL-HDBK-1797*, US Department of Defense, USA

Mitchell, D.G., et. al., (2004). Evolution, Revolution and Challenges of Handling Qualities, *J. of Guidance, Control and Dynamics*, 27(1), pp. 12-28

Murphy, P.; Bailey, M. & Ostroff (1991). A. Candidate Control Design Metrics for an Agile Fighter, NASA-TM-4238

Padfield, Gareth D. (1998). The Making of Helicopter Flying Qualities; A Requirements Perspective. *The Aeronautical Journal of Royal Aeronautical Society*, Vol. 102, No. 1018, pp. 409-437

Padfield, G.D. (2007). *Helicopter Flight Dynamics*, Blackwell Science LTD

Padfield, G.D. (2008). Capturing requirements for tiltrotor handling qualities – case studies in virtual engineering, *The Aeronautical Journal*, Vol. 112(1134), Aug. 2008, pp. 433-448

Padfield, Gareth, D.& Hodkinson, John. The Influence of Flying Qualities on Operational Agility", *AGARD Flight Mechanics Panel Symposium Technologies for Highly Manoeuvrable Aircraft*, Annapolis, Maryland, Oct. 1993, Published as AGARD-CP-548

Pavel, Marilena, D. & Padfield, Gareth, D. (2002). Defining Consistent ADS-33-Metrics for Agility Enhancement and Structural Loads Alleviation, *58th American Helicopter Society Conference*, Montreal, Canada, 11-14 June 2002

Pavel, Marilena D. & Padfield, Gareth. D. (2003). Progress in the Development of Complementary Handling and Loading Metrics for ADS-33 Maneuvers", *59th American Helicopter Society Conference*, Phoenix, Arizona, 6-8 May 2003

Padfield, Gareth. D. & Meyer, Michael (2003) Progress in Civil Tilt-Rotor Handling Qualities", *29th European Rotorcraft Forum*, Friedrichshafen, Germany, 15-17 October 2003.

Sahasrabudhe, V. & Celi, R., (1997). Efficient Treatment of Moderate Amplitude Constraints for Helicopter Handling Qualities Design Optimization, *J. of Aircraft*, 34(6), pp. 730-739

Tischler, M.B., et.al., (1997). CONDUIT- A New Multidisciplinary Integration Environment for Flight Control Development, NASA TM-112203

Tischler, M.B., Williams, J.N., Ham & J.A., (1996). Comments on The New ADS-33 Process: Cautions for Implementation, Technical Note, *J. of the American Helicopter Society*, 41, 194-195

'Veilig Vliegen' magazine (2003). *Hoe de zaken liepen in 2002' ('How the things evolved in 2002') , May 2003,* pp. 133-138

WAAS (2002). *World Aircraft Accident Summary – 1990-2001 Index-Cap 479*, Civil Aviation Authority, Airclaims Limited

9

Concurrent Subspace Optimization for Aircraft System Design

Ke-shi Zhang
Northwestern Polytechnical University
China

1. Introduction

Concurrent Subspace Optimization (CSSO) is one of the main decomposition approaches in Multidisciplinary Design Optimization (MDO). It supports a collaborative and distributed multidisciplinary design optimization environment among different disciplinary groups. Sobieski first proposed the subspace optimization method (Sobieszczanski-Sobieski, 1988), and Sobieski's blueprint was further developed by Bloebaum and subsequently named the concurrent subspace optimization method (Bolebaum, 1991). Renaud developed a second-order variant of the Global Sensitivity Equation (GSE) method and an alternative potential coordination procedure for the CSSO method (Renaud & Gabriele, 1993a, 1993b, 1994). Sellar proposed to replace GSE with the neutral-network based response surface method (Sellar et al., 1996).

The CSSO method allows a complex couple system to be decomposed into smaller, temporarily decoupled subsystems, each corresponding to different disciplines (subspaces). Each subspace optimization minimizes the system objective function subject to its own constraints as well as constraints contributed from the other subspaces. Each subspace optimization use its own high-fidelity analysis tools as well as given surrogate models or low-fidelity analysis tool provided by the other subspaces for analysis. Subsequently, the subspace optimizations can be performed concurrently. The system-level coordination optimization will be implemented completely based on approximation analysis tools. The subspace optimizations and the coordination optimization will be alternatively performed until results are finally decided by the coordination optimization. Therefore, the CSSO method is particularly suited to applications in a design organization where tasks are distributed among different design groups.

The CSSO method was developed initially for a single objective MDO problem. However, most MDO problems are essentially multi-objective. In recent years more work (Aute & Azarm, 2006; Huang & Bloebaum, 2004; McAllister et al., 2000; McAllister et al., 2004; Orr & Hajela, 2005; Parashar & Bloebaum, 2006; Tappeta & Renaud, 1997; Zhang et al., 2008) has focused on extending existing MDO method to handle such multi-objective MDO problems, by means of integrating a multi-objective optimization method within the MDO framework. This kind of method can be called a multi-objective MDO method.

It is an effective way to integrate multi-objective optimization method within the CSSO framework to develop the multi-objective MDO method. CSSO was extended to solve multi-objective MDO problems, including the Multi-objective Pareto CSSO (MOPCSSO) method,

the Multi-objective Range CSSO (MORCSSO) method, the Multi-objective Target CSSO (MOTCSSO) method, the Multi-objective Genetic Algorithm CSSO (MOGACSSO) method and Adaptive Weighted Sum based CSSO (AWSCSSO). In MOPCSSO the Constraint method is integrated within CSSO framework (Huang & Bloebaum, 2007). In MORCSSO and MOTCSSO the concept of designer preference is introduced (Huang & Bloebaum, 2004). In MOGACSSO the Genetic Algorithm is combined with CSSO and in the hope of improving the computational efficiency (Parashar & Bloebaum, 2006). In AWSCSSO the Adaptive Weighted Sum method is introduced into CSSO (Zhang et al., 2008).

2. General description of MDO problem for aircraft system design

Aircraft is a complex engineering system in which multiple disciplines (such as aerodynamics, structure, thrust, noise, electronics, cost, et al.) are included. Actually different disciplines are not independent of each other. For example, deformation of a wing structure affects the aerodynamic lift distribution on the wing and in turn a new deformation is caused, which is the well-known aeroelastic problem. The designers of each discipline cannot work without consideration of other disciplines, which makes the aircraft system design become complicated. Especially in the aircraft preliminary design, the specialists of different disciplines have to often work together and discuss with each other to decide many design variables so that higher performance can be achieved. Aircraft system design is a typical multidisciplinary design problem.

In recent years, industry has paid more attention to improving efficiency in the design of complex systems, such as aircraft. MDO has emerged as an engineering discipline that focuses on the development of new design and optimization strategies for the complex systems. MDO researchers strive to reduce the time and cost associated with the coupling interaction among several disciplines. "Decomposition approaches provide many advantages for the solution of complex MDO problems, as they enable a partitioning of a large coupled problem into smaller, more manageable sub-problems. The resulting computational benefits, besides the obvious one associated with the solution of smaller problems, include creating a potential distributed processing environment. The primary benefit, however, pertains to the savings in personal hours, because groups are no longer required to wait around for other groups in the process to complete their design tasks."(Huang & Bloebaum, 2007)

Mathematically for minimization problems the general form for MDO can be represented as follows:

$$\begin{aligned} &\text{Min} \quad \boldsymbol{F}\left(\boldsymbol{X},\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_{\mathrm{N}}\right) \\ &\text{s.t.} \quad \boldsymbol{G}_i\left(\boldsymbol{X},\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_{\mathrm{N}}\right) \le \boldsymbol{0} \\ &\qquad \boldsymbol{Y}_i = f_i\left(\boldsymbol{X},\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_{i-1},\boldsymbol{Y}_{i+1},\cdots,\boldsymbol{Y}_{\mathrm{N}}\right) \\ &\qquad i = 1,2,\cdots,\mathrm{N} \\ &\qquad \mathrm{X}=\left\{x_1,x_2,\cdots,x_{\mathrm{NV}}\right\} \end{aligned} \tag{1}$$

Where $\boldsymbol{F}$ is the objective function vector that is composed of one or more objectives, $\boldsymbol{G}_i$ is the constraint vector provided by subsystem i, $\boldsymbol{Y}_i$ is the coupling vector of subsystem i, and $\boldsymbol{X}$ is the design vector. The objective function and the constraints can be expressed by the

function of $\boldsymbol{X}$ and $\boldsymbol{Y}_i$. In this chapter the MDO problem expressed in Eq. (1) will be taken as the example to discuss the CSSO methods.

3. Different frameworks for concurrent subspace optimization

Global Sensitivity Equation based CSSO and Response Surface based CSSO will be discussed in this chapter as the typical CSSO methods.

3.1 Global Sensitivity Equation based CSSO (GSECSSO)

GSECSSO is a bi-level optimization method. As an example, the GSECSSO method for a problem with subsystem 1 and 2 (Eq. (1): one objective and two coupled subsystems) is stated in following paragraphs.

The mathematical models of subspace optimizations of GSECSSO can be written as

$$\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } F\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) & \text{Min } \hat{F}\left(\boldsymbol{X}_2\right) \\
\text{s.t. } C_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) \le C_1^0\left(s_1\left(1-r_1^1\right)+\left(1-s_1\right)t_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right) \le C_1^0\left(s_1\left(1-r_2^1\right)+\left(1-s_1\right)t_2^1\right) \\
\quad \hat{C}_2\left(\boldsymbol{X}_1\right) \le C_2^0\left(s_2\left(1-r_1^2\right)+\left(1-s_2\right)t_1^2\right) & \quad C_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right) \le C_2^0\left(s_2\left(1-r_2^2\right)+\left(1-s_2\right)t_2^2\right) \\
\quad \boldsymbol{Y}_1 = f_1\left(\boldsymbol{X}_1,\hat{\boldsymbol{Y}}_2\right) & \quad \boldsymbol{Y}_2 = f_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1\right)
\end{array} \tag{2}$$

Where $\boldsymbol{X}_i$, $\boldsymbol{Y}_i$ and C_i are design variable vector, state variable vector and cumulative constraint of subspace i, respectively. C_i^0 is the value of C_i at the starting point $\boldsymbol{X}^0$. The value with symbol '^' above is a linearly approximated one. r_k^p, t_k^p and s_p are responsibility, trade-off and switch coefficients, respectively. In the first iteration, the responsibility coefficients are initialized using the sensitivities of cumulative constraints with respect to design variables (Bolebaum, 1991) and the switch coefficients are set to one. In the following iterations the coordination coefficients are decided by the coordination optimizations.

By means of Kresselmeier-Steinhauser (KS) function (Kreisselmeier & Steinhauser, 1979), the cumulative constraint, C_i, which represents the constraints of subsystem i, is expressed as

$$C_i = g_{i\max} + \frac{1}{\rho}\ln\left(\sum_{k=1}^{\mathrm{m}}\exp\left(\rho\left(g_{ik}\left(\boldsymbol{X}\right)-g_{i\max}\right)\right)\right) \tag{3}$$

Where $\left[g_{i1},g_{i2},\cdots,g_{i\mathrm{m}}\right]=\boldsymbol{G}_i$ and ρ is a positive user-prescribed value.

Global Sensitivity Equation (GSE) (Sobieszczanski-Sobieski, 1990) is a method for computing sensitivity derivatives of state (output) variables with respect to independent (input) variables for complex, internally coupled systems, while avoiding the cost and inaccuracy of finite differencing performed on the entire system analysis. By using GSE expressed below, the global sensitivities are calculated, by

$$\begin{bmatrix} \boldsymbol{I} & -\dfrac{\partial \boldsymbol{Y}_1}{\partial \boldsymbol{Y}_2} \\ -\dfrac{\partial \boldsymbol{Y}_2}{\partial \boldsymbol{Y}_1} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \dfrac{d\boldsymbol{Y}_1}{d\boldsymbol{X}} \\ \dfrac{d\boldsymbol{Y}_2}{d\boldsymbol{X}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \boldsymbol{Y}_1}{\partial \boldsymbol{X}} \\ \dfrac{\partial \boldsymbol{Y}_2}{\partial \boldsymbol{X}} \end{bmatrix} \tag{4}$$

The responsibility coefficients divide the responsibility of satisfying constraints among all participating subspaces. r_k^p represents the responsibility allocated to the k-th subspace for satisfying the p-th cumulative constraint. In the first iteration, they are initialized with the sensitivities of cumulative constraints with respect to design variables (Bolebaum, 1991). The trade-off coefficients allow for the violation of a constraint in one subspace optimization in order to gain a large reduction of the objective function. The switch coefficients enable or disable the responsibility coefficients and the trade-off coefficients decided by whether the cumulative constraints are satisfied.

The mathematical model of coordination optimization of GSECSSO can be written as

$$\begin{aligned} \text{Min} \quad & F(\boldsymbol{r},\boldsymbol{t}) = F^* + \sum_{p=1}^{2}\sum_{k=1}^{2} \left.\frac{dF}{dr_k^p}\right|_{\boldsymbol{X}=\boldsymbol{X}^*} \Delta r_k^p + \sum_{p=1}^{2}\sum_{k=1}^{2} \left.\frac{dF}{dt_k^p}\right|_{\boldsymbol{X}=\boldsymbol{X}^*} \Delta t_k^p \\ \text{s.t.} \quad & \sum_{k=1}^{2} r_k^p = 1, \quad \sum_{k=1}^{2} t_k^p = 0 \\ & 0 \le r_k^p \le 1, \quad -1 \le t_k^p \le 1, \quad p,k = 1,2 \end{aligned} \tag{5}$$

Where F^* is the objective value at X^* (X^* is the value of design variable vector after subspace optimizations). $\left.\frac{dF}{dr_k^p}\right|_{\boldsymbol{X}=\boldsymbol{X}^*}$ and $\left.\frac{dF}{dt_k^p}\right|_{\boldsymbol{X}=\boldsymbol{X}^*}$ are the optimum sensitivity of F with respect to responsibility and trade-off coefficients.

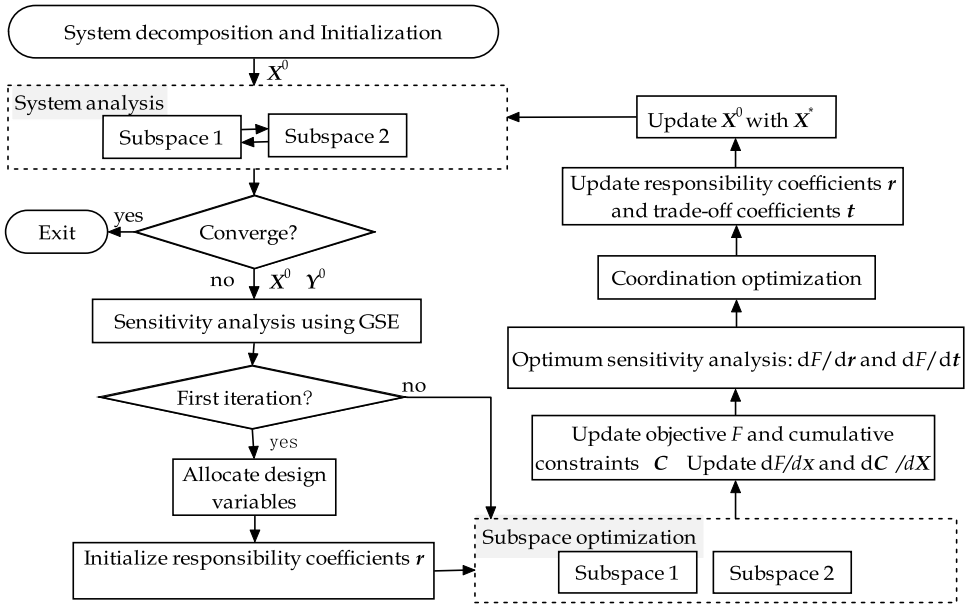


Fig. 1. Flowchart of GSECSSO method

The flowchart of GSECSSO is shown in Fig. 1. The original problem is decomposed and the mathematical models of subspace optimizations and coordination optimization are established as Eq. (2) and Eq. (5). Based on the initial design $\boldsymbol{X}^0$, the system analysis is performed and the GSE is used to find the system derivatives. At the first iteration, the design variables are allocated to the appropriate subspaces, as well as the responsibility, trade-off and switch coefficients, $\boldsymbol{r}$, $\boldsymbol{t}$, $\boldsymbol{s}$, are initialized. Subsequently, each subspace optimization minimizes the system objective function subject to its own constraints as well as constraints contributed from the other subspaces. The subspace optimizations are performed concurrently with respect to a disjoint subset of design variables. The updated design vector is the simple combination of local optimal design sub-vectors. After that the global sensitivities and optimum sensitivities are updated. Then the system-level coordination optimization is implemented to optimize coordination parameters.

In GSECSSO, the subspace optimizations are implemented concurrently with respect to a disjoint subset of design variables, which substantially reduces the complexity of the optimization problem within each disciplinary group. The updated design vector is the simple combination of local optimal design sub-vectors. This provides designers with a significant potential benefit in terms of computational effort. However, sometimes the convergence of GSECSSO is oscillatory and premature. It is due to that the trade-off coefficients and the KS parameter ρ with inappropriate value may lead to bad convergence (Huang & Bloebaum, 2004) as main reasons.

3.2 A variant of GSECSSO

The variant of GSECSSO (Huang & Bloebaum, 2004) is much more efficient than the original GSECSSO. It is adopted in multiple multi-objective CSSO methods (Huang & Bloebaum, 2004; Parashar & Bloebaum, 2006; Zhang et al., 2008).

In the variant of GSECSSO, several modifications are made: (1) the trade-off coefficients are abandoned since all trade-offs will occur directly with minimization of the objective functions within each subspace; (2) The KS parameter, ρ, is set to be increased from a small value; (3) an infeasibility minimization is appended to move the point into the feasible region; (4) the coordination optimization are abandoned and the responsibility coefficients are directly calculated via their initialization strategy.

Taking the MDO problem in Eq. (1) (one objective, two coupled subsystems, i.e. N=2) as an example, the mathematical models of subspace optimizations for the modified CSSO can be written as

$$
\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } F\left(\boldsymbol{X}_1, \boldsymbol{Y}_1, \hat{\boldsymbol{Y}}_2\right) & \text{Min } \hat{F}\left(\boldsymbol{X}_2\right) \\
\text{s.t. } C_1\left(\boldsymbol{X}_1, \boldsymbol{Y}_1, \hat{\boldsymbol{Y}}_2\right) \le C_1^0\left(1 - r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right) \le C_1^0\left(1 - r_2^1\right) \\
\quad \hat{C}_2\left(\boldsymbol{X}_1\right) \le C_2^0\left(1 - r_1^2\right) & \quad C_2\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \boldsymbol{Y}_2\right) \le C_2^0\left(1 - r_2^2\right) \\
\quad \boldsymbol{Y}_1 = f_1\left(\boldsymbol{X}_1, \hat{\boldsymbol{Y}}_2\right) & \quad \boldsymbol{Y}_2 = f_1\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1\right)
\end{array} \tag{6}
$$

Where $\boldsymbol{X}_i$, $\boldsymbol{Y}_i$ and C_i are design variable vector, state variable vector and cumulative constraint of subspace i, respectively. C_i^0 is the value of C_i at the starting point $\boldsymbol{X}^0$. The value with symbol '^' above is a linearly approximated one. r_k^p is responsibility coefficients.

the flowchart of modified CSSO method is shown in Fig. 2. In the first stage, based on the initial design variables the system analysis is implemented. If the initial scenario does not satisfy the constraints, a minimization will be required to reduce initial infeasibility of the constraints as much as possible. In the subsequent stage, all the sensitivities are computed by using GSE. Based on the sensitivity information, the impacts of design variables upon each subspace can be analyzed and the design variables will be allocated to the subspace of the greatest impact. In the third stage, the subspace optimizations are performed concurrently. There is no system optimization in this method. The updated design vector is the simple combination of the local optimal design sub-vectors.

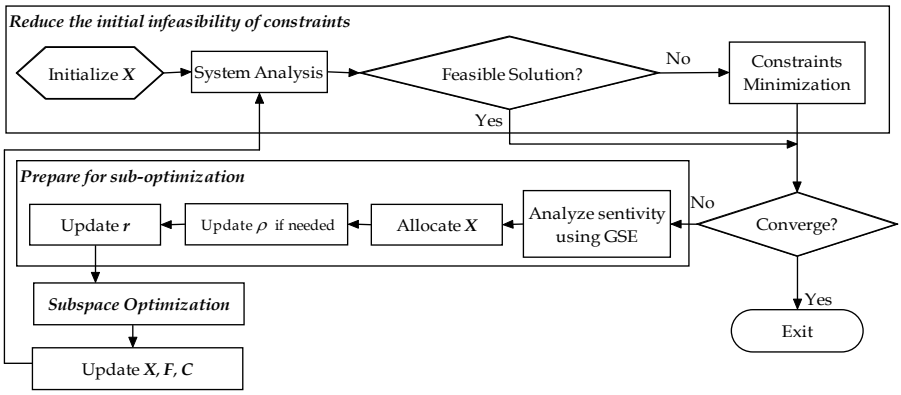


Fig. 2. Flowchart of the modified CSSO method

The variant of GSECSSO behaves better than GSECSSO on convergence performance. Nevertheless, how to set the start value and increase step of the KS parameter is still hard to say. These two values will affect convergence to some extent and need to be further investigated. The constraint minimization will bring extra computation cost. Furthermore, the linear approximation may sometimes cause oscillatory and premature convergence.

3.3 Response Surface based CSSO (RSCSSO)

The RSCSSO method performs optimization in bi-level framework. Taking the MDO problem in Eq. (1) (one objective, two coupled subsystems, i.e. N=2) as an example, the mathematical models of subspace optimizations for RSCSSO can be written as

$$\begin{array}{l|l} \text{Sub-optimization 1} & \text{Sub-optimization 2} \\ \text{Min } F\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) & \text{Min } F\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right) \\ \text{s.t. } \boldsymbol{G}_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right)\le \boldsymbol{0} & \text{s.t. } \boldsymbol{G}_1\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)\le \boldsymbol{0} \\ \boldsymbol{G}_2\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right)\le \boldsymbol{0} & \boldsymbol{G}_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)\le \boldsymbol{0} \\ \boldsymbol{Y}_1 = f_1\left(\boldsymbol{X}_1,\hat{\boldsymbol{Y}}_2\right) & \boldsymbol{Y}_2 = f_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1\right) \end{array} \tag{7}$$

Where $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ are design variable vector and state variable vector of subspace i, respectively. In subspace i, $\boldsymbol{Y}_i$ is calculated using the high-fidelity analysis tool while the

other state variables with a symbol '^' above are approximated by a quadratic response surface. Actually, other surrogate modeling techniques, such as kriging model or radial basis functions, can be used for approximation as well.
The mathematical model of the coordination optimization of RSCSSO can be written as

$$\begin{aligned} &\text{Min} \quad F\left(\boldsymbol{X},\hat{\boldsymbol{Y}}_1,\hat{\boldsymbol{Y}}_2\right) \\ &\text{s.t.} \quad \boldsymbol{G}_1\left(\boldsymbol{X},\hat{\boldsymbol{Y}}_1,\hat{\boldsymbol{Y}}_2\right) \le \boldsymbol{0} \\ &\qquad \boldsymbol{G}_2\left(\boldsymbol{X},\hat{\boldsymbol{Y}}_1,\hat{\boldsymbol{Y}}_2\right) \le \boldsymbol{0} \end{aligned} \tag{8}$$

In the system-level optimization, all the state variables are approximately calculated and the design variables of all subspaces are optimized.
The flowchart of RSCSSO is shown in Fig. 3. Firstly, the original problem is decomposed into subspace optimizations in Eq. (7) and coordination optimization in Eq. (8). Then couples of sample points of the design variable vector are generated via Design of Experiment (DOE) method, such as the orthogonal experiment design. The system analysis will be performed based on these data and the results will be stored in the database.

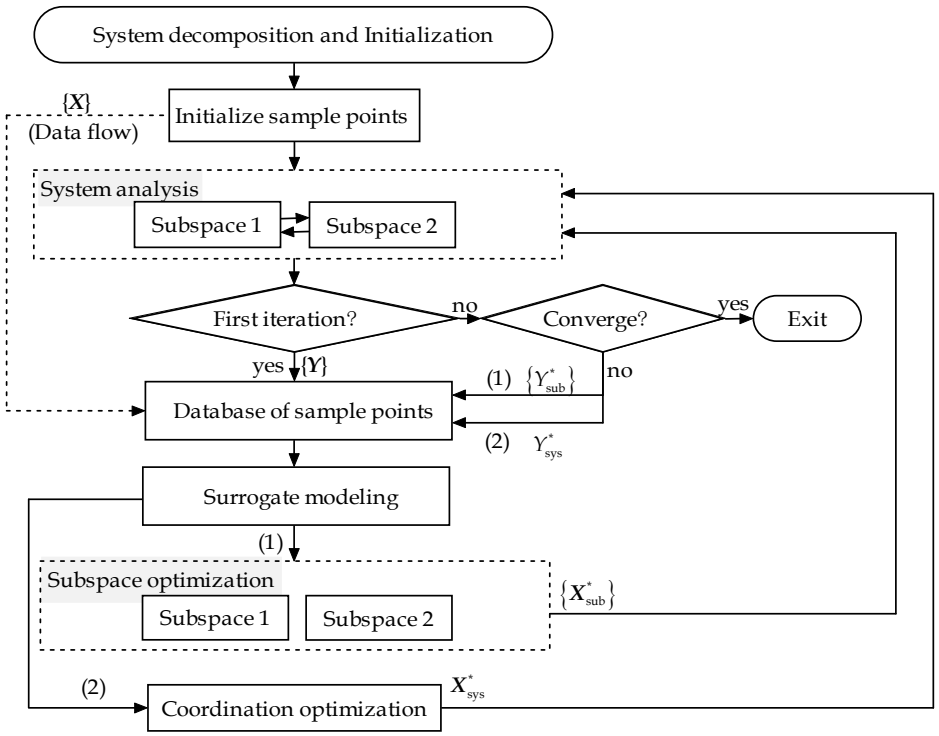


Fig. 3. Flowchart of RSCSSO method

Subsequently these data will be used to generate the response surface models for the state variables. After that the subspace optimizations are performed simultaneously, the results of

which will be analyzed and augmented into the database for updating the response surface models, as flow (1) in Fig. 3. Then all of the design variables will be optimized in the coordination optimization and the result will be used to update response surface models as well, as flow (2) in Fig. 3. The subspace optimizations and the coordination optimization are alternately performed until stable convergence is achieved.

In the RSCSSO method, the subspace optimizations are to generate sufficient design information for approximation. After the concurrent subspace optimizations, a global approximation problem is formulated about the current design vector using information stored in the design database. It is the coordination procedure of the global approximation that drives constraint satisfaction and the overall system optimization. A fast and robust convergence appears in the RSCSSO method. However, with an increase in design variables the sample points needed for creating response surface models will be greatly increased. Furthermore, augmenting optimal points is not so reasonable for improving response surface models. As the final result is completely decided by the coordination optimization, the subspace optimizations have little impact on the results and can be abandoned.

3.4 Examples

3.4.1 Example 1: An analytical example with coupling relationship

The mathematical model for an analytical MDO example considering two coupling disciplines is of the form

$$\begin{aligned} \text{Min}\ & f = x_2^2 + x_3 + y_1 + e^{-y_2} \\ \text{s.t.}\ & g_1 = 8 - y_1 \le 0 \\ & g_2 = y_2 - 10 \le 0 \\ & y_1 = x_1^2 + x_2 + x_3 - 0.2y_2 \\ & y_2 = \sqrt{y_1} + x_1 + x_3 \\ & x_1 \in [-10,10],\ x_2, x_3 \in [0,10] \end{aligned} \tag{9}$$

Where y_1 and y_2 are state variables belonging to subspace 1 and subspace 2, respectively; g_1 and g_2 are constrains provided by discipline 1 and discipline 2, respectively. The coupling relationship between discipline 1 and discipline 2 is depicted in Fig. 4.

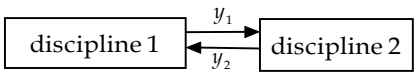


Fig. 4. Coupling relationship between discipline 1 and discipline 2 for a numerical example

The comparison of the results obtained by different methods is listed in Table 1. Sequential Quadratic Programming (SQP) method is taken as a reference for the convenience of this comparative study. From Table 1, several conclusions can be drawn as follows: (1) the system-level analysis can be remarkably reduced when using different CSSO methods, which shows that the CSSO methods are well suited for MDO problems; (2) the variant of GSECSSO method outperforms the RSCSSO method, in view of the optimization results as well as the required number of system- and disciplinary-level analysis.

variables	SQP	Variant of GSECSSO	RSCSSO
f	8.002860	8.002943	8.031944
x_1	3.028427	3.028442	2.993963
x_2	0.000000	0.000000	0.192359
x_3	0.000000	0.000000	0.000000
Number of system analysis	122	7	38
Number of discipline 1 analysis	0	72	179
Number of discipline 2 analysis	0	64	144

Table 1. Comparison of different CSSO methods with direct SQP for an analytical optimization problem

3.4.2 Example 2: Gear reducer optimization

This example is taken from the reference (Azarm & Li, 1989). The object is to minimize the overall weight, subject to the constraints for bending and torque stresses. The mathematical model for the optimization problem is as following:

$$
\begin{aligned}
\text{Min}\quad & f = 0.7854x_1x_2^2\left(3.3333x_3^2+14.9334x_3\text{-}43.0934\right)\text{-}1.508x_1\left(x_6^2+x_7^2\right)\\
& +7.477\left(x_6^3+x_7^3\right)+0.7854(x_4x_6^2+x_5x_7^2)\\
\text{s.t.}\quad & g_1 = 27x_1^{-1}x_2^{-2}x_3^{-1}-1\le 0\\
& g_2 = 397.5x_1^{-1}x_2^{-2}x_3^{-2}-1\le 0\\
& g_3 = 1.93x_2^{-1}x_3^{-1}x_4^3x_6^{-4}-1\le 0\\
& g_4 = 1.93x_2^{-1}x_3^{-1}x_5^3x_7^{-4}-1\le 0\\
& g_5 = 10x_6^{-3}\sqrt{\left(745x_2^{-1}x_3^{-1}x_4\right)^2+1.69\times 10^7}-1100\le 0\\
& g_6 = 10x_7^{-3}\sqrt{\left(745x_2^{-1}x_3^{-1}x_5\right)^2+1.575\times 10^8}-850\le 0\\
& g_7 = x_4^{-1}\left(1.5x_6+1.9\right)-1\le 0\\
& g_8 = x_5^{-1}\left(1.1x_7+1.9\right)-1\le 0\\
& g_9 = x_2x_3-40\le 0\\
& g_{10} = 5-x_1x_2^{-1}\le 0\\
& g_{11} = x_1x_2^{-1}-12\le 0\\
& x_1\in[2.6,3.6]\quad x_2\in[0.7,0.8]\quad x_3\in[17,28]\\
& x_4,x_5\in[7.3,8.3]\quad x_6\in[2.9,3.9]\quad x_7\in[5.0,5.5]
\end{aligned}
\tag{10}
$$

The optimization problem is classified into two disciplines: bearing discipline and shaft discipline, without any coupling relationship.

The optimization problem is solved by SQP and different CSSO methods. The results are depicted in Table 2. The convergence histories are shown in Fig. 5. From this comparative study, some conclusions can be drawn as follows: 1) Both GSECSSO and the variant of GSECSSO enable the reduction of system analysis, which in turn improve the efficiency; 2) No reduction of the system analysis is achieved by using RSCSSO. Nevertheless, the optimization procedure as well as the data-flow interface is much simpler, and the optimization can be potentially improved by implementing parallel system analysis for sample points and parallel subspace optimization; 3) The variant of GSECSSO method outperforms the GSECSSO method.

Variables	SQP	GSECSSO	Variant of GSECSSO	RSCSSO
f	2994.341316	2995.606759	2995.607422	2994.193811
x_1	3.5	3.500000	3.500000	3.500021
x_2	0.7	0.7	0.7	0.700000
x_3	17	17	17	17.000000
x_4	7.3	7.300000	7.3	7.300000
x_5	7.715320	7.733301	7.733330	7.715316
x_6	3.350215	3.352131	3.352156	3.349631
x_7	5.286654	5.287256	5.287246	5.286643
Number of system analysis	40	15	11	51
Number of bearing-discipline analysis	0	326	188	280
Number of shaft-discipline analysis	0	2986	1818	288

Table 2. Comparison of different CSSO methods with direct SQP for gear-box optimization

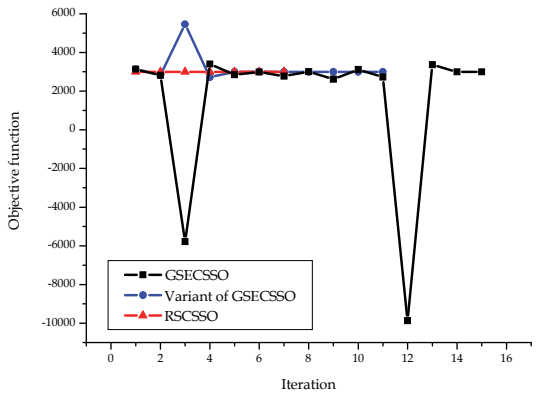


Fig. 5. Convergence history of CSSO optimization for gear reducer

4. Different frameworks for multi-objective concurrent subspace optimization

The MOPCSSO, MORCSSO and AWSCSSO methods will be discussed in this chapter as the typical multi-objective CSSO methods.

4.1 Multi-objective Pareto (MOPCSSO)

The constraint method is an effective multi-objective optimization method which optimizes preferred objective with others treated as constraints. MOPCSSO is developed by introducing the constraint method into the variant of GSECSSO.

Taking the MDO problem in Eq. (1) (two objectives, two coupled subsystems, i.e. N=2) as an example, the mathematical models of subspace optimizations of MOPCSSO can be written as

$$\begin{array}{l|l} \text{Sub-optimization 1} & \text{Sub-optimization 2} \\ \text{Min } F_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) & \text{Min } F_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right) \\ \text{s.t. } C_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right)\le C_1^0\left(1-r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right)\le C_1^0\left(1-r_2^1\right) \\ \hat{C}_2\left(\boldsymbol{X}_1\right)\le C_2^0\left(1-r_1^2\right) & C_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)\le C_2^0\left(1-r_2^2\right) \\ \hat{F}_2\le F_2^0 & \hat{F}_1\le F_1^0 \\ \boldsymbol{Y}_1=f_1\left(\boldsymbol{X}_1,\hat{\boldsymbol{Y}}_2\right) & \boldsymbol{Y}_2=f_1\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1\right) \end{array} \tag{11}$$

The variant of GSECSSO method described in Sub-section 3.2 is adopted in the single objective optimization for each objective function. The flowchart of MOPCSSO is same as that in Fig. 2. During the system optimization, all objective functions are improved simultaneously or individual objective function are improved without worsening the others. The optimization continues until no further improvement can be made so that the Pareto optimum can be obtained finally.

4.2 Multi-objective Range CSSO (MORCSSO)

The goal programming method is one of the most popular multi-objective optimization techniques that consider designer's preference. MORCSSO and Multi-objective Target CSSO (MOTCSSO) are both developed by combining idea of the goal programming method with MOPCSSO.

The $\left[F_2^{\min},F_2^{\max}\right]$ is supposed to be the preferred range of objective function F_2. The MDO problem in Eq. (1) (two objectives, two coupled subsystems, i.e. N=2) is taken as an example. The framework of MORCSSO is shown in Fig. 6. The variant of GSECSSO method described in Sub-section 3.2 is adopted for sub-optimizations and system-level coordination. The optimization is performed in two stages to obtain a preferred Pareto point.

1. If the design point does not meet the preference at beginning, the first-stage optimizations are implemented to obtain a design in the preferred range from the starting point.

The mathematical models of subspace optimizations of MORCSSO can be written as

If $F_2^0 < F_2^{\min}$,

$$
\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } F_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) & \text{Min } \left|F_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)-F_2^{\min}\right| \\
\text{s.t. } C_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right)\le C_1^0\left(1-r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right)\le C_1^0\left(1-r_2^1\right) \\
\quad \hat{C}_2\left(\boldsymbol{X}_1\right)\le C_2^0\left(1-r_1^2\right) & \quad C_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)\le C_2^0\left(1-r_2^2\right) \\
\quad -\hat{F}_2\le -F_2^0 & \quad \boldsymbol{Y}_2=f_1\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1\right) \\
\quad \boldsymbol{Y}_1=f_1\left(\boldsymbol{X}_1,\hat{\boldsymbol{Y}}_2\right) &
\end{array}
\tag{12}
$$

If $F_2^0 > F_2^{\max}$,

$$
\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } F_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) & \text{Min } \left|F_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)-F_2^{\max}\right| \\
\text{s.t. } C_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right)\le C_1^0\left(1-r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right)\le C_1^0\left(1-r_2^1\right) \\
\quad \hat{C}_2\left(\boldsymbol{X}_1\right)\le C_2^0\left(1-r_1^2\right) & \quad C_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)\le C_2^0\left(1-r_2^2\right) \\
\quad \hat{F}_2\le F_2^0 & \quad \boldsymbol{Y}_2=f_1\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1\right) \\
\quad \boldsymbol{Y}_1=f_1\left(\boldsymbol{X}_1,\hat{\boldsymbol{Y}}_2\right) &
\end{array}
\tag{13}
$$

2. In the second stage the design point is optimized closer to the Pareto frontier gradually within the preferred range.

After the optimization in the first stage, the design is in the designer's preferred objective range. Then the mathematical models of subspace optimizations of MORCSSO is same as those of MOPCSSO in Eq. (11).

In the course of optimization in the second stage, the design point maybe flees out of the preferred objective range again. In such a case, the optimization below should be performed.

If $F_2^0 < F_2^{\min}$,

$$
\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } F_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) & \text{Min } \left|F_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)-F_2^{\min}\right| \\
\text{s.t. } C_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right)\le C_1^0\left(1-r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right)\le C_1^0\left(1-r_2^1\right) \\
\quad \hat{C}_2\left(\boldsymbol{X}_1\right)\le C_2^0\left(1-r_1^2\right) & \quad C_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)\le C_2^0\left(1-r_2^2\right) \\
\quad -\hat{F}_2\le -F_2^0 & \quad \hat{F}_1\le F_1^0 \\
\quad \boldsymbol{Y}_1=f_1\left(\boldsymbol{X}_1,\hat{\boldsymbol{Y}}_2\right) & \quad \boldsymbol{Y}_2=f_1\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1\right)
\end{array}
\tag{14}
$$

If $F_2^0 > F_2^{\max}$,

$$
\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } F_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right) & \text{Min } \left|F_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)-F_2^{\max}\right| \\
\text{s.t. } C_1\left(\boldsymbol{X}_1,\boldsymbol{Y}_1,\hat{\boldsymbol{Y}}_2\right)\le C_1^0\left(1-r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right)\le C_1^0\left(1-r_2^1\right) \\
\hat{C}_2\left(\boldsymbol{X}_1\right)\le C_2^0\left(1-r_1^2\right) & C_2\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1,\boldsymbol{Y}_2\right)\le C_2^0\left(1-r_2^2\right) \\
\hat{F}_2\le F_2^0 & \hat{F}_1\le F_1^0 \\
\boldsymbol{Y}_1=f_1\left(\boldsymbol{X}_1,\hat{\boldsymbol{Y}}_2\right) & \boldsymbol{Y}_2=f_1\left(\boldsymbol{X}_2,\hat{\boldsymbol{Y}}_1\right)
\end{array}
\tag{15}
$$

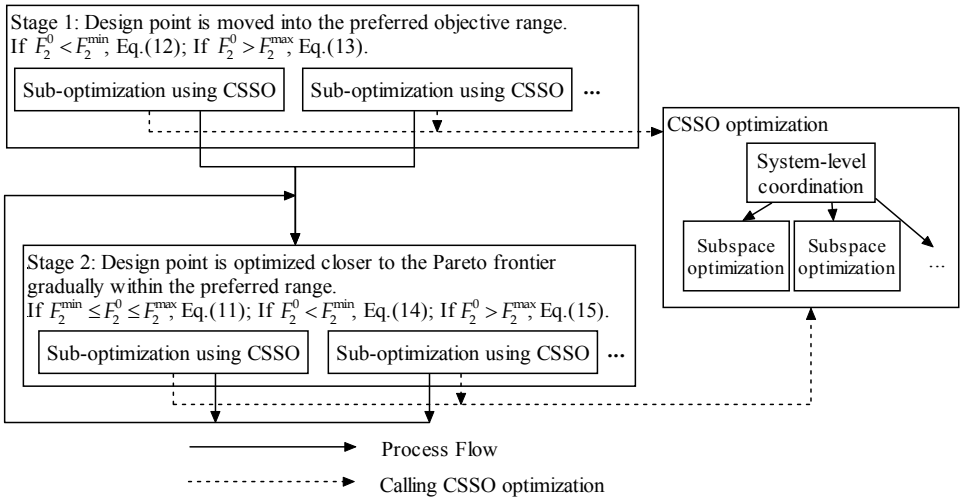


Fig. 6. Framework of MORCSSO method

For the MORCSSO method, in the case of $F_2^0 > F_2^{\max}$ in the second-stage optimizations, the optimizations may converge to the Pareto frontier above the preferred objective range, i.e. the Pareto point that $F_2 > F_2^{\max}$ is obtained.

Why does the optimization fail in this case? It can be analyzed from Eq. (15) in the objective space shown in Fig. 7. For Eq. (15), (1) $\boldsymbol{X}_0$ should be closer to the line $F_2 = F_2^{\max}$ according to Min $\left|F_2 - F_2^{\max}\right|$, (2) $F_2(\boldsymbol{X}_0)$ should be no more than F_2^0 according to $\hat{F}_2 \le F_2^0$, (3) $\boldsymbol{X}_0$ should optimized to decrease F_1^0 according to $\hat{F}_1 \le F_1^0$ and Min F_1, (4) $\boldsymbol{X}_0$ should be in the feasible region. As shown in Fig. 7, (2) and (3) forces $\boldsymbol{X}_0$ to move along a direction of $\vec{n}_1$ in the lower-left shadowed region, in which direction the design point will impossibly move into the preferred objective range. Only along a direction of $\vec{n}_2$ in the lower-right shadowed region the preferred range can be achieved. In such a case $\boldsymbol{X}_0$ may move down straightly along the direction of $\vec{n}$ to arrive at the Pareto front. The semi-infinite region between $F_1 \le F_{\max 1}$ and the Pareto front is named as Blind Region in this chapter, which means the point falling into this region will not converge to the Pareto front in the preferred region any more. This error will happen in the case of three and more objectives as well, as shown in Fig. 8.

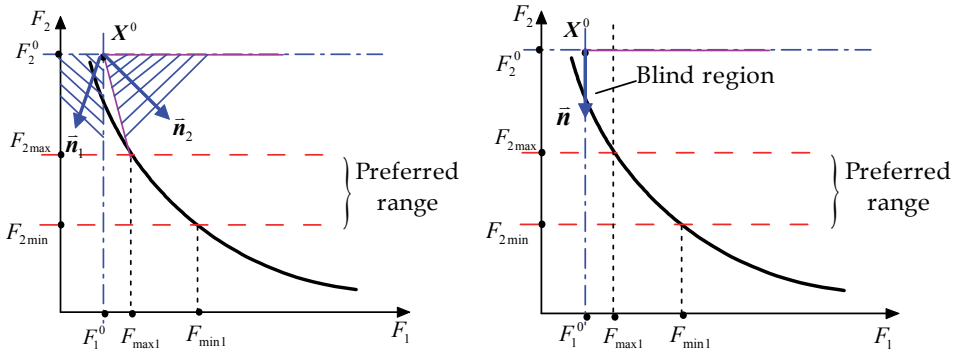


Fig. 7. The analysis of Eq. (15) in bi-objective space

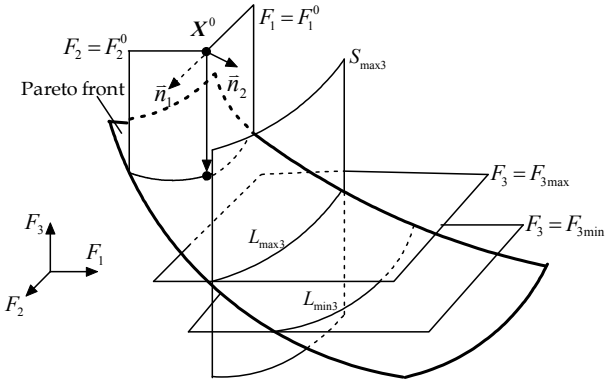


Fig. 8. The analysis of Eq. (15) in three-objective space

How to solve this problem? From Fig. 7, we only need to move the line $F_1 = F_1^0$ right a little bit, then the two shadowed region will be crossed each other. In the mathematical meaning $F_1 \le F_1^0$ is relaxed to $F_1 \le F_1^0 + \omega\left|F_1^0\right|$ in Eq. (15). This strategy is proven to be effective.

4.3 Adaptive Weighted Sum based CSSO (AWSCSSO)

The procedure of solving the Pareto front by AWSCSSO is similar to Adaptive Weighted Sum (AWS) method (Kim & de Weck, 2004, 2005). As an example, the AWSCSSO method for a generic bi-objective problem, with subsystem 1 and 2, is stated in the following paragraphs (Eq. (1): two objectives and two coupled subsystems). In the same way AWSCSSO can be also applied for the multi-objective problem with three or more subsystems.

1. In the first stage a rough profile of the Pareto front is determined.

The variant of CSSO method described in Sub-section 3.2 is adopted in the single objective optimization for each objective function and objective function is normalized as following:

$$\overline{J}_i = \frac{J_i - J^{\text{Nadir}}}{J^{\text{Utopia}} - J^{\text{Nadir}}} \tag{16}$$

When $\boldsymbol{X}^{i*}$ is the optimal solution vector for the single objective optimization of the ith objective function J_i, the utopia point and pseudo nadir point are defined as

$$J^{\text{Utopia}} = \left[J_1\left(\boldsymbol{X}^{1*}\right), J_2\left(\boldsymbol{X}^{2*}\right), \cdots, J_{\text{m}}\left(\boldsymbol{X}^{\text{m}*}\right) \right] \tag{17}$$

$$J^{\text{Nadir}} = \left[J_1^{\text{Nadir}}, J_2^{\text{Nadir}}, \cdots, J_{\text{m}}^{\text{Nadir}} \right] \tag{18}$$

Where $J_i^{\text{Nadir}} = \max\left[J_i\left(\boldsymbol{X}^{1*}\right) \; J_i\left(\boldsymbol{X}^{2*}\right) \cdots \; J_i\left(\boldsymbol{X}^{\text{m}*}\right) \right]$ and m is the number of objective functions.

Then with a large step size of the weighting factor the usual weighted sum method is used in the variant of GSECSSO to approximate the Pareto front quickly. The subspace optimization for AWSCSSO can be expressed as

$$\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } W_1 F_1\left(\boldsymbol{X}_1, \boldsymbol{Y}_1, \hat{\boldsymbol{Y}}_2\right) + W_2 \hat{F}_2\left(\boldsymbol{X}_1\right) & \text{Min } W_1 \hat{F}_1\left(\boldsymbol{X}_2\right) + W_2 F_2\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \boldsymbol{Y}_2\right) \\
\text{s.t. } C_1\left(\boldsymbol{X}_1, \boldsymbol{Y}_1, \hat{\boldsymbol{Y}}_2\right) \le C_1^0\left(1 - r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right) \le C_1^0\left(1 - r_2^1\right) \\
\quad \hat{C}_2\left(\boldsymbol{X}_1\right) \le C_2^0\left(1 - r_1^2\right) & \quad C_2\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \boldsymbol{Y}_2\right) \le C_2^0\left(1 - r_2^2\right) \\
\quad \boldsymbol{Y}_1 = f_1\left(\boldsymbol{X}_1, \hat{\boldsymbol{Y}}_2\right) & \quad \boldsymbol{Y}_2 = f_2\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \hat{\boldsymbol{Y}}_3\right)
\end{array} \tag{19}$$

Where the value with symbol '^' above is a linearly approximated one, C_1 and C_2 are cumulative constraints of $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$, respectively, and r_k^p represents responsibility assigned to the k-th subsystem for reducing the violation of C_p. The value with superscript '0' is corresponding to the starting point $\boldsymbol{X}_0$. W_1 and W_2 are weighting factors for objective function vector F_1 and F_2, respectively.

By estimating the size of each Pareto patch, the refined regions in the objective space are determined. An example of the mesh refinement in AWSCSSO is shown in Fig. 9. Where hollow points represent the newly refined node P_E (expected solution) while solid points represent initial four nodes that define the patch. As shown in Fig. 9, the quadrilateral patch is taken as an example. If the line segment that connects two neighboring nodes of the patch is too long, it is divided into only two equal line segments. The central point becomes the

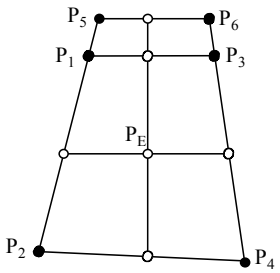
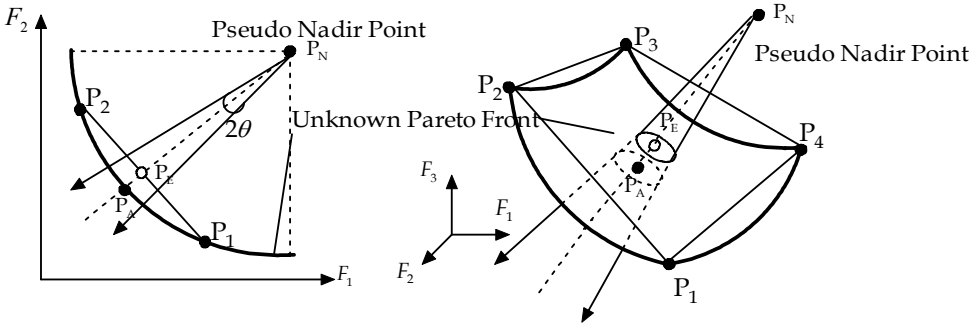


Fig. 9. Refine patches of AWSCSSO method

new refined node. These refined nodes are connected to form a refined mesh. Then the sub-optimizations in Eq. (19) are performed using different additional constraints for different refined nodes and the new Pareto points are obtained. In next step, according to the prescribed density of Pareto points, the Pareto-front patch that is too large will be refined again in the same way. In subsequent steps, the refinement and sub-optimizations are repeated until the number of Pareto points does not increase anymore.

2. In the subsequent stage only these regions are specified as feasible domains for sub-optimization problem with additional constraints. Each Pareto front patch is then refined by imposing additional equality constraints that connect the pseudo nadir point (P_N) and the expected Pareto optimal solutions (P_E) on a piecewise planar surface in the objective space (as shown in Fig. 10).



a) 2-D representation b) 3-D representation

Fig. 10. AWSCSSO method for multidimensional problems

Sub-optimizations are defined by imposing additional constraint H to Eq. (19) as

$$
\begin{array}{l|l}
\text{Sub-optimization 1} & \text{Sub-optimization 2} \\
\text{Min } W_1 F_1\left(\boldsymbol{X}_1, \boldsymbol{Y}_1, \hat{\boldsymbol{Y}}_2\right) + W_2 \hat{F}_2\left(\boldsymbol{X}_1\right) & \text{Min } W_1 \hat{F}_1\left(\boldsymbol{X}_2\right) + W_2 F_2\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \boldsymbol{Y}_2\right) \\
\text{s.t. } C_1\left(\boldsymbol{X}_1, \boldsymbol{Y}_1, \hat{\boldsymbol{Y}}_2\right) \le C_1^0\left(1 - r_1^1\right) & \text{s.t. } \hat{C}_1\left(\boldsymbol{X}_2\right) \le C_1^0\left(1 - r_2^1\right) \\
\hat{C}_2\left(\boldsymbol{X}_1\right) \le C_2^0\left(1 - r_1^2\right) & C_2\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \boldsymbol{Y}_2\right) \le C_2^0\left(1 - r_2^2\right) \\
\boldsymbol{Y}_1 = f_1\left(\boldsymbol{X}_1, \hat{\boldsymbol{Y}}_2\right) & \boldsymbol{Y}_2 = f_2\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \hat{\boldsymbol{Y}}_3\right) \\
H\left(\boldsymbol{X}_1, \boldsymbol{Y}_1, \hat{\boldsymbol{Y}}_2\right) \le 0 & H\left(\boldsymbol{X}_2, \hat{\boldsymbol{Y}}_1, \hat{\boldsymbol{Y}}_3\right) \le 0
\end{array}
\tag{20}
$$

The additional inequality constraint is

$$
H = -\frac{\left(\bar{\boldsymbol{F}}_E - \bar{\boldsymbol{F}}_N\right) \cdot \left(\bar{\boldsymbol{F}}(\boldsymbol{X}) - \bar{\boldsymbol{F}}_N\right)}{\left|\bar{\boldsymbol{F}}_E - \bar{\boldsymbol{F}}_N\right| \left|\bar{\boldsymbol{F}}(\boldsymbol{X}) - \bar{\boldsymbol{F}}_N\right|} + \mathrm{L} \le 0 \tag{21}
$$

Where L is the adaptive relax factor that is less than 1. $\bar{\boldsymbol{F}}_E$, $\bar{\boldsymbol{F}}_N$ and $\bar{\boldsymbol{F}}(\boldsymbol{X})$ are the normalized position vector of node P_E, P_N and the current design point X respectively. In

AWSCSSO, L is set to be increased with the rise of the distribution density of Pareto points.

Fig. 11 shows the framework of AWSCSSO. In Fig. 11, W^{1i} and W^{2i} are weighting factors in stage 1 and stage 2, respectively; H is the additional constraint. The optimization problem is performed in two stages in the AWSCSSO method. In the first stage the Pareto front is approximated quickly with large step size of weight factors. The optimization problems of this stage are defined in Eq. (19). In the subsequent stage, by calculating the distances between neighboring solutions on the front in objective space, the refined regions are identified and the refined mesh is formulated. Only these regions then become the feasible regions for optimization by imposing additional constraints in the objective space. The optimization problems of this stage are defined in Eq. (20). The different locations of new Pareto points are defined by the different additional constraints. Optimization is performed in each of the regions and the new solution set is acquired. Being a MDO problem, the optimization is performed by the variant of GSECSSO method.

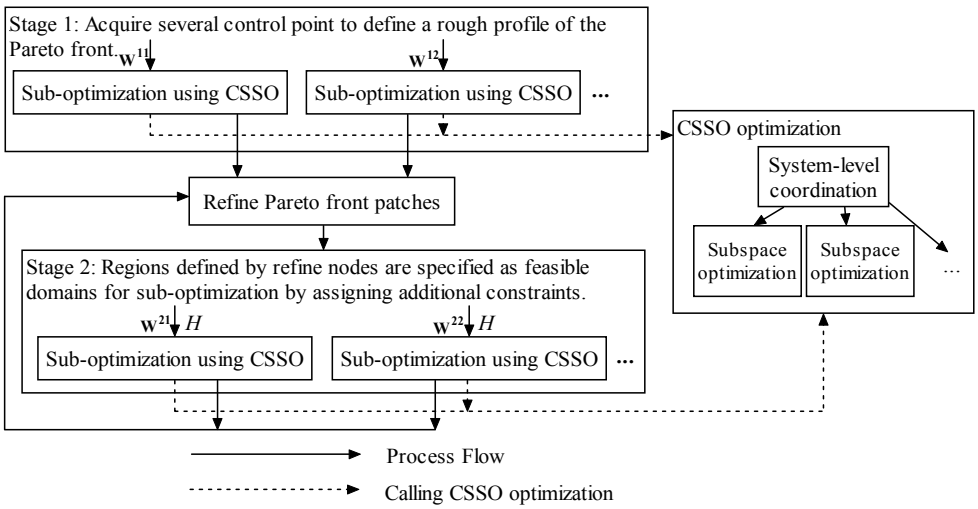


Fig. 11. Framework of AWSCSSO method

4.4 Examples

4.4.1 Example 1: Convex Pareto front

This problem is taken from a test problem (Huang, 2003). This is a test problem available in the NASA Langley Research Center MDO Test Suite. It has two objectives, F_1 and F_2, to be minimized. It consists of ten inequality constraints, four coupled state variables and ten design variables in two coupled subsystems. The mathematical model is not listed here for concision. We refer the readers to the test problem 1 in the corresponding references.

The comparison of the solution obtained by MOPCSSO and AWSCSSO is shown in Fig. 12. It can be concluded that for the problem with convex Pareto front a uniformly-spaced wide-distributed and smooth Pareto front can be obtained by AWSCSS method. When using MOPCSSO I have not captured the whole range.

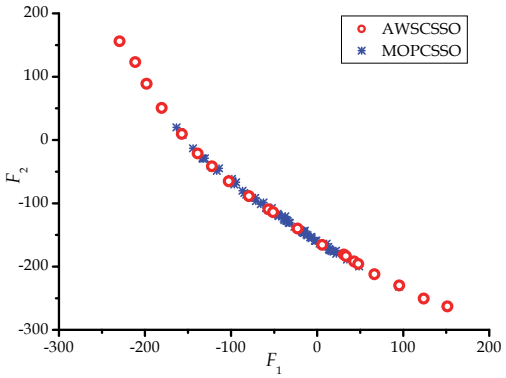


Fig. 12. Comparison of Pareto front obtained by using AWSCSSO and MOPCSSO

4.4.2 Example 2: Non-convex Pareto front

This problem consists of two objective functions, six design variables and six constraints. Two objectives, F_1 and F_2 need to be minimized. The model problem is defined as

$$
\begin{aligned}
&\text{Min} \quad F_1(\boldsymbol{x}) = -\left(25(x_1-2)^2 + (x_2-2)^2 + (x_3-1)^2 + (x_4-4)^2 + (x_5-1)^2\right) \\
&\text{Min} \quad F_2(\boldsymbol{x}) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 \\
&\text{s.t.} \quad c_1(\boldsymbol{x}) = x_1 + x_2 - 2 \geq 0 \\
&\qquad c_2(\boldsymbol{x}) = 6 - x_1 - x_2 \geq 0 \\
&\qquad c_3(\boldsymbol{x}) = 2 + x_1 - x_2 \geq 0 \\
&\qquad c_4(\boldsymbol{x}) = 2 - x_1 + 3x_2 \geq 0 \\
&\qquad c_5(\boldsymbol{x}) = 4 - (x_3-3)^2 - x_4 \geq 0 \\
&\qquad c_6(\boldsymbol{x}) = (x_5-3)^2 + x_6 - 4 \geq 0 \\
&\qquad 0 \leq x_1, x_2, x_6 \leq 10, \quad 1 \leq x_3, x_5 \leq 5, \quad 0 \leq x_4 \leq 6
\end{aligned}
\tag{22}
$$

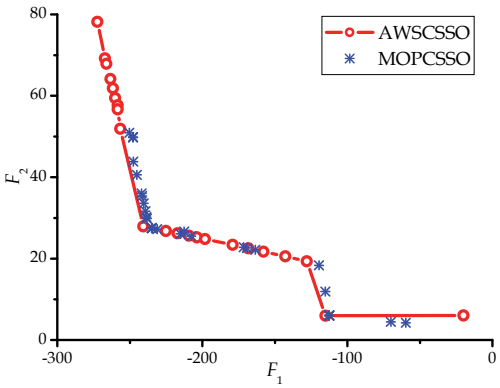


Fig. 13. Comparison of Pareto fronts obtained by using AWSCSSO and MOPCSSO methods

The comparison of Pareto front obtained by AWSCSSO and MOPCSSO is shown in Fig. 13. It is concluded that, for the problem with non-convex Pareto front, the more uniformly-spaced, more widely-distributed and smoother Pareto front is also obtained by the AWSCSSO method.

4.4.3 Example 3: Conceptual design of a subsonic passenger aircraft

The mathematical model of this problem is defined as

$$\begin{aligned} &\text{Max } U \\ &\text{Max } L/D_{\mathrm{C}} \\ &\text{s.t. } C_{\mathrm{d0L}} \le 0.2,\ C_{\mathrm{d0C}} \le 0.02 \\ &\quad R_{\mathrm{f}} \ge 1 \\ &\quad q_{\mathrm{To}} \ge 0.027,\ q_{\mathrm{L}} \ge 0.024 \\ &\quad D_{\mathrm{To}} \le 1981,\ D_{\mathrm{L}} \le 1371 \end{aligned} \tag{23}$$

The objective functions in Eq. (23) are to maximize useful load fraction (U) and lift-to-drag ratio for the cruising condition (L/D_{C}). The constraints in Eq.(23) are as follows. (1) The drag coefficient for the take-off condition and landing condition (C_{d0L}) is no more than 0.2 and that for the cruising condition (C_{d0C}) is no more than 0.02. (2) The overall fuel balance coefficient (R_{f}) is no less than 1. (3) The achievable climb gradient for the take-off condition (q_{To}) is greater than 0.027 and that for the landing condition (q_{L}) is greater than 0.024. (4) The take-off field length (D_{To}) is less than 1981m and the landing field length (D_{L}) is less than 1371m. The overall fuel balance coefficient is defined as the ratio of the fuel weight required for mission to that available for mission. The design variables are listed in Table 3.

Design Variable/unit	Symbol	Lower limit	Upper limit
Wing area/m²	S	111.48	232.26
Aspect ratio	AR	9.5	10.5
Design gross weight/10³kg	W_{dg}	63.504	113.400
Installed thrust/10³kg	T_{i}	12.587	24.948

Table 3. List of design variables

Two disciplines, aerodynamics and weight, are considered in this problem. The dataflow between and in subsystems is analyzed in Fig. 14, where L/D_{To}, L/D_{L}, L/D_{C} are the lift-to-drag ratios for the take-off, landing and cruising conditions respectively, V_{br} is the cruise velocity with the longest range, R_{fr} is the fuel weight fraction required for mission, and C_{d0C} is the zero-lift drag coefficient for the cruising condition.

Two disciplines, aerodynamics and weight, are coupled. When the state variables in aerodynamics such as cruise velocity with the longest range, lift coefficients, zero-lift drag coefficients, skin-friction drag coefficients, lift-to-drag ratio are computed, some state variables in Weight such as R_{fr} should be known. Similarly, when the state variables in Weight such as useful load fraction, overall fuel balance coefficient, achievable climb gradient on take-off and landing, take-off field length and landing field length are computed, some state variables in Aerodynamics such as L/D_{To}, L/D_{L}, L/D_{C} and V_{br} should also be provided. In the Aerodynamics discipline, V_{br} is coupled with C_{d0C}. The dataflow between state variables and design variables can be seen in Fig. 15. Many details of equations in the aerodynamic discipline model and weight discipline model can refer to the

reference (Zhang et al., 2008). The full description of them can be found in the references (Lewis & Mistree, 1995; Lewis, 1997).

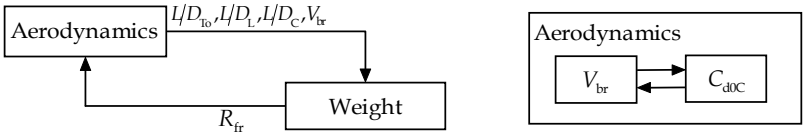


Fig. 14. Dataflow between and in subspaces

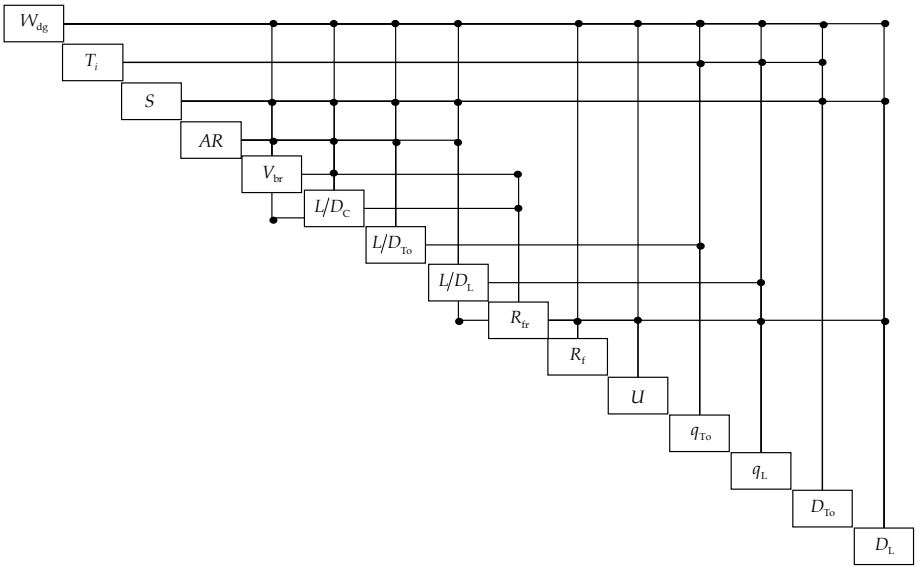


Fig. 15. Dataflow between state variables and design variables

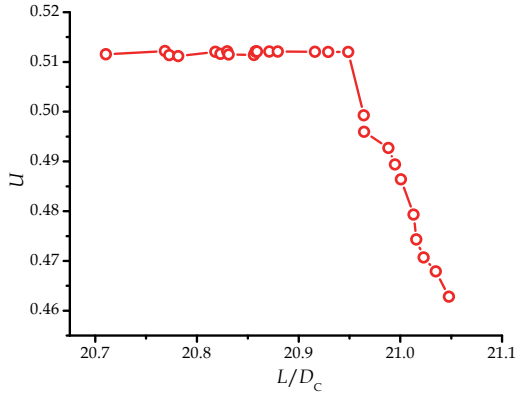


Fig. 16. Pareto front obtained by using AWSCSSO

The Pareto front obtained using AWSCSSO is shown in Fig. 16. Each solution on Pareto front is obtained using CSSO with iterative subspace optimizations. Taking one of the optimal designs as example, the values of the design variables are: S=232.3m², AR=10.5, W_{dg}=113.4×10³kg, T_i=16.75×10³kg. The performance parameters of the aircraft in optimal design are as follows: C_{d0C}=0.01777, L/D_C=21.05, V_{br}=183.43m/s, q_{To}=0.03303, q_L=0.08804, D_{To}=1823m and D_L=1086m. Several conclusions can be made from these results. 1) The AWSCSSO method is primarily proved to be applicable for aircraft conceptual design. 2) The distribution of Pareto points is not so uniform as expected. These results are still very encouraging in general. The non-uniformity may be due to the additional constraint that changes the location to expected solution. Further study is still needed on how to achieve the balance between uniformity and convergence.

5. Conclusion

The CSSO method is one of the main bi-level MDO methods. Couples of CSSO methods for single- and multi-objective MDO problems are discussed in this chapter. It can be concluded that, (1) number of the system analysis can be greatly reduced by using the CSSO methods, which in turn improve the efficiency; (2) the CSSO methods enable concurrent design and optimization of different design groups, which can greatly improve efficiency; (3) the CSSO methods are effective and applicable in solving not only single-objective but also multi-objective MDO problems.

For the CSSO methods, although the RSCSSO method behaves more robust, it will actually reduce to a single-level surrogate modeling based MDO method since the subspace optimizations have little impact on the results. So the GSECSSO method is more promising as a bi-level method and worth further studying. The future study on the GSECSSO method will focus on improving its robustness and efficiency. For the multi-objective CSSO methods, the AWSCSSO method behaves better on obtaining widely-distributed Pareto points. The future work on the multi-objective CSSO methods will focus on improving the solution quality and also on testing it for more realistic engineering design problems.

6. References

Aute, V. and Azarm, S. A, "Genetic Algorithms Based Approach for Multidisciplinary Multiobjective Collaborative Optimization," *11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Virginia, 2006, AIAA 2006-6953.

Azarm, S. and Li, W.C., "Multi-level Design Optimization using Global Monotonicity Analysis," *ASME Journal of Mechanisms and Automation in Design*, 1989, Vol.111, No.2, pp.259-263.

Bolebaum, C.L., "Formal and Heuristic System Decomposition in Structural Optimization," NASA-CR-4413, 1991.

Huang, C. H., "Development of Multi-Objective Concurrent Subspace Optimization and Visualization Methods for Multidisciplinary Design," Ph.D. Dissertation, The State University of New York, New York, 2003.

Huang, C. H. and Bloebaum, C. L., "Incorporation of Preferences in Multi-Objective Concurrent Subspace Optimization for Multidisciplinary Design," *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, New York, 2004, AIAA 2004-4548.

Huang, C. H. and Bloebaum, C. L., "Multi-Objective Pareto Concurrent Subspace Optimization for Multidisciplinary Design," *AIAA Journal*, Vol. 45, No. 8, 2007, pp. 1894-1906.

Kreisselmeier, G., Steinhauser, R., "Systematic Control Design by Optimizing a Vector Performance Index," *IEAC Symposium on Computer Aided Design of Control Systems*, Zurich, Switzerland, 1979.

Kim, I. Y. and de Weck, O. L., "Adaptive Weighted Sum Method for Multiobjective Optimization," *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, New York, 2004, AIAA 2004-4322.

Kim, I. Y. and de Weck, O. L., "Adaptive Weighted-sum Method for Bi-objective Optimization: Pareto Front Generation," *Structural and Multidisciplinary Optimization*, No. 29, 2005, pp. 149-158.

Lewis, K., Mistree, F. "Designing Top-level Aircraft Specifications-A Decision-based Approach to A Multiobjective, Highly Constrained Problem," *6th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Bellevue, WA, AIAA 1995-1431.

Lewis, K., "An Algorithm for Integrated Subsystem Embodiment and System Synthesis," Ph.D. Dissertation, Georgia Institute of Technology, Atlanta, Georgia, August 1997.

McAllister, C. D., Simpson, T. W. and Yukesh, M. "Goal Programming Applications in Multidisciplinary Design Optimization," 8th *AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, CA, 2000, AIAA 2000-4717.

McAllister, C. D., Simpson, T. W., Lewis, K. and Messac, A., "Robust Multiobjective Optimization through Collaborative Optimization and Linear Physical Programming," *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, New York, 2004, AIAA 2004-4549.

Orr, S. A. and Hajela, P., "Genetic Algorithm Based Collaborative Optimization of A Tilrotor Configuration," *46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference*, Texas, 2005, AIAA 2005-2285.

Parashar, S. and Bloebaum, C. L., "Multi-objective Genetic Algorithm Concurrent Subspace Optimization (MOGACSSO) for Multidisciplinary Design," *47th AIAA/ASME/ ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Rhode Island, 2006, AIAA 2006-2047.

Renaud, J.E. and Gabriele, G.A., "Second Order Based Multidisciplinary Design Optimization Algorithm Development," *Advance in Design Automation*, 1993, Vol.65, No.2, pp.347-357.

Renaud, J.E. and Gabriele, G.A. "Improved coordination in non-hierarchic system optimization," AIAA Journal, 1993, Vol.31, No.12, pp.2367-2373.

Renaud, J.E. and Gabriele, G.A., "Approximation in non-hierarchic system optimization," *AIAA Journal*, 1994, Vol.32, No.1, pp.198-205.

Sellar, R. S, Batill, S. M and Renaud, J. E., "Response surface based, concurrent subspace optimization for multidisciplinary system design," *34th AIAA Aerospace Sciences Meeting*, 1996, AIAA 96-0714.

Sobieszczanski-Sobieski, J., "Optimization by Decomposition: A Step from Hierarchic to Non-hierarchic Systems," Recent Advances in Multidisciplinary Analysis and Optimization, NASA CP-3031, Hampton, 1988.

Sobieszczanski-Sobieski, J., "Sensitivity of Complex, Internally Coupled Systems," *AIAA Journal*, Vol. 28, No. 1, 1990, pp. 153-160.

Tappeta, R. V. and Renaud, J. E., "Multiobjective Collaborative Optimization," *ASME Journal of Mechanical Design*, No. 3. 1997, pp. 403-411.

Zhang, K.S., Han, Z.H., Li, W.J., and Song, W.P., "Bilevel Adaptive Weighted Sum Method for Multidisciplinary Multi-Objective Optimization," *AIAA Journal*, 2008, Vol.46 No.10, pp. 2611-2622.

10

The Assessment Method for Multi-Azimuth and Multi-Frequency Dynamic Integrated Stealth Performance of Aircraft

Ying Li, Jun Huang, Nanyu Chen and Yang Zhang
Beijing University of Aeronautics and Astronautics, Beijing, China

1. Introduction

Stealth technology of aircraft, known as one of the three technological revolutions together with high-energy laser weapons and cruise missiles in the development history of military science since 1980s, has become the third milestone after jet engines and swept wings technology in modern aviation history. Stealth aircraft has been also considered as one of the ten greatest inventions of the U.S. ADARPA (Defense Advance Research Project Agency). Nowadays, stealth technology has become one key technology. The countries all over the world have paid great attention and strived to develop the stealth technology. Reasonable assessment method for stealth performance plays a crucial role in the development of advanced stealth technology. For example, the result of the stealth performance assessment of aircraft can provide reference for modifying the aircraft's stealth design to achieve a higher stealth performance. Meanwhile, it can also provide reference for making some specific strategies to increase the probability of successful tasks by reducing the detecting probability of the radar or radar network. Nowadays, the electronic battlefield is becoming more complex and it is urgent to build up a new method to analyze the multi-azimuth and multi-frequency dynamic integrated stealth performance of aircraft, under the complex electronic environment.

The existing stealth performance assessment methods include two types. One is the static stealth performance assessment method, the other is the stealth assessment method based on the effectiveness of combat simulation. The former just uses the average RCS value of target circumferential area, or that of some critical radar detecting areas, under some important radar frequencies as the basis. And the latter uses aircraft's survival probability (including detection probability, hitting probability and damage probability) in specific combat tasks as the basis to assess the stealth performance of aircraft. Each method can reflect the characteristics of the target's stealth performance well. However, there are still some limitations such as: the results based on the two methods can't reflect the impact on aircraft stealth performance caused by different target scattering characteristics. As not taking the certain combat task and detecting environment into consideration, the results also fail to reflect the dynamic characteristics of stealth performance, which happens in the entire

proceeding of different combat tasks of aircraft, and the multi-azimuth and multi-frequency integrated stealth performance under complex electronic environment. According to the above, the development of modern stealth technology urgently requires a new type of stealth performance assessment method to provide the reliable basis.

In this paper, a new type assessment method for multi-azimuth and multi-frequency dynamic integrated stealth performance of aircraft was established by building the multi-azimuth and multi-frequency dynamic integrated stealth performance assessment models and the series of assessment criteria. With these efforts, on the one hand, a more reasonable analysis result of stealth performance based on the multi-azimuth and multi-frequency dynamic detecting environment can be provided, on the other hand, the limitations of the existing stealth performance assessment methods can be overcame.

2. The existing stealth performance assessment method

Although lots of countries home and abroad started researches about the stealth technology early, there are just only a few assessment methods for stealth performance of aircraft. One of the two main existing methods is static stealth performance assessment method, the other is the stealth performance assessment method based on the effectiveness of combat.

2.1 The static stealth performance assessment method

2.1.1 The theory of the static stealth performance assessment

The classical static stealth performance assessment method includes two aspects: one is the static testing assessment method and the other is theoretical calculation assessment method. The American scholar Knott E.F has made a number of deep studies into the radar cross section calculation and testing technologies. Chinese scholars, such as Ruanying Zheng, Kao Zhang, Dongli Ma and so on, also have done researches on radar cross section calculation and testing, they propose the concept of critical RCS reduction region of aircraft, which has been widely used. At present, the radar cross section testing technology home and abroad has been used widely. The static radar cross section testing is a method, by doing the outfield or laboratory RCS testing on made full-scale models or reduce-scale models to get a basic understanding of the target scattering characteristics. The existing methods of the radar cross section theoretical calculation mainly include three types and they are the high frequency approximation, finite difference time domain and finite difference time domain.

The detailed steps of static stealth performance assessment are as follows: First of all, obtaining the RCS curve of the target under different radar detecting frequencies through static testing method or theoretical calculation method, then analyzing the stealth performance of aircraft, according to the average RCS of target circumferential area, or that of some critical radar detecting area, or the RCS value of certain radar detecting azimuth, under some important radar frequencies.

The assessment criteria of the static stealth performance assessment method is that the lower the average RCS of target circumferential area, or that of some critical radar detecting area, or the RCS of specific azimuth, under some important radar frequencies, the better the stealth performance of aircraft has. Technology flow chart of the static stealth performance assessment method is shown as Fig.1.

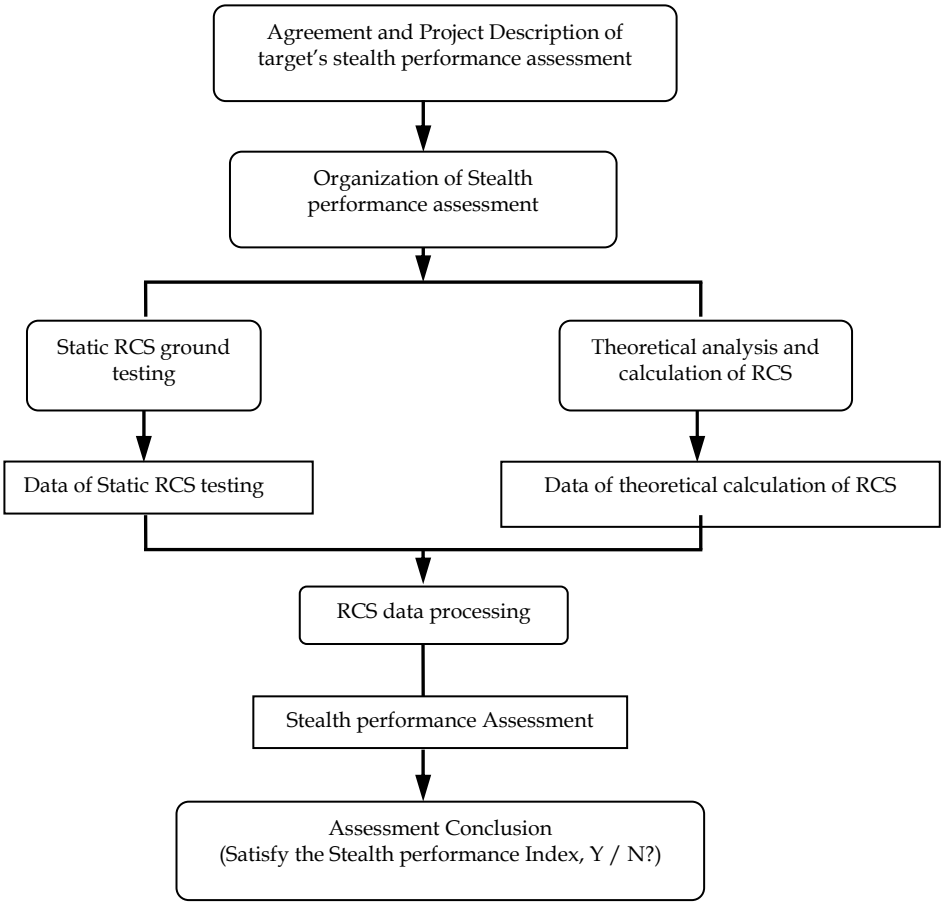


Fig. 1. Flow Chart of Static Stealth performance Assessment Method

The overall average RCS of both model one and model two is 0.5㎡. There are three RCS curve peaks at $\pm 30^{\circ}$ and 180° azimuth for model one. The maximum value of RCS curve peak is 10 ㎡ and the azimuth-width is 4°. There are four RCS curve peaks at $\pm 45^{\circ}$, 135° and 225° azimuth. The maximum value of RCS curve peak is 20㎡ and its peak azimuth takes up 4° as well.

A penetration testing is carried out in this section. The locations of every single radar in the radar network and the penetration destination are shown in Table 1, where GR means single radar and Basement stands for the penetration destination. The penetration testing angle is set from -60^{0} to $+60^{0}$ and the interval angles is 5°.

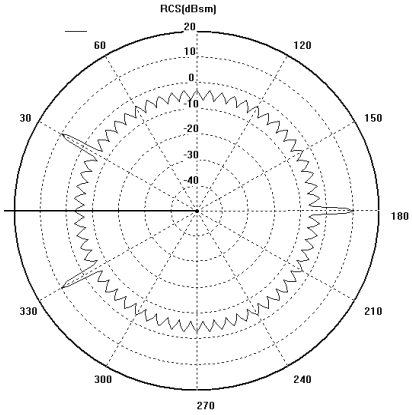


Fig. 2. Circumferential scattering distributions of model one

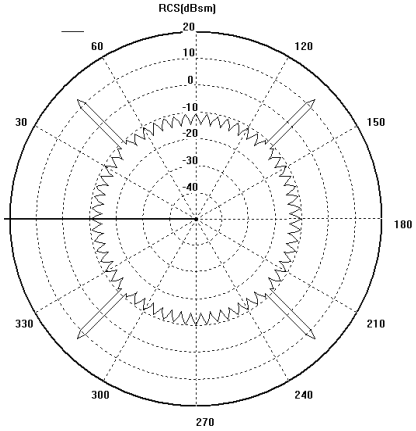
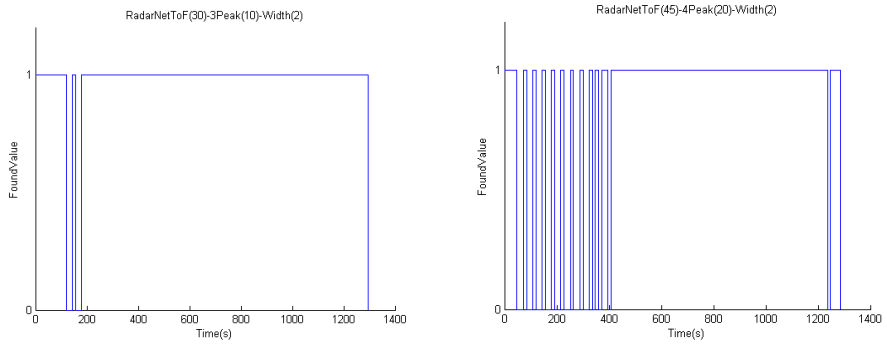


Fig. 3. Circumferential scattering distributions of model two

Name	GR1	GR2	GR3	GR4	GR5	GR6	GR7	Basement
Longitude	119.6330	121.550	120.4830	121.5330	121.6170	121.05	121.9667	121.5
Latitude	23.5667	24.0667	22.70	25.0330	24.0167	25.0667	24.8	25.0

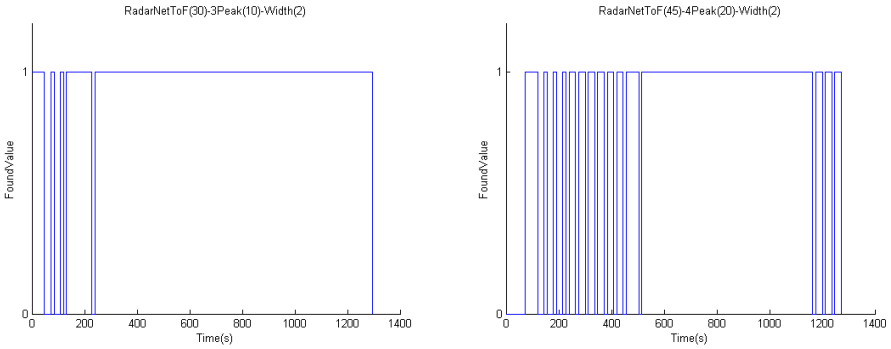
Table 1. Locations of Each Single Radar and the Penetration Destination

Fig.4 to Fig.7 just show several simulation results of these two models. The Y axis of the testing diagram the FoundValue stands for the radar network detection results. When FoundValue is equal to 0, it means that no target has been found by radar network. When FoundValue is equal to 1, it means the target has been found. The X axis of the diagram the Time stands for the time length of the penetration testing process. Fig (a) is the radar detection results of simulation model one and Fig (b) is the radar detection results of simulation model two.



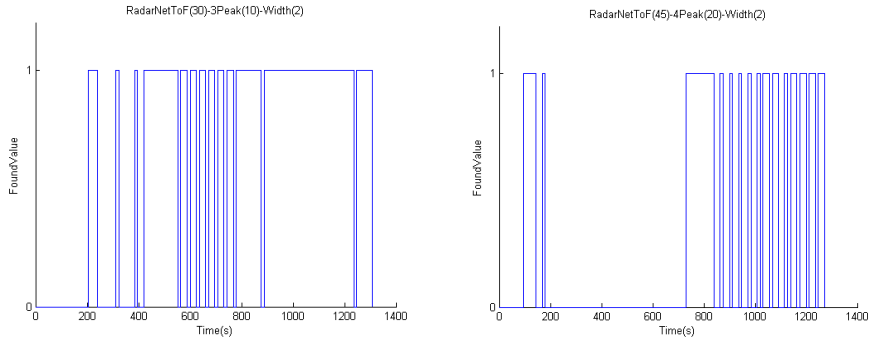
(a) (b)

Fig. 4. Simulation results comparison of model one and model two (penetration angle -60°)



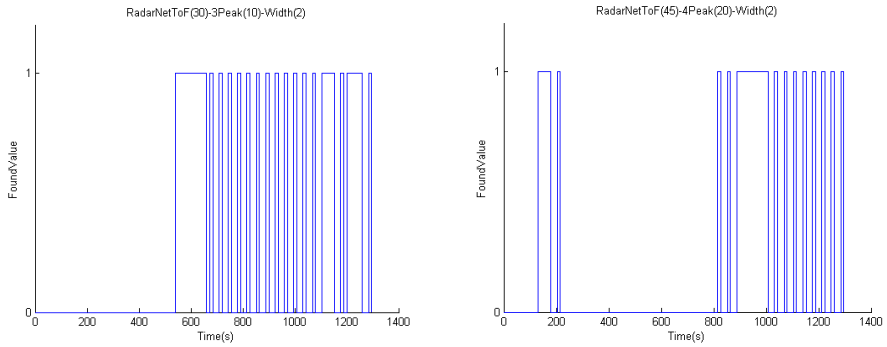
(a) (b)

Fig. 5. Simulation results comparison of model one and model two (penetration angle -45°)



(a) (b)

Fig. 6. Simulation results comparison of model one and model two (penetration angle 0°)



(a) (b)

Fig. 7. Simulation results comparison of model one and model two (penetration angle 15°)

It can be seen from the results that although the overall average RCS of the two simulation models is 0.5m^2, the dynamic stealth performance of these two models differs much from each other. Such as it is shown in Fig.6 and Fig.7, the simulation model two would be discovered earlier by radar network than the model one at the same penetration angle. After the model two was discovered for the first time, it was lost by the radar network for a long time, while the model one was detected continuously by radar network after its first being found.

Above all, the static stealth performance assessment method still has some limitations. Conclusions drawn from the tests are listed as follows:

1. Different circumferential scattering distribution can makes the aircraft has absolutely different stealth performance, even if they have the same circumferential average RCS;
2. According to the stealth performance of aircraft with different circumferential scattering distribution, the average RCS of aircraft requirement could be appropriately relaxed.

2.3 The assessment method based on the effectiveness of combat

Based on the radar simulation and target signal simulation technology, the steps of assessment method based on the effectiveness of combat are listed: first of all, calculating the aircraft survival probability during the whole combat mission, then summarizing the effects of RCS characteristic on aircraft survivability and evaluating the stealth performance of aircraft. The rules of this method are the aircraft with higher survivability has better stealth performance. The assessment method based on aircraft effectiveness of combat has been investigated early abroad. Ball R E made deep studies into the RCS reducing technology which would improve the survivability of aircraft greatly. At the same time, some software corporations abroad developed many kinds of analysis software that were applied for analyzing survivability and vulnerability of aircraft. For example, several kinds of software developed by SURVIAC center could quantitatively and comprehensively evaluate the survivability and vulnerability of aircraft in the situation of one to one air battle. In China there are a great deal of research about the effects of aircraft RCS on its survivability have been done. For example, Zhang Kao and Ma Dongli proposed the method of calculating the survivability of stealth aircraft that carries out given mission. Aimed at analyzing the effects of reducing aircraft RCS on survivability of aircraft.

Assessment method based on the effectiveness of combat reflects the stealth performance of aircraft carrying out the given mission, but still has some limitations: firstly, this method couldn't reflect the process and dynamic character of the stealth performance of aircraft during the whole mission. For example, the target 'flashing signal' caused by the different distributions of strong scattering source. Secondly, the assessment method based on the survivability could not reflect the character of multi-azimuth and multi-frequency dynamic comprehensive stealth performance of aircraft, and it also ignores characters of the multi-azimuth and multi-frequency electronic detecting environment.

3. The multi-azimuth and multi-frequency dynamic integrated stealth performance assessment method

Because of the limitations of the existing assessment method told in chapter 2, it should build up a new assessment method for stealth performance of aircraft. This method could evaluate the multi-azimuth and multi-frequency dynamic comprehensive stealth performance of aircrafts, not just gives the evaluation conclusions based on the average RCS of target circumferential area, or that of some critical radar detecting areas. The aircraft with different scattering characteristic could meet requirements of different missions. In order to evaluate the multi-azimuth and multi-frequency dynamic comprehensive stealth performance of aircraft, it should build up the RCS scattering model and the multi-azimuth and multi-frequency dynamic detecting environment model when the aircraft carries the given mission, then make the assessment rules for analyzing the multi-azimuth and multi-frequency dynamic comprehensive stealth performance of aircraft.

3.1 The multi-azimuth and multi-frequency dynamic comprehensive assessment models

This section describes the building steps of the multi-azimuth and multi-frequency dynamic comprehensive assessment models in detail, the assessment model includes the typical complex dynamic detecting environment simulation model based on the given mission and the RCS scattering model of aircraft.

3.1.1 The RCS scattering model of aircraft

In order to build a model that could reflect the RCS scattering character of aircraft and use this model to analyze the stealth performance, we need to guarantee the accuracy of the model.

Aircrafts executing different missions will encounter different detecting threat at different azimuths from land, sea, air and space. And even if using the same detector with the same working mode, the signal of aircraft the radar detected may still changes at any moment in the mission. So the method of building reduced-scale model of aircraft and doing experiments is considered. We can acquire the corresponding data for building up the scattering model of aircraft. The RCS database should include the data of aircraft RCS at different azimuths, under different frequencies. It is impossible to meet the needs of building the database by the way of building model and testing it because of the limitation of experiment condition. So the feasible way is to combine the data from both experiment and theoretical calculating. Fig 8 shows the detailed steps.

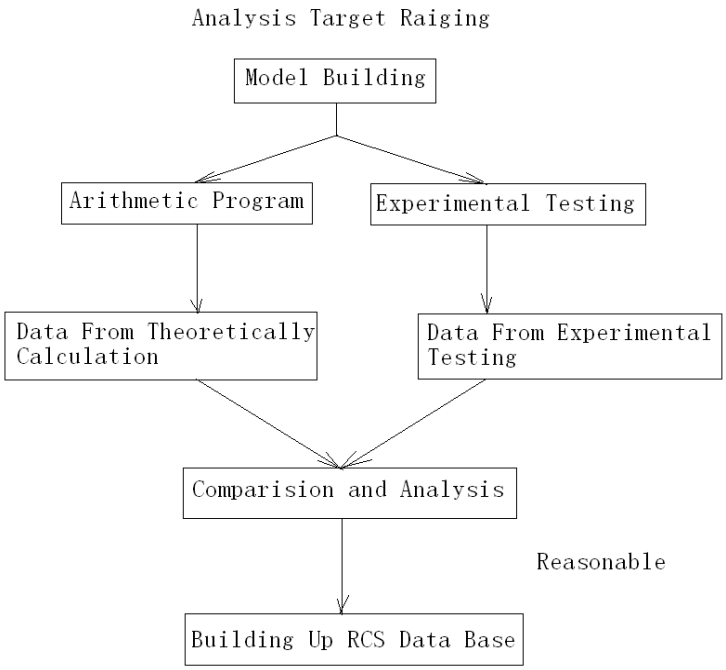


Fig. 8. Flow chart of building the RCS database

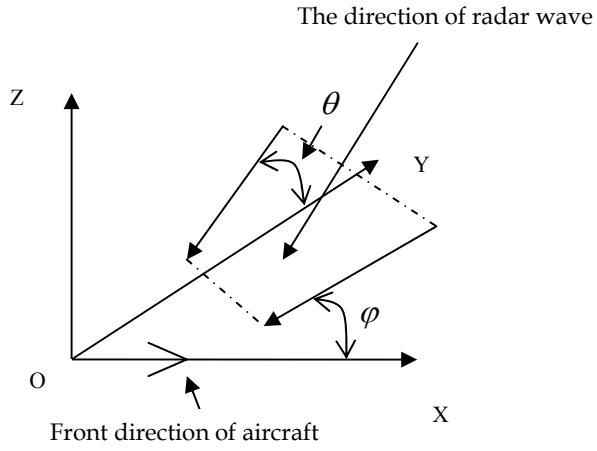


Fig. 9. Schematic illustration of the aircraft scattering model

The aircraft scattering model can be built up based on the adequate RCS data. The detailed steps of building up the aircraft scattering model are showed as Fig.9. The coordinate system in Fig.9 is defined just the same as aircraft body axes coordinate system for building

up equation of motion, which follows the right-handed screw rule. θ and φ are two parameters of radar detecting wave. φ is defined as the angle between the project of radar wave on the XOY plane and the X-axes. θ is defined as the angle between the project of radar wave on the YOZ plane and the Y-axes. θ and φ together decide the location of radar wave in aircraft body axes coordinate system.

The planform of one aircraft scattering model is shown as Fig10. Fig11 shows the test RCS curve of this model. We can see from the two figures that the model's RCS scattering character is consistent with experiment result which looks like a butterfly. This building method of the RCS scattering model is feasible. The method of building a scattering model is an innovation, which accurately reflects the RCS scattering character of the aircraft under different frequencies. This method could be applied for analyzing different kinds of stealth aircraft.

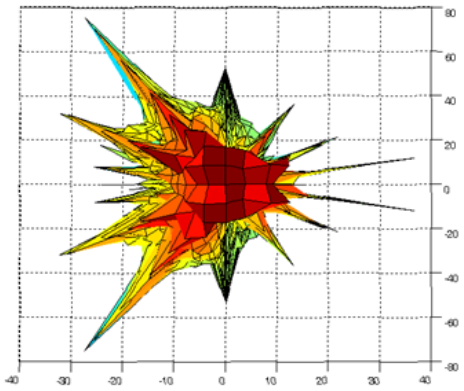


Fig. 10. Planform of the Aircraft Scattering Model

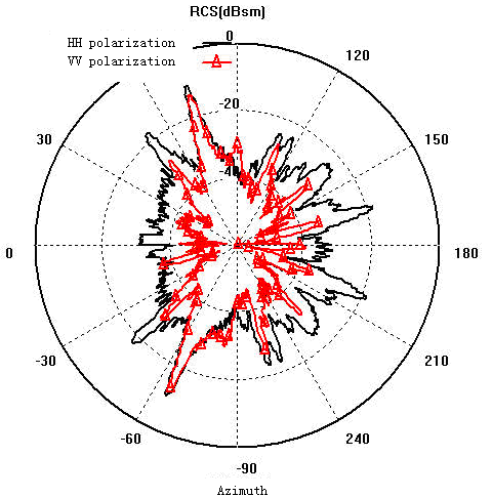


Fig. 11. Circumferential Test RCS curve of Aircraft

3.1.2 The typical complex dynamic detecting environment model

The typical detecting environment is different when the aircraft executes different missions. So if we want to evaluate the multi-azimuth and multi-frequency dynamic comprehensive stealth performance of aircraft, we must consider the typical complex detecting environment. For example, when an aircraft is executing a penetration mission, the main radar detecting thread is from the head or tail direction of aircraft.

The detection environment of warfare is becoming more complex, so it is extremely difficult to describe it completely and accurately. By studying the performance of radar detection systems, we can summarize the typical detection environment. Here are the main elements to describe the typical complex and dynamic detection environmental model including two types, one is the main tactical applications, such as (1) characters of the two sides of combat. (2) Threat which the opposing sides may meet. (3) Interference and anti-interference measures of the opposing sides. The other is the related information of electronic equipment in typical detection environment, such as: (1) Number of detectors. (2) Spatial distribution of detectors. (3) Density of the detectors (time domain). (4) Parameters of the detector's signal. (5) Frequency and scope of the detector's signal. (6) The power of detector. (7) Working mode of detector. Using these model parameters above, we can accurately describe the typical complex dynamic detection environment when aircraft carries a specific mission.

3.2 The multi-azimuth and multi-frequency dynamic comprehensive stealth assessment rules

As mentioned above, the existing assessments method is based on the average RCS of aircraft. In order to make sure that the aircraft stealth performance assessments conclusions are applicable, which can be used to guide the new type of stealth aircraft design and stealth performance analysis, we should consider the specific tasks the aircraft carries. When performing different tasks, the aircraft may encounter different typical detection environment. So it is not reasonable to use the same stealth performance assessment rules. Therefore, this section will establish the multi-azimuth and multi-frequency dynamic integrated stealth performance assessment rules based on the given mission. Based on specific combat mission, the assessment rules should be built up by analyzing various types of detection threat the aircraft may encounter and the influence of the aircraft scattering characteristic. Based on the conclusions got from the assessment rules above, we can compare and analyze the stealth performance of different aircrafts which perform the same task. The task-based aircraft multi-azimuth and multi- frequency dynamic integrated stealth performance assessment rules should satisfy two conditions:

1. Assessment rules should reflect the character of specific tasks carried by aircraft. For example, the stealth performance assessment rules for penetration aircraft should reflect the character of penetration task firstly. In different stages of penetration task, the influence of the aircraft survivability and successful mission probability is different. For example, the first time when penetration aircraft is found by radar (or radar network) decides how many times the aircraft would probably be attacked by enemy firepower. When penetration aircraft was first discovered by radar network, the distance from the penetration destination determines whether the plane could use the remote attack weapon. Therefore, the establishment of penetration stealth performance assessment rules should be combined with the penetration characters, so that the evaluation conclusions reflect the differences of stealth performance among different aircraft when executing penetration.

2. The conclusion got from the assessment rules should be able to reflect the scattering character of aircraft. For example, during the process of penetration, aircraft mainly get detecting thread from its front azimuth by various types of detectors in the enemy air defense. If the aircraft goes through the first line of air defense system undetected and continues to fly to the enemy's air defense system, it may also be detected by radars at its two side azimuths. Therefore, when establishing stealth performance assessment rules, we should focus on the stealth performance of the front and two side azimuths of aircraft. The stealth performance of penetration aircraft is changing all the time. Therefore, the assessment rules should not only combine with combat characteristics of the penetration mission, but also need to consider the dynamic scattering character of penetration aircraft.

4. The effects of different aircraft circumferential RCS scattering characters

The aircrafts with different circumferential scattering characters is suitable for different combat missions, caused by various configuration design parameters. The stealth performance evaluation conclusions of aircraft with different circumferential RCS scattering characters can provide reference for a reasonable layout design of new type of aircraft. This section will combine the multi-azimuth and multi-frequency dynamic comprehensive assessment rules developed in chapter 3 to analyze and evaluate the aircrafts with different circumferential RCS scattering characters.

4.1 The new target integrated circumferential RCS scattering model

4.1.1 The relevant model parameters

There are several requirements which the model should meet for the new analysis method: 1) quantified the overall and partial RCS scattering characters of target; 2) setting up the relations between each different radar detecting areas; 3) controlling the RCS scattering changing trends of model through inducing several model RCS scattering control parameters. Considering that the RCS value can reflect the quantified target scattering characters and the differences between the RCS values may up to the magnitude order for the different target azimuth, the average RCS value of target circumferential area is introduced as one of the RCS scattering characteristic parameters of model and the unit is dBsm, and its symbol is δ_{ave}.

The existing analysis method does not take the effects of the changing relations between different important radar detecting region on the target stealth ability into account. For example, the aircraft could be excellent in depth penetration mission, if it has lower RCS value at the front azimuth and higher RCS values at other azimuths. Moreover, if the stealth aircraft has lower RCS value at its two side azimuths compared with that for the rest azimuths, it can carry out penetration mission with a smaller horizontal distance arriving at the enemy bases. In order to build up the relations between different target important detecting areas, the new analysis method uses the average RCS value of target front important radar detect area as the basis. Furthermore, the average RCS values of another radar detecting areas are introduced into the model as the target local RCS scattering characteristic parameters and the corresponding symbols are $\overline{\delta_i}$ (i=0,1,…). The subscript i represents the sequence of radar detecting areas. For describing the relations between the

target front and other direction important radar detecting areas, the new model defines a set of relational parameters. Their symbols are k_{δ_i} (i = 0,1,…) and can be written as:

$$k_{\delta_i} = \bar{\delta}_i / \bar{\delta}_0 \tag{1}$$

where $\bar{\delta}_0$ represents the average RCS value of the target front important radar detecting area.

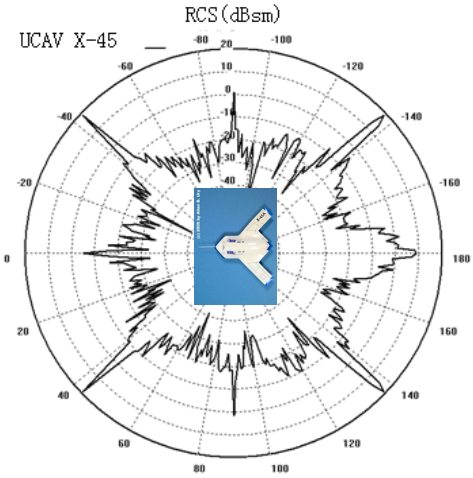


Fig. 12. RCS curve corresponding to one type of stealth aircraft (Under the S wave band).

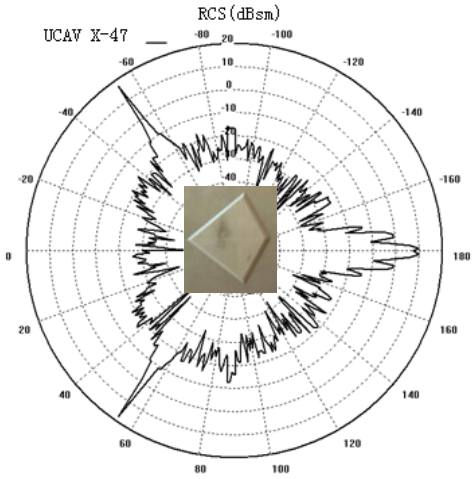


Fig. 13. RCS curve corresponding to one type of aircraft (Under the S wave band).

Due to different stealth design parameters the UCAV X-45 and X-47 have, so they have completely different RCS scattering characters. Their RCS curves differ much from each other, as shown in Fig.12 and Fig.13. So the same set of RCS scattering controlling parameters can not be used to describe the dissimilar RCS curve patterns and control the RCS scattering changing trends of various new models well. There are two requirements for the RCS scattering control parameters: one is that it is not advisable to introduce too many RCS scattering control parameters, the other is the model can satisfy all kinds of stealth ability analysis requirements. For example, building up the target circumferential RCS scattering model with triangle pattern character as shown in Fig.2 needs two RCS scattering control parameters, which can meet the requirements of controlling RCS scattering changing trends and conducting target integrated stealth performance analysis. These parameters are K_L and K_A respectively, K_L and K_A can be expressed as:

$$K_L = L_a / L_b \tag{2}$$

$$K_A = A_F / 360 \tag{3}$$

where L_a and L_b are the side lengths of model with triangle circumferential RCS scattering character. The parameter K_L can control the RCS scattering changing trends of target head and tail areas. By this way, it can satisfy the analysis requirements about the effects of different target head and tail stealth performance on its integrated stealth performance. A_F represents the angular region of target front important radar detecting area. Likewise, the models with different A_F values can meet the analysis requirements about the effects of different front important radar detecting areas on target integrated stealth performance.

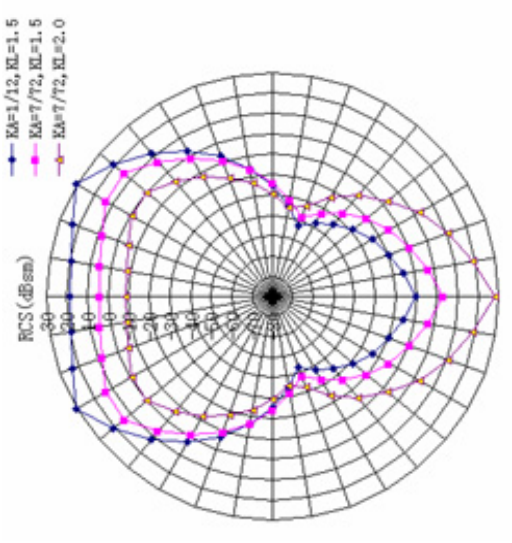


Fig. 14. RCS curves of different target RCS scattering models.

Fig.14 shows the models with various target circumferential RCS scattering characters and being built up by changing the values of K_L and K_A, when δ_{ave} is equal to -10dBsm.

4.1.2 The model building method

According to Fig.12, the detailed building steps of the new model are described in this section. First of all, defining the suitable target whole and local RCS scattering characteristic parameters and introducing several model RCS scattering controlling parameters according to the different target RCS scattering characters are necessary. So as it is shown from Fig.1, the average RCS value of target circumferential area and the heading direction within the angular region of -30^0 to $+30^0$ are introduced as the target entire and local RCS scattering characteristic parameters respectively. Their corresponding symbols are δ_{ave} and $\bar{\delta}_0$. K_L represents the ratio between long side and short side lengths. Secondly, different functions are used to describe the RCS curves of different radar detecting regions. The variable of curve function is angle value φ, its corresponding function value is length value R. But in the target RCS curve, the corresponding value of angle φ is the target RCS value. Therefore, a transform between the coordinate length value R and the target RCS value is needed. Based on this transform, RCS value in any direction of the target can be expressed by:

$$\delta_{(\varphi)} = \delta_{ave} + (R_{(\varphi)} - \frac{1}{N}\sum_{i=1}^{N} R_i) \times \frac{\delta_{ave} + \delta_{\min}}{\frac{1}{N}\sum_{i=1}^{N} R_i} \tag{4}$$

where $R_{(\varphi)}$ denotes the R value in any target azimuth, R_i is one of the series values of R in any target important radar detecting areas and subscript i represents the sequence of these R values, δ_{ave} is the average RCS value of target circumferential area or local radar detecting areas and $\delta_{\min}$ represents the RCS value of the coordinate origin. The stealth analysis about the target local RCS scattering character should be included in the conclusions of target integrated stealth performance analysis. So the last step is dividing the target RCS curve into several parts according to the target RCS scattering characters, then using different functions to describe these parts respectively. By this way, the analysis conclusion about how the target local stealth performance affects its integrated stealth ability can be reached. All the functions for every part of the curve can be written as:

$$\delta_{(\varphi)} = R_{(\varphi)} \times \frac{(\delta_{ave} + \delta_{\min})}{\frac{1}{N}\sum_{i=1}^{N} R_i} - \delta_{\min} \tag{5}$$

where $\delta_{i(\varphi)}$ and $R_{i(\varphi)}$ are the target RCS values and R values respectively and corresponding to the target azimuth of φ. The subscript i represents the sequence of target important radar detecting areas.

The functions corresponding to Fig.12 can be expressed as:

$$\delta_{i(\varphi)} = \begin{cases} (\delta_{ave} + \delta_{\min}) \times \dfrac{K_L}{2 \times \cos\varphi} \times \dfrac{91}{Sum} - \delta_{\min} \\ 0 \le \varphi \le \varphi_{Limit} \\ or(180 - \varphi_{Limit}) \le \varphi \le (180 + \varphi_{Limit}) \\ or(360 - \varphi_{Limit}) \le \varphi < 360 \\ (\delta_{ave} + \delta_{\min}) \times \dfrac{1}{2 \times \sin\varphi} \times \dfrac{91}{Sum} - \delta_{\min} \\ (180 + \varphi_{Limit}) < \varphi < (360 - \varphi_{Limit}) \\ or(360 - \varphi_{Limit}) \le \varphi < 360 \end{cases} \tag{6}$$

$$Sum = \sum_{\alpha=0}^{\varphi_{Limit}} \left| \frac{K_L}{2 \times \cos\alpha} \right| + \sum_{\alpha=\varphi_{Limit}+1}^{90} \left| \frac{1}{2 \times \sin\alpha} \right|$$

where δ_{ave} is the average RCS value of target circumferential area, $\delta_{\min}$ represents the RCS value of the coordinate origin, α is the radar detect angle and K_L is the model RCS scattering controlling parameter.

4.2 Examples and discussions

In this section, the new kind of target circumferential RCS scattering models will be built according to Fig.12 and Fig.13. Before that, disposing several radars in different azimuths of enemy base. Combing relevant dynamic models and integrated stealth analysis rules can give the detailed integrated stealth analysis conclusions. In these examples, aircraft flight altitude is 1000m and flight velocity is 500m/s. The aircraft carries out the penetration mission along a straight flight course at the azimuth of 90 degrees. The azimuths of these radars are 0, 30, 45, 60, 90, 120, 135 and 180 degrees respectively. A comparison is made between the integrated stealth performance of these two serial models are given blow.

4.2.1 Rectangular RCS scattering models

Combining the new modeling methods described in Section 4.1.2 and the relevant stealth performance analysis requirements, the serial models with absolutely different RCS scattering characters are modeled.

The values of relevant model RCS scattering parameters are δ_{ave} =-10dBsm and 0dBsm and K_L =0.5, 1.0 and 2.0 respectively. Fig.15 and Fig.16 compare the average radar detecting probability of these serial models. Fig.15 shows that when δ_{ave} is equal to -10dBsm, the RCS scattering characters of these three models differs much from each other. Among these models, the one corresponding to K_L =0.5 has the highest average radar detecting probability. When K_L =1.0, the corresponding model will have much lower average radar detecting probability. When K_L = 0.5, the radars located in the two sides of the enemy base will have much higher detecting probability than that for K_L =1.0. So when δ_{ave} is around -10dBsm, the condition of K_L =1.0 can make the penetration aircraft with rectangular RCS scattering character have excellent sidewise stealth ability. Then the aircraft can carry out

the penetration mission with a small transverse distance arriving at enemy base. The average radar detecting probability rises remarkably, when δ_{ave} goes up to 0dBsm.

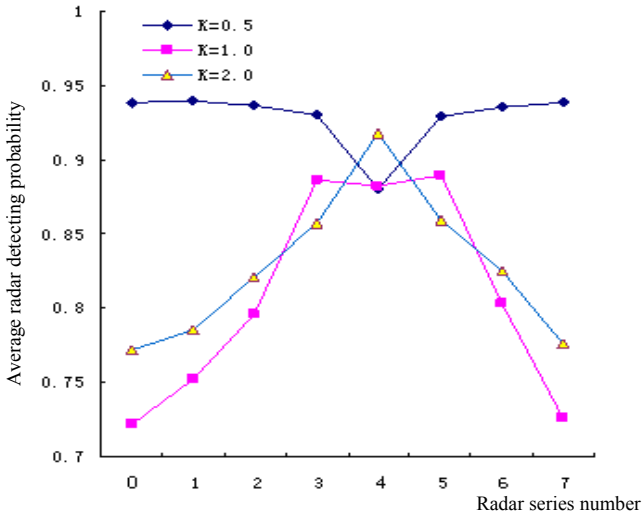


Fig. 15. Comparison of radar detecting probability corresponding to serial models (δ_{ave} =-10dBsm).

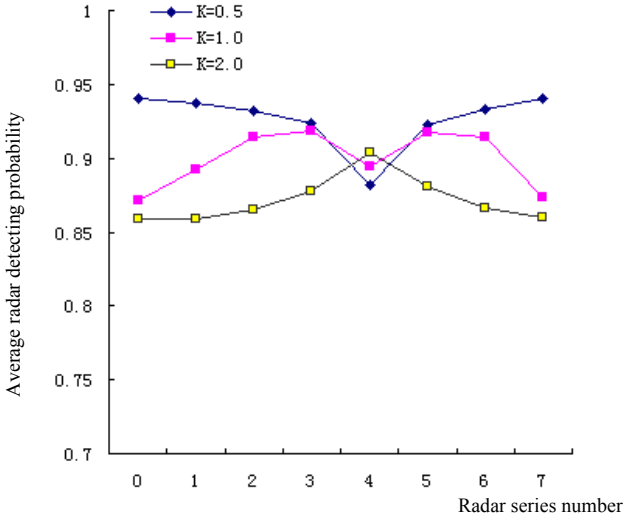


Fig. 16. Comparison of radar detecting probability corresponding to serial models (δ_{ave} =0dBsm).

K_L	Radar series number	Lost target number	T_{First}	Stable tracking number	Tracking duration (s)
0.5	0	1	289	1	315
0.5	1	1	201	1	291
0.5	2	1	174	1	263
0.5	3	1	161	1	225
0.5	4	1	167	1	159
0.5	5	1	162	1	223
0.5	6	1	175	1	261
0.5	7	1	290	1	313
1.0	0	2	194	1	139
				2	139
1.0	1	2	110	1	174
				2	174
1.0	2	1	87	1	436
1.0	3	1	79	1	389
1.0	4	1	89	1	315
1.0	5	1	81	1	385
1.0	6	1	89	1	433
1.0	7	2	196	1	141
				2	142
2.0	0	2	1	1	142
				2	300
2.0	1	2	1	1	105
				2	346
2.0	2	2	1	1	137
				2	399
2.0	3	2	1	1	197
				2	469
2.0	4	1	1	1	765
2.0	5	2	1	1	202
				2	474
2.0	6	2	1	1	142
				2	402
2.0	7	2	1	1	147
				2	304

Table 2. Relevant radar detecting data of serial models (δ_{ave} =-10dBsm)

Table 2 lists the radar detecting data of this serial models. The data shows that when δ_{ave} =-10dBsm, the lost target number will increase and T_{First} will decrease with the increase of K_L. Suppose that the aircraft will counter the threat of enemy firepower only when it is continuously found by radar more than a certain time threshold and the corresponding notation is T_{Limit}. If the time threshold is 200 seconds, when K_L =0.5, 1.0 and 2.0, the

corresponding valid stable track number are 7 , 5 and 8 respectively. So when K_L =1.0, the model has the least chances of meeting with enemy firepower.

K_L	Radar series number	Lost target number	T_{First}	Stable tracking number	Tracking duration
0.5	0	1	242	1	409
0.5	1	1	156	1	381
0.5	2	1	132	1	347
0.5	3	1	122	1	303
0.5	4	1	131	1	231
0.5	5	1	123	1	301
0.5	6	1	133	1	344
0.5	7	1	244	1	406
1.0	0	1	86	1	721
1.0	1	1	4	1	685
1.0	2	1	1	1	626
1.0	3	1	1	1	569
1.0	4	1	1	1	509
1.0	5	1	1	1	568
1.0	6	1	1	1	625
1.0	7	1	88	1	717
2.0	0	2	1	1	250
				2	966
2.0	1	2	1	1	201
				2	1003
2.0	2	1	1	1	221
				2	1049
2.0	3	1	1	1	1390
2.0	4	1	1	1	1347
2.0	5	1	1	1	1389
2.0	6	2	1	1	255
				2	1053
2.0	7	2	1	1	255
				2	970

Table 3. Relevant radar detecting data of serial models (δ_{ave} = 0dBsm)

The data in Table 3 shows that when δ_{ave} =0dBsm, the duration of aircraft being stable tracked rises obviously and the valid stable tracking number also increases. It is also learned from the data that the time of models first being found by radar corresponding to K_L =1.0 and 2.0 will be much earlier. Therefore, the data shows that the RCS scattering controlling parameter δ_{ave} has big effects on the integrated stealth performance of serial models.

4.2.2 Triangular RCS scattering models

The serial models with triangular RCS scattering characters are established according to the methods described in section 4.1.1. The meanings of δ_{ave}, K_L and K_A have been told in

section 4.1.1. The values of K_L and K_A are taken as K_L =1.5、2.0 and K_A = 1/12, 7/72 and 1/9 respectively. Fig.6 and Fig.7 show the comparison of average radar detecting probability between these serial models.

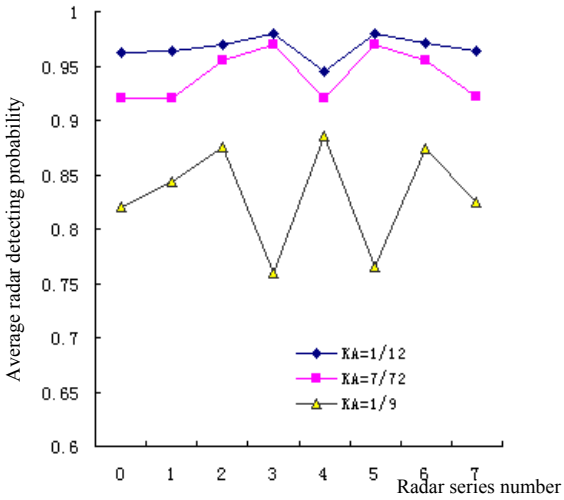


Fig. 17. Comparison of radar detecting probability of serial models (δ_{ave} =-10dBsm, K_L =1.5).

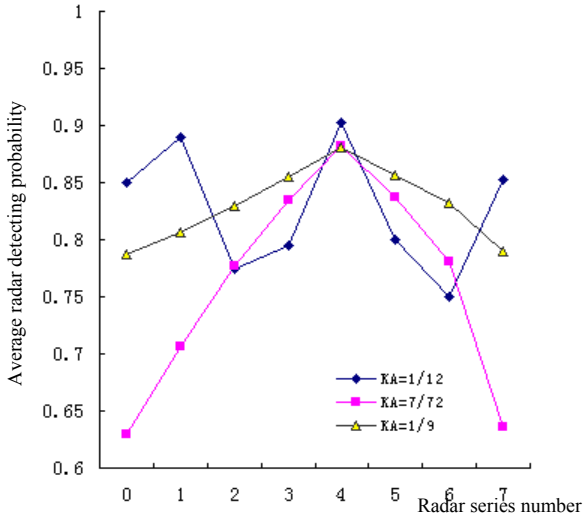


Fig. 18. Comparison of radar detecting probability of serial models (δ_{ave} =-10dBsm, K_L =2.0).

Fig.17 shows that when δ_{ave} =-10dBsm and K_L = 1.5, the model with RCS scattering parameters K_A =1/9 has lower radar detecting probability than both the models for K_A =1/12 and 7/72 respectively. The radar detecting probability curves in Fig.18 show that when K_L =2.0 and K_A adopts different values, the average radar detecting probability of all these serial models are decreased. Especially for the situation of K_A =7/72, the decreasing range is the maximum one. Both the two models corresponding to K_A = 7/72 and 1/9 respectively have lower radar detecting probability in their two sides than that in their heading direction. Especially when the RCS scattering controlling parameter K_A =7 /72.

K_L	Radar series number	Lost target number	T_{First}	Stable tracking number	Tracking duration
1.5	0	1	1	1	378
1.5	1	1	1	1	299
1.5	2	1	1	1	279
1.5	3	1	1	1	269
1.5	4	1	1	1	684
1.5	5	1	1	1	270
1.5	6	1	1	1	281
1.5	7	1	1	1	380
2.0	0	2	120	1	161
				2	773
2.0	1	2	37	1	197
				2	1020
2.0	2	2	16	1	224
				2	1244
2.0	3	2	9	1	243
				2	1476
2.0	4	1	17	1	2097
2.0	5	2	10	1	243
				2	1489
2.0	6	2	17	1	226
				2	1259
2.0	7	2	121	1	165
				2	799

Table 4. Relevant radar detecting data of series models（δ_{ave} =-10dBsm, K_A =7/72）

Table 4 lists the relevant radar detecting data corresponding to the situation for the RCS scattering controlling parameters δ_{ave} = -10dBsm and K_A =7/72 and K_L adopts different values.

The data in Table 4 shows that when δ_{ave} = -10 dBsm, the models with RCS scattering parameter K_L = 2.0 will have more chances of being tracked by radar and the duration of radar stable tracking will be longer than the models corresponding to K_L =1.5.

5. The effects of different aircraft RCS curve peak characters

Analyzing rules about effects of the aircraft RCS curve peaks with different shape on penetration stealth performance reasonably, is meaningful. These rules could be references for both the concept penetration aircraft's stealth configuration design and making effective aircraft RCS reduction plans.

5.1 The target RCS curve peak numerical simulation model

5.1.1 The relevant model parameters

The RCS curve peaks on the target circumferential RCS curve are formed by the total RCS of aircraft scattering sources, which are located in different radar detecting directions. The RCS curve peaks with different shape would cause different effects on the aircraft penetration stealth performance.

There are two requirements for the target RCS curve peak numerical simulation model: (1)The changing trends of RCS curve peak scattering characteristic can be controlled (2) The analysis requirements about effects on penetration dynamic stealth ability, which are caused by kinds of RCS curve peak scattering characteristic changes, can be satisfied.

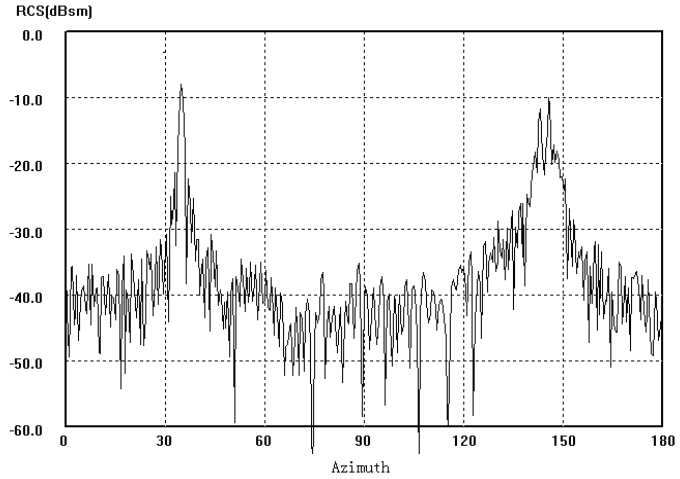


Fig. 19. The RCS curve peaks with different shape

The two RCS curve peaks with different shape in Fig19 show that RCS curve peak value and RCS curve peak azimuth-width are the main factors, which control the changing trends of RCS curve peaks. Changing the number or azimuth of aircraft scattering sources both can alter the RCS curve peak value or RCS curve peak azimuth-width. Rules about effects on the penetration dynamic stealth ability, which are caused by changes of the RCS curve peak value and RCS curve peak azimuth-width, could be references for reasonable stealth configuration design. As mentioned above, it defines two target RCS curve peak numerical simulation model parameters to control the changing trends of model's scattering characteristic, one is the peak value controlling parameter, the symbol is δ_{Max} . The other is the peak azimuth-width controlling parameter, the sign is φ_{Width} . This article also defines a

set of RCS curve peak location parameters, corresponding signs are φ_i (i=0,1, 2...), the subscript i represents the sequence of RCS curve peaks, which are located on the target circumferential RCS curve. As analyzing the rules about effects on aircraft penetration stealth ability, which are caused by changes of RCS curve peak scattering characteristic, influences of target scattering characteristic in another aircraft azimuth, should be eliminated. This article sets the RCS value in another model azimuth, as a constant value and less than the RCS peak value. As mentioned above, this article introduces the scattering characteristic controlling parameter, which is corresponding to another model azimuth, the symbol is δ_φ.

5.1.2 The model building method

Defining the model parameters according to section 5.1.1, the detailed model building method is told as below: (1) Defining the variable model parameters, according to the analysis requirements about influences on the aircraft penetration stealth ability, which are caused by different RCS curve peak scattering characteristic. For example, the RCS curve peak value would be reduced by the decrease of aircraft scattering sources. For analyzing the corresponding effects on the penetration stealth ability, it can build up a series of models with different RCS curve peak value, by changing the RCS curve peak value controlling parameters $\delta_{(i)Max}$ (i=0,1,2...), the subscript i represents the sequence of RCS curve peaks on the target circumferential RCS curve. Therefore, it can define $\delta_{(i)Max}$ as the variable model parameters. (2) Introducing the scattering characteristic controlling parameters and RCS curve peak location parameters into the model, then determining the value of these parameters. For example, according to the case that there are two RCS curve peaks with 6°azimuth-width, which are located on -30°and 30°azimuth of aircraft front area respectively, it can define the RCS curve peak azimuth-width controlling parameter: $\varphi_{(i)Width}$ =6°(i=0,1) and the RCS curve peak location parameters: φ_i (i=0,1), $\varphi_0 = 30^0$, $\varphi_1 = -30^0$. (3) Fixing up the value of scattering characteristic controlling parameter (δ_φ), which is corresponding to another model azimuth.

5.2 Examples and discussion

This section builds up a series of target RCS curve peak numerical simulation models with different peak value or azimuth-width. By testing and analyzing these models, influences of RCS curve peak value or RCS curve peak azimuth-width on the dynamic stealth ability are generalized.

5.2.1 Influences of RCS curve peak characters

For evaluating effects on the integrated stealth performance of penetration aircraft, caused by different target RCS curve peak characters. According to Fig.19, this section builds up series of models, with same circumferential RCS characters, but different RCS curve peak characters. The circumferential RCS character controlling parameters of these models are: δ_{ave} =-10dBsm, K_L =1.0, and the RCS curve peak azimuth-location parameters are: φ_0 =45°, φ_1 =135°, φ_2 =225° and φ_3 =315°. Table 1.lists another RCS curve peak character controlling factors.

	$\delta_{j(Max)}$ (j=0,1,2,3)	$\varphi_{j(Width)}$ (j=0,1,2,3)
Model 1	10 dBsm	2°
Model 2	10 dBsm	4°
Model 3	10 dBsm	6°
Model 4	20 dBsm	2°
Model 5	20 dBsm	4°
Model 6	20 dBsm	6°
Model 7	30 dBsm	2°
Model 8	30 dBsm	4°
Model 9	30 dBsm	6°

Table 5. The RCS curve peak controlling factors

Simulation conditions of this example are:
flight condition: flight altitude H=1000 m; flight velocity V=200 m/s, flight azimuth (relative to enemy base) φ_f = -40, -20, 0, 20 and 40 degrees; distance to penetration destination L=400Km.
condition of radar network : Table 2 lists positions of every ground to air radar (GR) in the network and location of the penetration destination (Basement).

Name	Longitude	Latitude
GR1	119.6330	23.5667
GR2	121.550	24.0667
GR3	120.4830	22.7
GR4	121.5330	25.0330
GR5	121.6170	24.0167
GR6	121.05	25.0667
GR7	121.9667	24.8
Basement	121.5	25.0

Table 6. Positions of radar and destination

Table7 lists the radar network detecting results of this serial models, corresponding to φ_f = -20 degree.

From Table7 we find that, φ_{Width} and δ_{Max} have different effects on the integrated stealth performance of penetration aircraft. Firstly, if φ_{Width} takes a fixed value, with the rise of δ_{Max}, T_{First} will decrease and target "flashing signal" will weaken ($T_{f(Find)}$ increases and $T_{(l)Lose}$ decreases). Furthermore, when φ_{Width} has a bigger fixed value, effects caused by increase of δ_{Max} would be more obvious. Adversely, if φ_{Width} takes a small value, namely, the azimuth-width of target RCS curve peak is narrow, δ_{Max} has little effect on the integrated stealth performance of penetration aircraft. Secondly, if δ_{Max} takes a fixed value,

T_{First} will decrease and the target "flashing signal" will weaken, with increase of φ_{Width}. Similarly, when δ_{Max} has a bigger fixed value, effects caused by increase of φ_{Width} would be more obvious. Lastly, when one of δ_{Max} and φ_{Width} takes a big fixed value, the other increases, effect on putting off T_{First}, caused by increase of φ_{Width} is more obvious.

	T_{First}	$N_{(f)Find}$	$T_{f(Find)}$	$N_{(l)Lose}$	$T_{(l)Lose}$
Model 1	209.4	1	150.5	1	559.1
		2	281		
Model 2	173.6	1	222.3	1	523.1
		2	281		
Model 3	173.6	1	258.3	1	475.1
		2	293		
Model 4	209.4	1	186.5	1	523.1
		2	281		
Model 5	173.6	1	258.3	1	487.1
		2	281		
Model 6	137.7	1	330.2	1	427.1
		2	305		
Model 7	209.4	1	186.5	1	523.1
		2	281		
Model 8	137.7	1	294.2	1	487.1
		2	281		
Model 9	101.8	1	366.1	1	427.1
		2	305		

Table 7. Radar network detecting results of serial models

The rules for reducing aircraft scattering sources and improving integrated stealth performance of penetration aircraft can be got from above conclusions:

1. reducing δ_{Max} firstly for putting off T_{First}.
2. plan about strengthening target "flashing signal" is: it should reduce δ_{Max} firstly and φ_{Width} secondly, when δ_{Max} takes a big value. However, when δ_{Max} is small, it should reduce φ_{Width} firstly.
3. there is no need to reduce δ_{Max} too much, if azimuth-width of RCS curve peak is narrow.

5.2.2 Influences of superposing RCS curve peaks

Because of the stealth design plan: "parallel leading-edge" and "parallel trailing-edge", the RCS curve peaks created by wing edges, could be located on the same azimuth. For evaluating effects of the plan, this section builds up two models, corresponding to "before superposing" and "after superposing" respectively (see Fig.10 and Fig.11). Circumferential RCS controlling parameters of these two models are: δ_{Max} = -20dBsm, K_L =1.0, and the RCS curve peak controlling factors are:

model one: $\varphi_0 = 57^0$, $\varphi_1 = 63^0$, $\varphi_2 = 180^0$, $\varphi_3 = 297^0$,
$\varphi_4 = 303^0$; $\varphi_{j(Width)} = 2^0$; $\delta_{j(Max)}$ =10dBsm, (j=0,1,2,3,4)

model two: $\varphi_0 = 60^0$, $\varphi_1 = 180^0$, $\varphi_2 = 300^0$; $\delta_{1(Max)}$ =10dBsm, $\delta_{0(Max)} = \delta_{2(Max)}$ =13.01dBsm;
$\varphi_{j(Width)} = 2^0$ (j=0,1,2).

Simulation conditions are the same as the example in section 5.2.1

Table8 lists the radar network detecting results corresponding to $\varphi_f = -20^0$, 0^0 and 20^0, respectively. These results show that, due to superposing RCS curve peaks, not only T_{First} is reduced a little, but also the target "flashing signal" is strengthened. Namely, the integrated stealth performance of penetration aircraft can be enhanced by superposing RCS curve peaks.

φ_f		T_{First}	$N_{(f)Find}$	$T_{f(Find)}$	$N_{(l)Lose}$	$T_{(l)Lose}$
-20°	Model 1	742.4	1	197.8	1	202.1
			2	57.7		
	Model 2	740	1	195.9	1	243.9
			2	20.2		
0°	Model 1	284.2	1	435.7	1	67.9
			2	196.1	2	92.3
			3	123.8		
	Model 2	230.2	1	363.5	1	218
			2	172.2	2	92.3
			3	123.8		
20°	Model 1	250.1	1	325.8	1	57.5
			2	566.6		
	Model 2	250.1	1	325.8	1	57.5
			2	566.6		

Table 8. Radar network detecting results of two models

6. Conclusions

The conclusions about integrated stealth performance of penetration aircraft, based on a reasonable assessment method, are meaningful for both new type of stealth penetration aircraft design and efficient penetration tactics making. This paper proposes a new integrated stealth performance assessment method for penetration aircraft, which constitutes of the target RCS numerical simulation model and dynamic assessment model. This method can assesses effects on the integrated stealth performance of penetration aircraft, caused by different target circumferential RCS characters and RCS curve peak characters. As a fruit, three examples in this article show that:

1. The target RCS numerical simulation models, performing different circumferential RCS characters or RCS curve peak characters, could satisfy various assessment requirements for the integrated stealth performance of aircraft.
2. Better configurations, which can improve integrated stealth performance of penetration aircraft, could be got from the laws for effects of different target RCS characters.

As a new method, it can be applied in the integrated stealth performance of aircraft assessment field. To get more credible assessment conclusions, the more accurate target RCS numerical simulation model and dynamic assessment model should be introduced.

7. References

Knott E.F.(1987). Radar cross section. Beijing: Publishing House of Electronics Industry, pp. 410-416 ISBN 7-5050

Wang Lingxiao, Zhao Deyun.(2009). Effective Path Planning Method for Low Detectable Aircraft. Journal of Systems Engineering and Electronics. Vol.20, No.4, pp.784- 785

John F. Dishman, Edward R. Beadle. SEVR: A LPD Metric for a 3-D Battle Space. 2007 IEEE Millitary Communication Conference. Orlando Oct 2007.

Guo Zhengxin , Jiang Jin. (2003). Computation Model of Radar Detection Probability With Various Distance.Journal of Air Force Radar Academy. Vol17, No4. pp.7-8

Yuan Qilun, Tan Shaoxian.(2002). A method of establishing model of radar detection probability. Morden Defence Technology. Vol.30, No.3, pp.1-3

Zhang Kao, Zhang Yunfei, Ma Dongli.(1997). The aircraft stealth Performance Analysis and Calculating Methods. Beijing: National Defence Industry Press, pp.85- 92. ISBN 7-118-02718-9

Wang Xu, Song Bifeng.(2008). Analytic Model of Aircraft Survivability detection Time for One-to-one Engagement Acta Aeronautical Et Astronautica Sinica. Vol.29, No.4, pp. 915-916.

Paterson J.(1999) Overview of Low Observable Technology and Its Effects on Combat Aircraft Survivability. Journal of Aircraft., Vol. 36, No.2, pp. 380-388.

Pei Y, Song B.(2003). Solving the Combinatorial Explosion ProblemWhen Calculating the Multiple-Hit Vulnerability of Aircraft. Journal of Aircraft, Vol.43, No.4, pp.1190-1194.

Li GuangMing, Tang Yemin, Jiang SuRong.(2006). Performance Evaluation of Radar Network for Counterchecking Stealth Aircraft-Comprehensive Detecting Probability of Radar Network. Modern Radar, Vol.28, No.1.pp.23-25

Li ShouAn, Zhang HengXi, Li ShuLin, Meng Ke.(2005). Research on Aircraft Survability Evaluation and Synthetic Trade off Method. Acta Aeronautical Et Astronautica Sinica, Vol. 26, No.1, pp.23-26.

11

Aircraft Gas-Turbine Engine's Control Based on the Fuel Injection Control

Alexandru-Nicolae Tudosie
University of Craiova, Avionics Department
Romania

1. Introduction

For a gas-turbine engine, particularly for a jet engine, the automatic control is one of the most important aspects, in order to assure to it, as aircraft's main part, an appropriate operational safety and highest reliability; some specific hydro-mechanical or electro-mechanical controller currently realizes this purpose.
Jet engines for aircraft are built in a large range of performances and types (single spool, two spools or multiple spools, single jet or twin jet, with constant or with variable exhaust nozzle's geometry, with or without afterburning), depending on their specific tasks (engines for civil or for combat aircraft). Whatever the engine's constructive solution might be, it is compulsory that an automatic control system assist it, in order to achieve the desired performance and safety level, for any flight regime (altitude and speed).
Regarding the nowadays aircraft engine, the more complex their constructive solution is, the bigger the number of their parameters is. Considering an engine as a controlled object (see figure 1.a), one has to identify among these parameters the most important of them, the easiest to be measured and, in the mean time, to separate them in two classes: control parameters and controlled parameters. There is a multitude of eligible controlled parameters (output parameters, such as: thrust, fuel consumption, spool(s) speed, combustor's temperature etc.), but only a few eligible control parameters (input parameters, such as: fuel flow rate, nozzle's exit area and/or inlet's area). It results a great number of possible combinations of control programs (command laws) connecting the input and the output parameters, in order to make the engine a safe-operating aircraft part; for a human user (a pilot) it is impossible to assure an appropriate co-ordination of these multiples command laws, so it is compulsory to use some specific automatic control systems (controllers) to keep the output parameters in the desired range, whatever the flight conditions are.

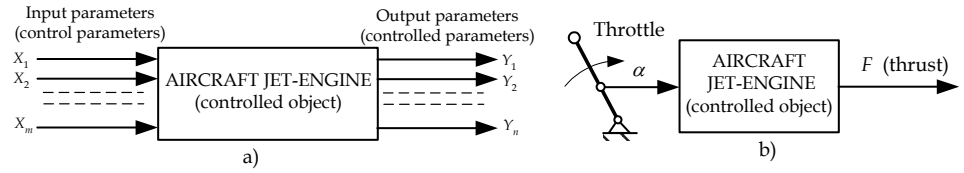


Fig. 1. Aircraft engines as controlled objects

In fact, the pilot has only a single engine's command possibility, a single input parameter - the throttle displacement and a single relevant output parameter – the engine's thrust (as shown in figure 1.b). Although, the engine's thrust is difficult to be measured and displayed, but it could be estimated and expressed by other parameters, such as engine's spool(s) speed(s) or gas temperature behind the turbine, which are measured and displayed much easier.

Consequently, most of aircraft engine command laws and programs are using as control parameters the fuel flow rate Q_c (which is the most important and the most used) and the exhaust nozzle throat and/or exit area A_5 and as controlled parameters the engine's spool (s) speed(s) and/or the engine's exhaust burned gas temperature. Meanwhile, in an engine control scheme, throttle's displacement becomes itself the input for a mixed (complex) setting block, which establishes the reference parameters for the engine's controller(s), as shown in figure 2. So, in this case, both engine's control parameters become themselves controlled parameters of the engine's controller(s), a complex engine control system having as sub-systems an exhaust nozzle exit area control system (Aron & Tudosie, 2001) as well as a fuel injection control system (Lungu & Tudosie, 1997).

Because of the fuel injection great importance, fuel injection controllers' issue is the main concern for pump designers and manufacturers.

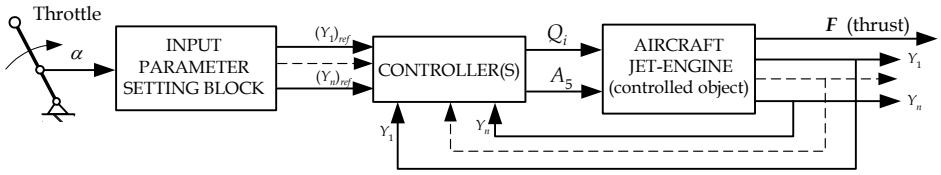


Fig. 2. Aircraft engine's automatic control system

2. Principles of the fuel flow rate control

Aircraft engines' fuel supply is assured by different type of pumps: with plungers, with pinions (toothed wheels), or with impeller. For all of them, the output fuel flow rate depends on their rotor speed and on their actuator's position; for the pump with plungers the actuator gives the plate's cline angle, but for the other pump type the actuator determines the by-pass slide-valve position (which gives the size of the discharge orifice and consequently the amount of the discharged fuel flow rate, as well as the fuel pressure).

The fuel flow rate through some injection scheme part (x) Q_x is given by the generic formula (where μ_x is the x-part flow co-efficient, depending on its inner channel shape and roughness, A_x – x-part injection effective area, ρ – fuel density, p_b – pressure before and p_a – pressure after the above-mentioned part):

$$Q_x = \mu_x A_x \sqrt{\frac{2}{\rho}} \sqrt{p_b - p_a} \,. \tag{1}$$

Consequently, the fuel flow rate through the injector(s) is

$$Q_i = \mu_i A_i \sqrt{\frac{2}{\rho}} \sqrt{p_i - p_{CA}} \,, \tag{2}$$

and through the dosage valve (before the injector) is, similarly,

$$Q_d = \mu_d A_d \sqrt{\frac{2}{\rho}} \sqrt{p_p - p_i} \,, \tag{3}$$

where μ_i and μ_r are the flow co-efficient, p_i – fuel injection pressure, p_{CA} – air/gases pressure in the engine's combustor, p_p – fuel pump supplying pressure, A_d – dosage valve's effective area (depending on the pump's actuator displacement), A_i – injector's effective area, assumed as circular $\left(A_i = \frac{\pi d_i^2}{4} \right)$, d_i – injector's diameter.

For a steady state regime the fuel flow rate is constant, so $Q_i = Q_r$, which leads to a new expression for the injected flow rate (Stoenciu, 1986), where $f = \frac{\mu_i A_i}{\mu_d A_d}$:

$$Q_r = \mu_d A_d \sqrt{\frac{2}{\rho(1+f)}} \sqrt{p_p - p_{CA}} \,. \tag{4}$$

As far as, for a constant engine operating regime p_{CA} can be assumed as constant, the fuel flow rate depends on p_p (or $p_p - p_i$) and A_r, therefore a fuel flow rate controller has to deal with one of these parameters, or with both of them simultaneously.

Nowadays common use basic fuel injection controllers are built, according to this observation, as following types (Stoicescu & Rotaru, 1999):

- with constant fuel pressure and adjustable fuel dosage valve;
- with constant fuel differential pressure and adjustable fuel dosage valve;
- with constant injector flow areas and adjustable fuel differential pressure.

Usually, the fuel pumps are integrated in the jet engine's control system; more precisely: the fuel pump is spinned by the engine's shaft (obviously, through a gear box), so the pump speed is proportional (sometimes equal) to the engine's speed, which is the engine's most frequently controlled parameter. So, the other pump control parameter (the plate angle or the discharge orifice width) must be commanded by the engine's speed controller.

Most of nowadays used aircraft jet engine controllers have as controlled parameter the engine's speed, using the fuel injection as control parameter, while the gases temperature is only a limited parameter; temperature limitation is realized through the same control parameter – the injection fuel flow rate (Moir & Seabridge, 2008; Jaw & Matingly, 2009). Consequently, a commanded fuel flow rate decrease, in order to cancel a temperature override, induces also a speed decrease.

3. Fuel injection controller with constant pressure chamber

A very simple but efficient fuel injection control constructive solution includes a fuel pump with constant pressure chamber in a control scheme for the engine's speed or exhaust gases

temperature. As far as the most important aircraft jet engine performance is the thrust level and engine's speed value is the most effective mode to estimate it, engine speed control becomes a priority.

Figure 3 presents a hydro-mechanical fuel injection control system, based on a fuel pump with plungers and constant pressure chamber.

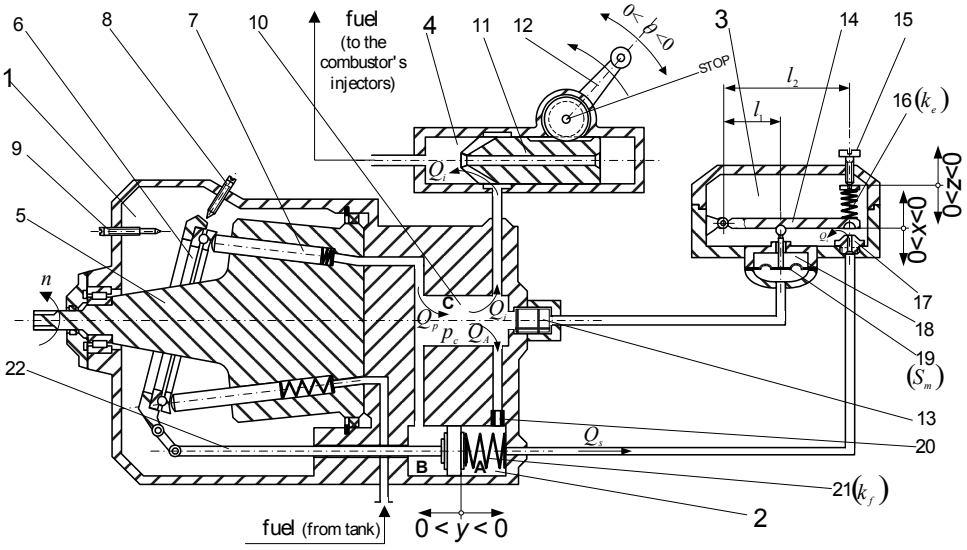


Fig. 3. Fuel injection controller with constant pressure chamber

3.1 System presentation

This type of fuel injection controller assures the requested fuel flow rate adjusting the dosage valve effective area, while the fuel pressure before it is kept constant.

Main parts of the system are: 1-fuel pump with plungers; 2-pump's actuator; 3-pressure sensor with nozzle-flap system; 4-dosage valve (dosing element). The fuel pump delivers a Q_p fuel flow rate, at a p_c pressure in a pressure chamber 10, which supplies the injector ramp through a dosage valve. This dosage valve slide 11 operates proportionally to the throttle's displacement, being moved by the lever 12. The pump is connected to the engine shaft, so its speed is n, or proportional to it. Pump 6 plate's angle is established by the actuator's rod 22-displacement y, given by the balance of the pressures in the actuator's chambers (A and B) and the 21 spring's elastic force. The pressure p_A in chamber A is given by the balance between the fuel flow rates through the drossel 20 and the nozzle 17 (covered by the semi-spherical flap, attached to the sensor's lever 14). The balance between two mechanical moments establishes the sensor lever's displacement x: the one given by the elastic force of the spring 16 (due to its z pre-compression) and the one given by the elastic force of the membrane 19 (displaced by the pressure in chamber, between the membrane and the fluid oscillations buffer 13).

The system operates by keeping a constant pressure in chamber 10, equal to the preset value (proportional to the spring 16 pre-compression, set by the adjuster bolt 15). The engine's necessary fuel flow rate Q_i and, consequently, the engine's speed n, are controlled by the co-relation between the p_c pressure's value and the dosage valve's variable slot (proportional to the lever's angular displacement θ).
An operational block diagram of the control system is presented in figure 4.

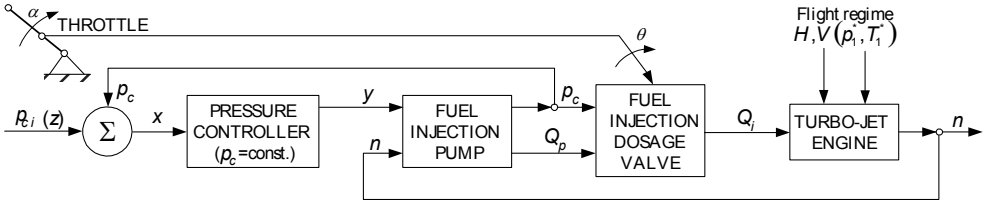


Fig. 4. Constant pressure chamber controller's operational block diagram

3.2 System mathematical model

The mathematical model consists of the motion equations for each sub-system, as follows:

a. fuel pump flow rate equation

$$Q_p = Q_p(n, y)\,, \tag{5}$$

b. constant pressure chamber equation

$$Q_i = Q_p - Q_A\,, \tag{6}$$

c. fuel pump actuator equations

$$Q_A = \mu_{dA}\frac{\pi d_A^2}{4}\sqrt{\frac{2}{\rho}}\sqrt{p_c - p_A}\,, \tag{7}$$

$$Q_A - Q_s = \beta V_{A0}\frac{\mathrm{d}p_A}{\mathrm{d}t} + S_A\frac{\mathrm{d}}{\mathrm{d}t}\left(y_s + y\right), \tag{8}$$

$$m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \xi\frac{\mathrm{d}y}{\mathrm{d}t} + k_f\left(y_s + y\right) = S_B p_c - S_A p_A\,, \tag{9}$$

d. pressure sensor equations

$$Q_s = -\mu_n \pi d_n\sqrt{\frac{2}{\rho}}(z + x)\sqrt{p_A - p_0}\,, \tag{10}$$

$$l_1 S_m p_c + l_2\frac{\pi d_n^2}{4}p_A = l_2 k_e(z + x)\,, \tag{11}$$

e. dosing valve equation

$$Q_i = \mu_i b_1 \frac{\theta_s + \theta}{\pi} \sqrt{\frac{2}{\rho}} \sqrt{p_c - p_{CA}}\,, \tag{12}$$

f. jet engine's equation (considering its speed n as controlled parameter)

$$n = n\left(Q_i, p_1^*, T_1^*\right), \tag{13}$$

where Q_p, Q_i, Q_A, Q_s are fuel flow rates, p_c -pump's chamber's pressure, p_A - actuator's A chamber's pressure, p_{CA} -combustor's internal pressure, p_0 -low pressure's circuit's pressure, μ_{dA}, μ_n, μ_i -flow rate co-efficient, d_A, d_n -drossels' diameters, S_A, S_B -piston's surfaces, $S_A \approx S_B$, S_m -sensor's elastic membrane's surface, k_f, k_e -spring elastic constants, V_{A0} -actuator's A chamber's volume, β -fuel's compressibility co-efficient, ξ -viscous friction co-efficient, m-actuator's mobile ensemble's mass, θ -dosing valve's lever's angular displacement (which is proportional to the throttle's displacement), x-sensor's lever's displacement, z-sensor's spring preset, y-actuator's rod's displacement, p_1^*, T_1^* -engine's inlet's parameters (total pressure and total temperature).

It's obviously, the above-presented equations are non-linear and, in order to use them for system's studying, one has to transform them into linear equations.

Assuming the small-disturbances hypothesis, one can obtain a linear form of the model; so, assuming that each X parameter can be expressed as

$$X = X_0 + \frac{\Delta X}{1!} + \frac{(\Delta X)^2}{2!} + \ldots + \frac{(\Delta X)^n}{n!}, \tag{14}$$

(where X_0 is the steady state regime's X-value and ΔX -deviation or static error) and neglecting the terms which contains $(\Delta X)^r, r \geq 2$, applying the finite differences method, one obtains a new form of the equation system, particularly in the neighborhood of a steady state operating regime (method described in Lungu, 2000, Stoenciu, 1986), as follows:

$$\Delta Q_A = k_A\left(\Delta p_c - \Delta p_A\right), \tag{15}$$

$$\Delta Q_i = k_{i\theta}\Delta\theta + k_{ic}\Delta p_c, \tag{16}$$

$$\Delta Q_s = k_{SA}\Delta p_A - k_s\Delta x - k_s\Delta z, \tag{17}$$

$$\Delta Q_i = \Delta Q_p - \Delta Q_A, \tag{18}$$

$$\Delta Q_A - \Delta Q_s = \beta V_{A0}\frac{\mathrm{d}}{\mathrm{d}t}\Delta p_A + S_A\frac{\mathrm{d}}{\mathrm{d}t}\Delta y, \tag{19}$$

$$k_e\left(\Delta x + \Delta z\right) = \frac{l_1}{l_2}S_m\Delta p_c, \tag{20}$$

$$\frac{k_f}{S_A}\left(\frac{m}{k_e}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Delta y + \frac{\xi}{k_e}\frac{\mathrm{d}}{\mathrm{d}t}\Delta y + 1\right) = \Delta p_c - \Delta p_A, \tag{21}$$

where the above used annotations are

$$k_A = \mu_{dA}\frac{\pi\sqrt{2}d_A^2}{8\sqrt{\rho\left(p_{c0}-p_{A0}\right)}},\ k_{i\theta} = \frac{\mu_i b_1\sqrt{2\rho p_{c0}}}{\pi\rho},\ k_{SA} = \frac{\mu_n\pi d_n\left(x_0+z_0\right)\sqrt{2\rho p_{A0}}}{2\rho p_{A0}},$$

$$k_{ic} = \frac{\mu_i b_1\left(\theta_s+\theta_0\right)\sqrt{2\rho p_{c0}}}{2\pi\rho p_{c0}},\ k_s = -\frac{\mu_n\pi d_n\sqrt{2\rho p_{A0}}}{\rho}. \tag{22}$$

Using, also, the generic annotation $\overline{X} = \frac{\Delta X}{X_0}$, the above-determined mathematical model can be transformed in a non-dimensional one. After applying the Laplace transformer, one obtains the non-dimensional linearised mathematical model, as follows

$$k_{PA}\left(\tau_A \mathrm{s}+1\right)\overline{p_A} + \tau_y \mathrm{s}\overline{y} + k_{cx}\overline{x} + k_{cz}\overline{z} = \overline{p_c}, \tag{23}$$

$$k_{cx}\overline{x} + k_{cz}\overline{z} = k_{zxc}\overline{p_c}, \tag{24}$$

$$k_{Ay}\left(T_y^2\mathrm{s}^2 + 2\omega_0 T_y \mathrm{s}+1\right)\overline{y} = k_{AC}\overline{p_c} - \overline{p_A}, \tag{25}$$

$$k_\theta\overline{\theta} + k_{pc}\overline{p_c} - k_{py}\overline{y} = \overline{p_A}, \tag{26}$$

$$k_{cQ}\overline{p_c} - k_{\theta Q}\overline{\theta} = \overline{Q_i}. \tag{27}$$

For the complete control system determination, the fuel pump equation (for Q_p) and the jet engine equation for n (Stoicescu & Rotaru, 1999) must be added. One has considered that the engine is a single-jet single-spool one and its fuel pump is spinned by its shaft; therefore, the linearised non-dimensional mathematical model (equations 23÷27) should be completed by

$$\overline{Q_p} = k_{pn}\overline{n} + k_{py}\overline{y}, \tag{28}$$

$$\left(\tau_M \mathrm{s}+1\right)\overline{n} = k_c\overline{Q_i} + k_{HV}\overline{p_1^*}. \tag{29}$$

For the (23)÷(29) equation system the used co-efficient expressions are

$$k_{AC} = \frac{p_{c0}}{p_{A0}},\ k_{PA} = \frac{\left(k_A+k_{SA}\right)}{k_A k_{AC}},\ k_{cx} = \frac{k_s x_0}{p_{c0}},\ k_{cz} = \frac{k_s z_0}{p_{c0}},\ \tau_A = \frac{\beta V_{A0}}{k_A+k_{SA}},\ \tau_y = \frac{S_A y_0}{k_A p_{c0}},\ k_{Ay} = \frac{k_e y_0}{S_A p_{c0}},$$

$$T_y = \sqrt{\frac{m}{k_e}},\ \omega_0 = \frac{\xi}{2T_y k_e},\ k_{zxc} = \frac{k_s S_m p_{c0}}{k_e}\frac{l_1}{l_2},\ k_\theta = \frac{k_{i\theta}\theta_0}{k_A p_{A0}},$$

$$k_{pc} = \frac{\left(k_A+k_{ic}\right)}{k_A k_{AC}},\ k_{Qp} = \frac{Q_{p0}}{k_A p_{A0}},\ k_{\theta Q} = \frac{k_{i\theta}\theta_0}{Q_{i0}},\ k_{cQ} = \frac{k_{ic}p_{c0}}{Q_{i0}} = k_c. \tag{30}$$

Based on some practical observation, a few supplementary hypotheses could be involved (Abraham, 1986). Thus, the fuel is a non-compressible fluid, so $\beta = 0$; the inertial effects are very small, as well as the viscous friction, so the terms containing m and ξ are becoming null. The fuel flow rate through the actuator Q_A is very small, comparative to the combustor's fuel flow rate Q_i, so $Q_p \approx Q_i$. Consequently, the new, simplified, mathematical model equations are:

- for the pressure sensor:

$$\overline{x} = k_l \overline{p_c} + k_z \overline{z}\,, \tag{31}$$

where

$$k_l = \frac{p_{c0}}{x_0}\left(\frac{\partial x}{\partial p_c}\right)_0 = \frac{p_{c0}}{x_0}\frac{l_1}{l_2}\frac{S_m}{k_e},\ k_z = \frac{z_0}{x_0}\left(\frac{\partial x}{\partial z}\right)_0, \tag{32}$$

or, considering that the imposed, preset value of p_c is $\overline{p_{ci}} = \frac{k_z}{k_l}\overline{z} = \frac{z_0}{p_{c0}}\left(\frac{\partial p_c}{\partial z}\right)_0$, one obtains

$$\overline{x} = -k_l\left(\overline{p_{ci}} - \overline{p_c}\right); \tag{33}$$

- for the actuator:

$$\left(\tau_y \mathrm{s} + 1\right)\overline{y} = -k_x \overline{x}\,, \tag{34}$$

where

$$\tau_y = \frac{S_A}{\left(\frac{\partial Q_s}{\partial y}\right)_0 - \left(\frac{\partial Q_A}{\partial y}\right)_0} = \frac{4S_A\sqrt{k_f y_0\left(p_{c0} - k_f y_0\right)}}{k_f \pi\left(2\mu_n d_n x_0 - \mu_{dA} d_A^2\right)},\ p_{A0} = p_{c0} - k_f y_0\,,$$

$$k_x = \frac{\frac{y_0}{x_0}\left(\frac{\partial Q_s}{\partial x}\right)_0}{\left(\frac{\partial Q_s}{\partial y}\right)_0 - \left(\frac{\partial Q_A}{\partial y}\right)_0} = \frac{y_0}{x_0}\frac{4\mu_n d_n p_{A0}\sqrt{k_f y_0}}{2\mu_n d_n x_0 - \mu_{dA} d_A^2}\,. \tag{35}$$

Simplified mathematical model's new form becomes

$$\left(\tau_M \mathrm{s} + 1\right)\overline{n} = k_c \overline{Q_i} + k_{HV}\overline{p_1^*}\,, \tag{36}$$

$$\overline{Q_i} \equiv \overline{Q_p} = k_{pn}\overline{n} + k_{py}\overline{y}\,, \tag{37}$$

$$\overline{x} = -k_l\left(\overline{p_{ci}} - \overline{p_c}\right); \tag{38}$$

$$\left(\tau_y \mathrm{s}+1\right)\overline{y}=-k_x\overline{x}\,, \tag{39}$$

$$\overline{p_c}=\frac{1}{k_p}\overline{Q_i}-\frac{k_\theta}{k_p}\overline{\theta}\,. \tag{40}$$

One can observe that the system operates by assuring the constant value of p_c, the injection fuel flow rate being controlled through the dosage valve positioning, which means directly by the throttle. So, the system's relevant output is the p_c-pressure in chambers 10.

For a constant flight regime, altitude and airspeed ($H=\text{const.}, V=\text{const.}$), which mean that the air pressure and temperature before the engine's compressor are constant $\left(p_1^*=\text{const.}, T_1^*=\text{const.}\right)$, the term in equation (36) containing $\overline{p_1^*}$ becomes null.

3.3 System transfer function

Based on the above-presented mathematical model, one has built the block diagram with transfer functions (see figure 5) and one also has obtained a simplified expression:

$$\left\{\tau_y\tau_M\mathrm{s}^2+\left[\left(1-k_ck_{pn}\right)\tau_y+\left(1+\frac{k_rk_{py}}{k_p}\right)\tau_M\right]\mathrm{s}+\left(1-k_ck_{pn}\right)+\frac{k_rk_{py}}{k_p}\right\}\overline{p_c}=$$

$$-\frac{k_\theta}{k_p}\left(\tau_y\mathrm{s}+1\right)\left(\tau_M\mathrm{s}+1-k_ck_{pn}\right)\overline{\theta}+\frac{k_{py}k_r}{k_p}\left(\tau_M\mathrm{s}+1\right)\overline{p_{ci}}\,, \tag{41}$$

where $k_r=k_xk_l$.

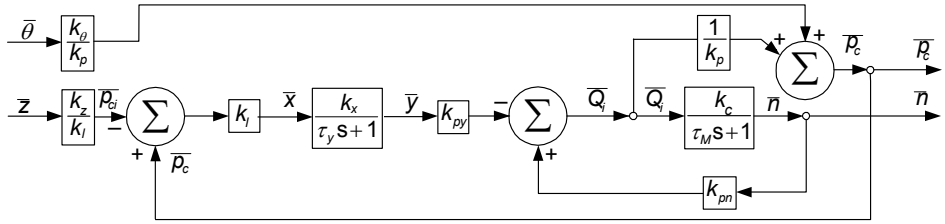


Fig. 5. System's block diagram with transfer functions

So, one can define two transfer functions:

a. with respect to the dosage valve's lever angular displacement $H_\theta(\mathrm{s})$;

b. with respect to the preset reference pressure p_{ci}, or to the sensor's spring's pre-compression z, $H_z(\mathrm{s})$.

While θ angle is permanently variable during the engine's operation, the reference pressure's value is established during the engine's tests, when its setup is made and

remains the same until its next repair or overhaul operation, so $\overline{z} = \overline{p_{ci}} = 0$ and the transfer function $H_z(\mathrm{s})$ definition has no sense. Consequently, the only system's transfer function remains

$$H_\theta(\mathrm{s}) = \frac{-\frac{k_\theta}{k_p}\left(\tau_y \mathrm{s}+1\right)\left(\tau_M \mathrm{s}+1-k_c k_{pn}\right)}{\tau_y \tau_M \mathrm{s}^2 + \left[\left(1-k_c k_{pn}\right)\tau_y + \left(1+\frac{k_r k_{py}}{k_p}\right)\tau_M\right]\mathrm{s} + 1 - k_c k_{pn} + \frac{k_r k_{py}}{k_p}}, \tag{42}$$

which characteristic polynomial's degree is 2.

3.4 System stability

One can perform a stability study, using the Routh-Hurwitz criteria, which are easier to apply because of the characteristic polynomial's form. So, the stability conditions are

$$\tau_y \tau_M > 0, \tag{43}$$

$$\left(1-k_c k_{pn}\right)\tau_y + \left(1+\frac{k_r k_{py}}{k_p}\right)\tau_M > 0, \tag{44}$$

$$1 - k_c k_{pn} + \frac{k_r k_{py}}{k_p} > 0. \tag{45}$$

The first condition (43) is obviously, always realized, because both τ_y and τ_M are strictly positive quantities, being time constant of the actuator, respectively of the engine.
The (40) and (41) conditions must be discussed.
The factor $1-k_c k_{pn}$ is very important, because its value is the one who gives information about the stability of the connection between the fuel pump and the engine's shaft (Stoicescu&Rotaru, 1999). There are two situation involving it:

a. $k_c k_{pn} < 1$, when the connection between the fuel pump and the engine shaft is a stable controlled object;
b. $k_c k_{pn} \geq 1$, when the connection fuel pump - engine shaft is an unstable object and it is compulsory to be assisted by a controller.

If $k_c k_{pn} < 1$, the factor $1-k_c k_{pn}$ is strictly positive, so $\left(1-k_c k_{pn}\right)\tau_y > 0$. According to their definition formulas (see annotations (35) and (30)), k_r, k_p, k_{py} are positive, so $\frac{k_r k_{py}}{k_p} > 0$ and $\left(1+\frac{k_r k_{py}}{k_p}\right)\tau_M > 0$, which means that both other stability requests, (44) and (45), are accomplished, that means that the system is a stable one for any situation.

If $k_c k_{pn} \geq 1$, the factor $1 - k_c k_{pn}$ becomes a negative one. The inequality (44) leads to

$$\tau_M < \frac{\left(k_c k_{pn} - 1\right)}{\left(1 + \frac{k_r k_{py}}{k_p}\right)} \tau_y \text{, or } \tau_y < \frac{\left(1 + \frac{k_r k_{py}}{k_p}\right)}{\left(k_c k_{pn} - 1\right)} \tau_M \text{,} \tag{46}$$

which offers a criterion for the time constant choice and establishes the boundaries of the stability area (see figure 6.a).

Meanwhile, from the inequality (45) one can obtain a condition for the sensor's elastic membrane surface area's choice, with respect to the drossels' geometry (d_A, d_n) and quality (μ_n, μ_A), springs' elastic constants (k_e, k_f), sensor's lever arms (l_1, l_2) and other stability co-efficient (k_c, k_{pn}, k_{py})

$$S_m > \frac{k_e}{k_f} \frac{l_2}{l_1} \frac{k_c k_{pn} - 1}{k_{py}} \frac{\left(2\mu_n d_n x_0 - \mu_{dA} d_A^2\right)\sqrt{2\rho k_f y_0}}{4\mu d_n p_{A0}}. \tag{47}$$

Another observation can be made, concerning the character of the stability, periodic or non-periodic. If the characteristic equation's discriminant is positive (real roots), than the system's stability is non-periodic type, otherwise (complex roots) the system's stability is periodic type. Consequently, the non-periodic stability condition is

$$\left[\left(1 - k_c k_{pn}\right)\tau_y + \left(1 + \frac{k_r k_{py}}{k_p}\right)\tau_M\right]^2 - 4\tau_y \tau_M + \frac{k_r k_{py}}{k_p}\Bigg) > 0 \text{,} \tag{48}$$

which leads to the inequalities

$$\frac{\tau_y}{\tau_M} < \frac{k_p\left(k_c k_{pn} - 1\right)\left[k_r k_{py} - k_p - \sqrt{2\left(k_p^2 + k_r^2 k_{py}^2\right)}\right]}{k_p^2 + k_r k_{py}\left(2k_p + k_r k_{py}\right)}, \tag{49}$$

$$\frac{\tau_y}{\tau_M} > \frac{k_p\left(k_c k_{pn} - 1\right)\left(k_r k_{py} - k_p + \sqrt{2\left(k_p^2 + k_r^2 k_{py}^2\right)}\right)}{k_p^2 + k_r k_{py}\left(2k_p + k_r k_{py}\right)}, \tag{50}$$

representing two semi-planes, which boundaries are two lines, as figure 6.b shows; the area between the lines is the periodic stability domain, respectively the areas outside are the non-periodic stability domains.

Obviously, both time constants must be positive, so the domains are relevant only for the positives sides of τ_y and τ_M axis.

Both figures (6.a and 6.b) are showing the domains for the pump actuator time constant choice or design, with respect to the jet engine's time constant.

The studied system can be characterized as a 2nd order controlled object. For its stability, the most important parameters are engine's and actuator's time constants; a combination of a small τ_y and a big τ_M, as well as vice-versa (until the stability conditions are accomplished), assures the non periodic stability, but comparable values can move the stability into the periodic domain; a very small τ_y and a very big τ_M are leading, for sure, to instability.

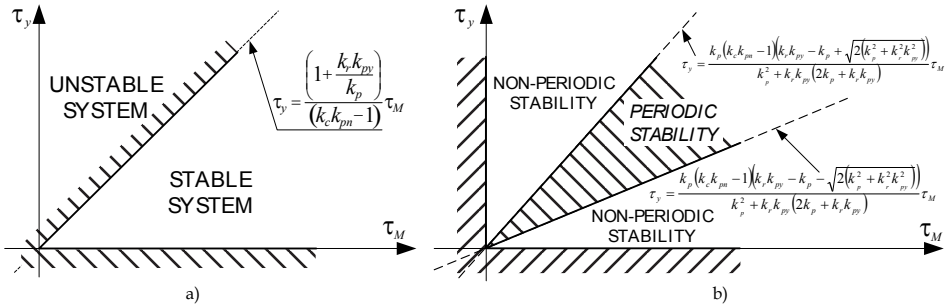


Fig. 6. System's stability domains

3.4 System quality

As the transfer function form shows, the system is static one, being affected by static error.
One has studied/simulated a controller serving on an engine RD-9 type, from the point of view of the step response, which means the system's behavior for step input of the dosage valve's lever's angle θ.
System's time responses, for the fuel injection pressure p_c and for the engine's speed *n* are

$$\overline{p_c}(t) = \frac{-k_\theta}{k_p}\left(1 - \frac{1}{1 + \frac{k_r k_{py}}{1 - k_c k_{pn}}}\right)\overline{\theta}(t), \tag{51}$$

$$\overline{n}(t) = \frac{k_\theta k_c k_r}{k_p\left(1 - k_c k_{pn}\right) + k_r k_{py}}\overline{\theta}(t), \tag{52}$$

as shown in figure 7.a). One can observe that the pressure p_c has an initial step decreasing, $p_c(0) = -\frac{k_\theta}{k_p}$, then an asymptotic increasing; meanwhile, the engine's speed is continuous asymptotic increasing.
One has also performed a simulation for a hypothetic engine, which has such a co-efficient combination that $k_c k_{pn} \geq 1$; even in this case the system is a stable one, but its stability happens to be periodic, as figure 7.b) shows. One can observe that both the pressure and the speed have small overrides (around 2.5% for *n* and 1.2% for p_c) during their stabilization.

The chosen RD-9 controller assures both stability and asymptotic non-periodic behavior for the engine's speed, but its using for another engine can produce some unexpected effects.

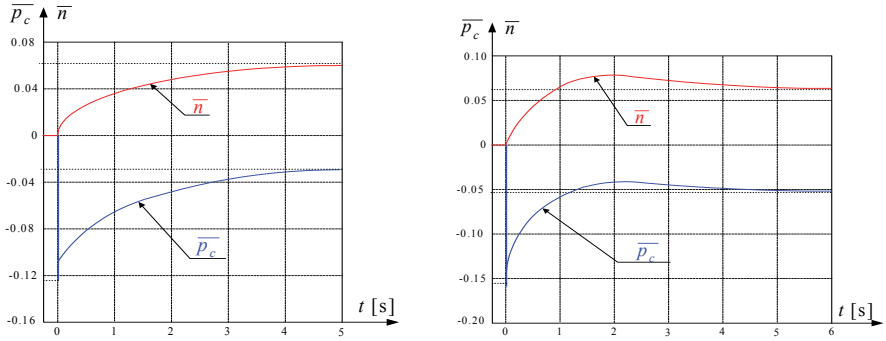


Fig. 7. System's quality (system time response for θ -step input)

4. Fuel injection controller with constant differential pressure

Another fuel injection control system is the one in figure 8, which assures a constant value of the dosage valve's differential pressure $p_c - p_i$, the fuel flow rate amount Q_i being determined by the dosage valve's opening.

As figure 8 shows, a rotation speed control system consists of four main parts: I-fuel pump with plungers (4) and mobile plate (5); II-pump's actuator with spring (22), piston (23) and rod (6); III-differential pressure sensor with slide valve (17), preset bolt (20) and spring (18); IV-dosage valve, with its slide valve (11), connected to the engine's throttle through the rocking lever (13).

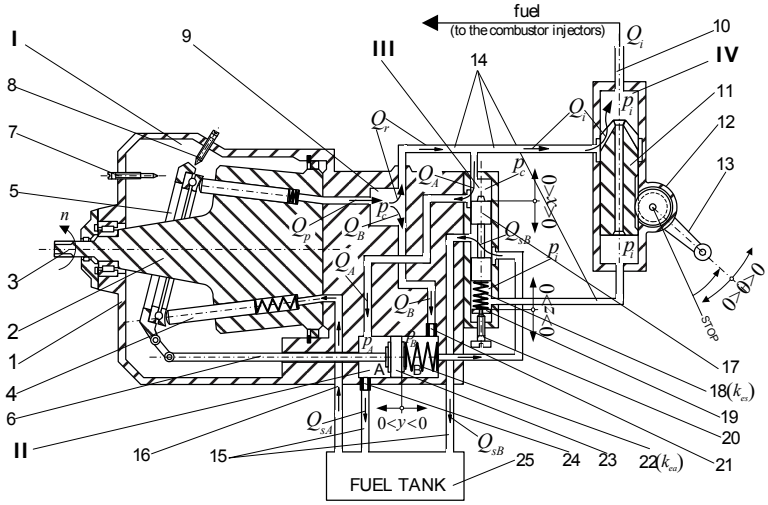


Fig. 8. Fuel injection controller with constant differential pressure $p_r = p_c - p_i$

The system operates by keeping a constant difference of pressure, between the pump's pressure chamber (9) and the injectors' pipe (10), equal to the preset value (proportional to the spring (18) pre-compression, set by the adjuster bolt (20)). The engine's necessary fuel flow rate Q_i and, consequently, the engine's speed n, are controlled by the co-relation between the $p_r = p_c - p_i$ differential pressure's amount and the dosage valve's variable slot opening (proportional to the (13) rocking lever's angular displacement θ).

4.1 Mathematical model and transfer function

The non-linear mathematical model consists of the motion equations for each above described sub-system In order to bring it to an operable form, assuming the small perturbations hypothesis, one has to apply the finite difference method, then to bring it to a non-dimensional form and, finally, to apply the Laplace transformer (as described in 3.2). Assuming, also, that the fuel is a non-compressible fluid, the inertial effects are very small, as well as the viscous friction, the terms containing m, β and ξ are becoming null. Consequently, the simplified mathematical model form shows as follows

$$\left(\tau_p \mathrm{s} + 1\right)\left(\overline{p_B} - \overline{p_A}\right) = -k_{px}\overline{x}\,, \tag{53}$$

$$\overline{p_B} - \overline{p_A} = k_{AB}\overline{y}\,, \tag{54}$$

$$\overline{x} = \frac{1}{k_{pic}}\left(\overline{p_c} - \overline{p_i}\right) - \frac{k_{iz}}{k_{pic}}\overline{z}\,, \tag{55}$$

$$\overline{p_c} - \overline{p_i} = \frac{1}{k_{Qp}}\left(\overline{Q_p} - k_\theta\overline{\theta} - k_{Qx}\overline{x}\right), \tag{56}$$

$$\overline{Q_i} = \overline{Q_p} - k_\theta\overline{\theta}\,, \tag{57}$$

$$\overline{Q_p} = k_{pn}\overline{n} + k_{py}\overline{y}\,. \tag{58}$$

The model should be completed by the jet engine as controlled object equation

$$\left(\tau_M \mathrm{s} + 1\right)\overline{n} = k_c\overline{Q_i} + k_{HV}\overline{p_1^*}\,, \tag{59}$$

where, for a constant flight regime, the term $k_{HV}\overline{p_1^*}$ becomes null.

The equations (53) to (59), after eliminating the intermediate arguments $\overline{p_A}, \overline{p_B}, \overline{p_c}$, $\overline{p_i}, \overline{Q_i}, \overline{Q_p}, \overline{y}, \overline{x}$, are leading to a unique equation:

$$\left[\left(k_{pic} + k_{Qx}\right)\frac{k_{AB}}{k_{px}}\left(\tau_p \mathrm{s} + 1\right)\frac{\tau_M \mathrm{s} + \left(1 - k_c k_{pn}\right)}{k_c k_{py}} - \frac{\tau_M \mathrm{s} + 1}{k_c}\right]\overline{n} = k_{iz}\overline{z} + \left(k_{pic} + k_{Qx}\right)\frac{k_{AB}k_\theta}{k_{px}k_{py}}\left(\tau_p \mathrm{s} + 1\right)\overline{\theta}\,. \tag{60}$$

System's transfer function is $H_\theta(s)$, with respect to the dosage valve's rocking lever's position θ. A transfer function with respect to the setting z, $H_z(s)$, is not relevant, because the setting and adjustments are made during the pre-operational ground tests, not during the engine's current operation.
So, the main and the most important transfer function has the form below

$$H_\theta(s) = \frac{f_1 s + f_0}{g_2 s^2 + g_1 s + g_0}, \tag{61}$$

where the involved co-efficient are $f_1 = k_c k_\theta \tau_p$, $f_0 = k_c k_\theta$, $g_2 = \tau_p \tau_M$,

$$g_1 = \tau_p\left(1 - k_c k_{pn}\right) + \tau_M\left[1 - \frac{k_{px}k_{py}}{k_{AB}\left(k_{pic} + k_{Qx}\right)}\right], \quad g_0 = \left(1 - k_c k_{pn}\right) - \frac{k_{px}k_{py}}{k_{AB}\left(k_{pic} + k_{Qx}\right)}. \tag{62}$$

4.2 System quality

As the transfer function shows, the system is a static-one, being affected by static error.
One has studied/simulated a controller serving on a single spool jet engine (VK-1 type), from the point of view of the step response, which means the system's dynamic behavior for a step input of the dosage valve's lever's angle θ.
According to figure 9.a), for a step input of the throttle's position α, as well as of the lever's angle θ, the differential pressure $p_r = p_c - p_i$ has an initial rapid lowering, because of the initial dosage valve's step opening, which leads to a diminution of the fuel's pressure p_c in the pump's chamber; meanwhile, the fuel flow rate through the dosage valve grows. The differential pressure's recovery is non-periodic, as the curve in figure 9.a) shows.
Theoretically, the differential pressure re-establishing must be made to the same value as before the step input, but the system is a static-one and it's affected by a static error, so the new value is, in this case, higher than the initial one, the error being 4.2%. The engine's speed has a different dynamic behavior, depending on the $k_c k_{pn}$ particular value.

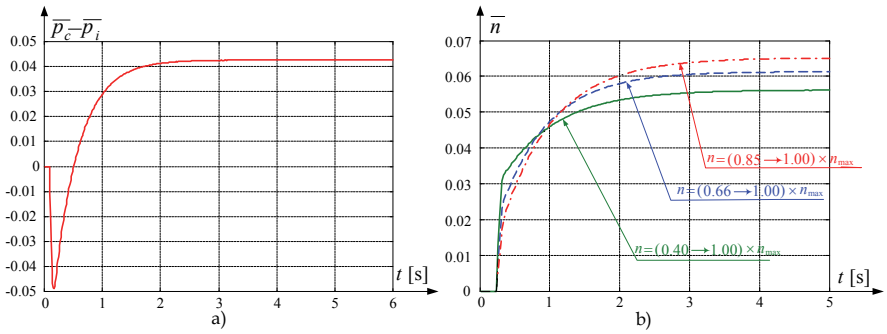


Fig. 9. System's quality (system time response for θ-step input)

One has performed simulations for a VK-1-type single-spool jet engine, studying three of its operating regimes: a) full acceleration (from idle to maximum, that means from

$0.4 \times n_{max}$ to n_{max}); b) intermediate acceleration (from $0.65 \times n_{max}$ to n_{max}); c) cruise acceleration (from $0.85 \times n_{max}$ to n_{max}).

If $k_c k_{pn} < 1$, so the engine is a stable system, the dynamic behavior of its rotation speed n is shown in figure 9.b). One can observe that, for any studied regime, the speed n, after an initial rapid growth, is an asymptotic stable parameter, but with static error. The initial growing is maxim for the full acceleration and minimum for the cruise acceleration, but the static error behaves itself in opposite sense, being minimum for the full acceleration.

5. Fuel injection controller with commanded differential pressure

Unlike the precedent controller, where the differential pressure was kept constant and the fuel flow rate was given by the dosage valve opening, this kind of controller has a constant injection orifice and the fuel flow rate variation is given by the commanded differential pressure value variation. Such a controller is presented in figure 10, completed by two correctors (a barometric corrector VII and an air flow rate corrector VIII, see 5.3).

The basic controller has four main parts (the pressure transducer I, the actuator II, the actuator's feed-back III and the fuel injector IV); it operates together with the fuel pump V, the fuel tank VI and, obviously, with the turbo-jet engine.

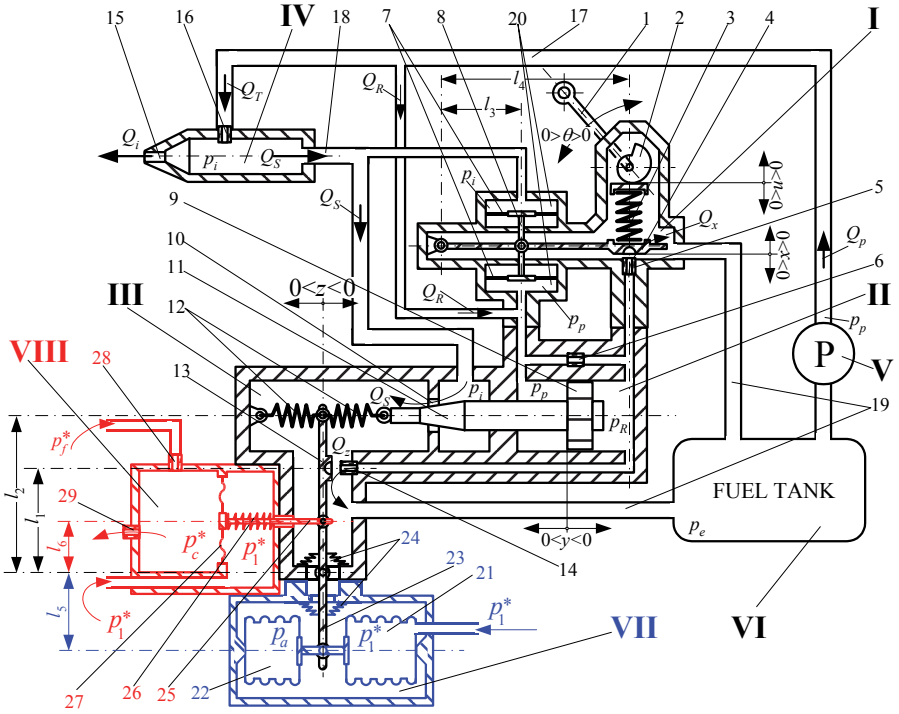


Fig. 10. Fuel injection controller with commanded differential pressure $p_r = p_c - p_i$ (basic controller), with barometric corrector and air flow rate corrector

Controller's duty is to assure, in the injector's chamber, the appropriate p_i value, enough to assure the desired value of the engine's speed, imposed by the throttle's positioning, which means to co-relate the pressure difference $p_p - p_i$ to the throttle's position (given by the 1 lever's θ -angle).
The fuel flow rate Q_i, injected into the engine's combustor, depends on the injector's diameter (drossel no. 15) and on the fuel pressure in its chamber p_i. The difference $p_p - p_i$, as well as p_i, are controlled by the level of the discharged fuel flow Q_S through the calibrated orifice 10, which diameter is given by the profiled needle 11 position; the profiled needle is part of the actuator's rod, positioned by the actuator's piston 9 displacement.
The actuator has also a distributor with feedback link (the flap 13 with its nozzle or drossel 14, as well as the springs 12), in order to limit the profiled needle's displacement speed.
Controller's transducer has two pressure chambers 20 with elastic membranes 7, for each measured pressure p_p and p_i; the inter-membrane rod is bounded to the transducer's flap 4. Transducer's role is to compare the level of the realized differential pressure $p_p - p_i$ to its necessary level (given by the 3 spring's elastic force, due to the (lever1+cam2) ensemble's rotation). So, the controller assures the necessary fuel flow rate value Q_i, with respect to the throttle's displacement, by controlling the injection pressure's level through the fuel flow rate discharging.

5.1 System mathematical model and block diagram with transfer functions

Basic controller's linear non-dimensional mathematical model can be obtained from the motion equations of each main part, using the same finite differences method described in chapter 3, paragraph 3.2, based on the same hypothesis.
The simplified mathematical model form is, as follows

$$\overline{p_R} = k_{1p}\overline{p_p} - k_{1x}\overline{x} - k_{1y}\left(\tau_y \mathrm{s} + 1\right)\overline{y}\,, \tag{63}$$

$$\overline{p_p} = k_{2p}\overline{p_i} + k_{2R}\overline{p_R} + k_{2Q}\overline{Q_p}\,, \tag{64}$$

$$\overline{y} = k_{yR}\overline{p_R} - k_{yp}\overline{p_p}\,, \tag{65}$$

$$\overline{p_i} = k_{3p}\overline{p_p} - k_{3y}\overline{y}\,, \tag{66}$$

$$\overline{Q_i} = k_i\overline{p_i}\,,\ \overline{u} = k_{u\theta}\overline{\theta}\,, \tag{67}$$

together with the fuel pump and the engine's speed non-dimensional equations

$$\overline{Q_p} = k_{pn}\overline{n}\,, \tag{68}$$

$$\left(\tau_m \mathrm{s} + 1\right)\overline{n} = k_c\overline{Q_i} + k_{HV}\overline{p_1^*}\,. \tag{69}$$

where the used annotations are

$$k_{PT} = \mu_d \frac{\pi d_{16}^2}{4}\frac{1}{\sqrt{2\rho\left(p_{p0} - p_{i0}\right)}}, k_{xx} = \mu_4 \pi d_4 \sqrt{\frac{2p_{R0}}{\rho}}, k_{Qi} = \mu_d \frac{\pi d_i^2}{4}\frac{1}{\sqrt{2\rho p_{p0}}}, k_{zz} = \mu_{11}\pi d_{11}\sqrt{\frac{2p_{R0}}{\rho}},$$

$$k_{RP}=\mu_7\frac{\pi d_7^2}{4}\frac{1}{\sqrt{2\rho\left(p_{p0}-p_{R0}\right)}},\ k_{sy}=\mu\frac{\pi}{4}\left[d_0^2-d_1^2-4\left(d_1-y_0\tan\alpha\right)\tan\alpha\right]\sqrt{\frac{2p_{i0}}{\rho}},$$

$$k_{si}=\mu\frac{\pi}{4}\left[d_0^2-d_1^2-4\left(d_1-y_0\tan\alpha\right)\tan\alpha\right]\frac{1}{\sqrt{2\rho p_{i0}}},\ k_{zR}=\mu_{11}\pi d_{11}(z_0-z_s)\sqrt{\frac{1}{2\rho p_{R0}}},$$

$$k_{xR}=\mu_4\pi d_4(x_0-x_s)\sqrt{\frac{1}{2\rho p_{R0}}},\ k_{u\theta}=k_\theta\frac{\theta_0}{u_0},\ k_{xd}=\frac{S_m l_3 p_{d0}}{k_{rs}l_4 x_0},\ k_u=\frac{u_0}{x_0},\ k_{dp}=\frac{p_{p0}}{p_{d0}},\ k_{di}=\frac{p_{i0}}{p_{d0}},$$

$$k_{1p}=\frac{k_{RP}p_{p0}}{\left(k_{RP}+k_{xR}+k_{zR}\right)p_{R0}},\ \tau_y=\frac{S_R l_2}{k_{zz}l_1},\ k_{1y}=\frac{k_{zz}l_1 y_0}{p_{R0}l_2},\ k_{1x}=\frac{k_{xx}x_0}{\left(k_{RP}+k_{xR}+k_{zR}\right)p_{R0}},\ k_i=\frac{k_{Qi}p_{i0}}{Q_{i0}}$$

$$k_{2p}=\frac{k_{PT}p_{i0}}{\left(k_{RP}+k_{PT}\right)p_{p0}},\ k_{2R}=\frac{k_{RP}p_{R0}}{\left(k_{RP}+k_{PT}\right)p_{p0}},\ k_{2Q}=\frac{Q_{p0}}{\left(k_{RP}+k_{PT}\right)p_{p0}},\ k_{yR}=\frac{S_R p_{R0}}{\left(k_{r1}+k_{r2}\right)y_0},$$

$$k_{yp}=\frac{S_P p_{p0}}{\left(k_{r1}+k_{r2}\right)y_0},\ k_{3p}=\frac{k_{PT}p_{p0}}{\left(k_{PT}+k_{si}+k_{Qi}\right)p_{i0}},\ k_{3y}=\frac{k_{sy}y_0}{\left(k_{PT}+k_{si}+k_{Qi}\right)p_{i0}},\ k_{pn}=\frac{n_0}{Q_{p0}}\left(\frac{\partial Q_p}{\partial n}\right)_0. \quad (70)$$

Furthermore, if the input signal u is considered as the reference signal forming parameter, one can obtain the expression

$$\overline{x}=-k_{xd}\left(\overline{p_{d_{ref}}}-\overline{p_d}\right), \quad (71)$$

where $\overline{p_{d_{ref}}}$ is the reference differential pressure, given by $\overline{p_{d_{ref}}}=\frac{k_\theta k_{rs}l_4}{S_m l_3 p_{d0}}\overline{\theta}=k_{r\theta}\overline{\theta}$.

A block diagram with transfer functions, both for the basic controller and the correctors' block diagrams (colored items), is presented in figure 11.

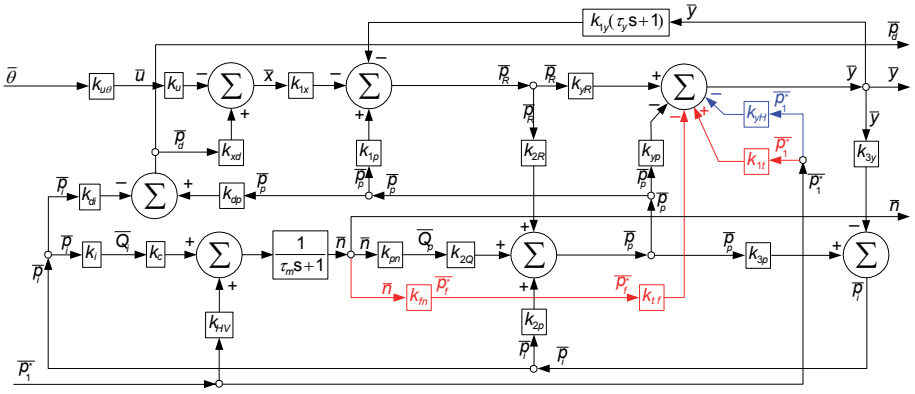


Fig. 11. Block diagram with transfer functions

5.2 System quality

As figures 10 and 11 show, the basic controller has two inputs: a) throttle's position - or engine's operating regime - (given by θ-angle) and b) aircraft flight regime (altitude and airspeed, given by the inlet inner pressure p_1^*). So, the system should operate in case of disturbances affecting one or both of the input parameters $\left(\overline{\theta}, \overline{p_1^*}\right)$.

A study concerning the system quality was realized (using the co-efficient values for a VK-1F jet engine), by analyzing its step response (system's response for step input for one or for both above-mentioned parameters). As output, one has considered the differential pressure $\overline{p_d}$, the engine speed $\overline{n}$ (which is the most important controlled parameter for a jet engine) and the actuator's rod displacement $\overline{y}$ (same as the profiled needle).

Output parameters' behavior is presented by the graphics in figure 12; the situation in figure 12.a) has as input the engine's regime (step throttle's repositioning) for a constant flight regime; in the mean time, the situation in figure 12.b) has as input the flight regime (hypothetical step climbing or diving), for a constant engine regime (throttle constant position). System's behavior for both input parameters step input is depicted in figure 12.a).

One has also studied the system's behavior for two different engine's models: a stable-one (which has a stabile pump-engine connection, its main co-efficient being $k_c k_{pn} < 1$, situation in figure 13.a) and a non-stable-one (which has an unstable pump-engine connection and $k_c k_{pn} > 1$, see figure 13.b).

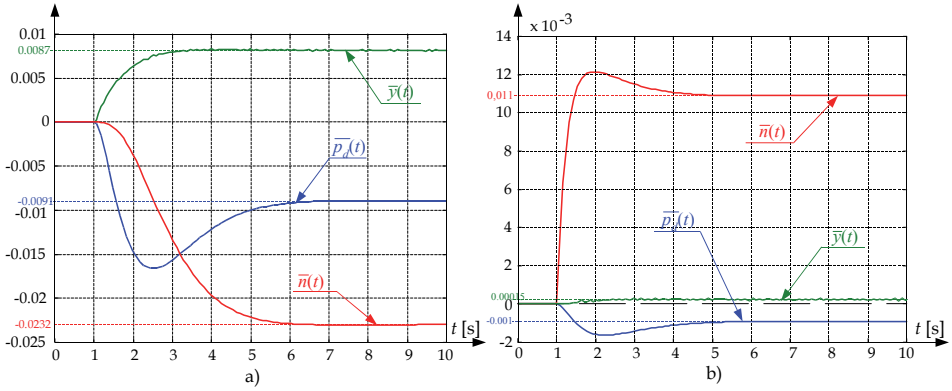


Fig. 12. Basic system step response a) step input for $\overline{\theta}$ $\left(\overline{p_1^*} = 0\right)$; b) step input for $\overline{p_1^*}$ $\left(\overline{\theta} = 0\right)$

Concerning the system's step response for throttle's step input, one can observe that all the output parameters are stables, so the system is a stable-one. All output parameters are stabilizing at their new values with static errors, so the system is a static-one. However, the static errors are acceptable, being fewer than 2.5% for each output parameter. The differential pressure and engine's speed static errors are negative, so in order to reach the engine's speed desired value, the throttle must be supplementary displaced (pushed).

For immobile throttle and step input of $\overline{p_1^*}$ (flight regime), system's behavior is similar (see figure 12.b), but the static errors' level is lower, being around 0.1% for $\overline{p_d}$ and for $\overline{y}$, but higher for $\overline{n}$ (around 1.1%, which mean ten times than the others).
When both of the input parameters have step variations, the effects are overlapping, so system's behavior is the one in figure 13.a).
System's stability is different, for different analyzed output parameters: $\overline{y}$ has a non-periodic stability, no matter the situation is, but $\overline{p_d}$ and $\overline{n}$ have initial stabilization values overriding. Meanwhile, curves in figures 12.a), 12.b) and 13.a) are showing that the engine regime has a bigger influence than the flight regime above the controller's behavior.
One also had studied a hypothetical controller using, assisting an unstable connection engine-fuel pump. One has modified k_c and k_{pn} values, in order to obtain such a combination so that $k_c k_{pn} > 1$. Curves in figure 13.b are showing a periodical stability for a controller assisting an unstable connection engine-fuel pump, so the controller has reached its limits and must be improved by constructive means, if the non-periodic stability is compulsory.

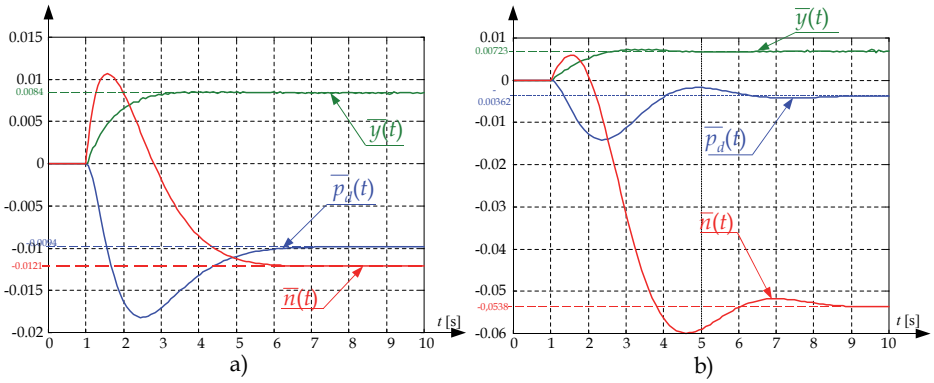


Fig. 13. Compared step response between a) stabile fuel pump-engine connection $\left(k_c k_{pn} < 1\right)$ and b) unstable fuel pump-engine connection $\left(k_c k_{pn} > 1\right)$

5.3 Fuel injection controller with barometric and air flow rate correctors

5.3.1 Correctors using principles

For most of nowadays operating controllers, designed and manufactured for modern jet engines, their behavior is satisfying, because the controlled systems become stable and their main output parameters have a non-periodic (or asymptotic) stability. However, some observations regarding their behavior with respect to the flight regime are leading to the conclusion that the more intense is the flight regime, the higher are the controllers' static errors, which finally asks a new intervention (usually from the human operator, the pilot) in order to re-establish the desired output parameters levels. The simplest solution for this issue is the flight regime correction, which means the integration in the control system of

new equipment, which should adjust the control law. These equipments are known as barometric (bar-altimetric or barostatic) correctors.
In the mean time, some unstable engines or some unstable fuel pump-engine connections, even assisted by fuel controllers, could have, as controlled system, periodic behavior, that means that their output main parameters' step responses presents some oscillations, as figure 14 shows. The immediate consequence could be that the engine, even correctly operating, could reach much earlier its lifetime ending, because of the supplementary induced mechanical fatigue efforts, combined with the thermal pulsatory efforts, due to the engine combustor temperature periodic behavior.
As fig. 14 shows, the engine speed n and the combustor temperature T_3^* (see figure 14.b), as well as the fuel differential pressure p_d and the pump discharge slide-valve displacement y (see figure 14.a) have periodic step responses and significant overrides (which means a few short time periods of overspeed and overheat for each engine full acceleration time).
The above-described situation could be the consequence of a miscorrelation between the fuel flow rate (given by the connection controller-pump) and the air flow rate (supplied by the engine's compressor), so the appropriate corrector should limit the fuel flow injection with respect to the air flow supplying.

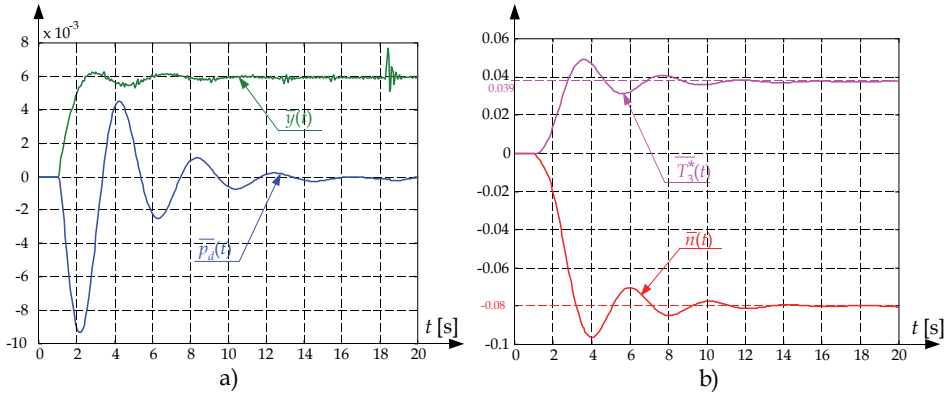


Fig. 14. Step response for an unstable fuel pump-engine connection assisted by a fuel injection pressure controller

The system depicted in figure 10 has as main control equipment a fuel injection controller (based on the differential pressure control) and it is completed by a couple of correction equipment (correctors), one for the flight regime and the other for the fuel-air flow rates correlation.
The correctors have the active parts bounded to the 13-lever (hemi-spherical lid's support of the nozzle-flap actuator's distributor). So, the 13-lever's positioning equation should be modified, according to the new pressure and forces distribution.

5.3.2 Barometric corrector

The barometric corrector (position VII in figure 10) consists of an aneroid (constant pressure) capsule and an open capsule (supplied by a p_1^* - total pressure intake), bounded by a common rod, connected to the 13-lever.

The total pressure p_1^* (air's total pressure after the inlet, in the front of the engine's compressor) is an appropriate flight regime estimator, having as definition formula

$$p_1^* = p_H \Pi(M_H) \sigma_c^*, \tag{72}$$

where p_H is the air static pressure of the flight altitude H, σ_c^* – inlet's inner total pressure lose co-efficient (assumed as constant), M_H – air's Mach number in the front of the inlet, k – air's adiabatic exponent and $\Pi(M_H) = \left(1 + \frac{k-1}{2} M_H^2\right)^{\frac{k}{k-1}}$.

The new equation of the 13-lever becomes

$$S_R p_R - S_p p_p = m_s \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \xi \frac{\mathrm{d}y}{\mathrm{d}t} + (k_{r1} + k_{r2}) y - S_H \left(p_1^* - p_a\right) \frac{l_5}{l_2}, \tag{73}$$

where p_a is the aneroid capsule's pressure and, after the linearization and the Laplace transformer applying, its new non-dimensional form becomes

$$k_{yR} \overline{p_R} - k_{yp} \overline{p_p} - k_{yH} \overline{p_1^*} = \left(T_y^2 \mathrm{s}^2 + 2\omega_0 T_y \mathrm{s} + 1\right) \overline{y} \cong \overline{y}, \tag{74}$$

and will replace the (65)-equation (see paragraph 5.2), where $k_{yH} = \frac{S_H p_{10}^*}{(k_{r1} + k_{r2}) y_0} \frac{l_5}{l_2}$.

5.3.3 Air flow-rate corrector

The air flow-rate corrector (position VIII in figure 10) consists of a pressure ratio transducer, which compares the realized pressure ratio value for a current speed engine to the preset value. The air flow-rate Q_a is proportional to the total pressure difference $p_2^* - p_1^*$, as well as to the engine's compressor pressure ratio $\pi_c^* = \frac{p_2^*}{p_1^*}$. According to the compressor universal characteristics, for a steady state engine regime, the air flow-rate depends on the pressure ratio and on the engine's speed $Q_a = Q_a\left(\pi_c^*, n\right)$ (Soicescu&Rotaru, 1999). The air flow-rate must be correlated to the fuel flow rate Q_i, in order to keep the optimum ratio of these values. When the correlation is not realized, for example when the fuel flow rate grows faster/slower than the necessary air flow rate during a dynamic regime (e.g. engine acceleration/deceleration), the corrector should modify the growing speed of the fuel flow rate, in order to re-correlate it with the realized air flow rate growing speed.

Modern engines' compressors have significant values of the pressure ratio, from 10 to 30, so the pressure difference $p_2^* - p_1^*$ could damage, even destroy, the transducer's elastic membrane and get it out of order. Thus, instead of p_2^*-pressure, an intermediate pressure p_f^*, from an intermediate compressor stage "*f*", should be used, the intermediate pressure ratio $\pi_f^* = \frac{p_f^*}{p_1^*}$ being proportional to π_c^*. The intermediate stage is chosen in order to obtain

a convenient value of p_f^*, around $4\times p_1^*$. Both values of p_f^* and p_2^* are depending on compressor's speed (the same as the engine speed *n*), as the compressor's characteristic shows; consequently, the air flow rate depends on the above-mentioned pressure (or on the above defined π_c^* or π_f^*). The transducer's command chamber has two drossels, which are chosen in order to obtain critical flow through them (Soicescu&Rotaru, 1999), so the corrected pressure p_c^* is proportional to the input pressure:

$$p_c^* = \frac{S_{28}}{S_{29}} p_f^* , \tag{75}$$

where S_{28}, S_{29} are 28 and 29-drossels' effective area values. Consequently, the transducer operates like a π_c^*-based corrector, correlating the necessary fuel flow-rate with the compressor delivered air flow-rate. So, the corrector's equations are:

$$p_2^* = p_2^*(n) \text{ or } p_f^* = p_f^*(n), \tag{76}$$

and become, after transformations,

$$\overline{p_f^*} = k_{fn}\bar{n} , \tag{76'}$$

$$\overline{x_f} = k_\pi \overline{\pi_f^*} = k_{sf}(\overline{p_f^*} - \frac{1}{k_{1p}}\overline{p_1^*}) . \tag{77}$$

The new form of (65)-equation becomes

$$S_R p_R - S_p p_p = m_s \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \xi \frac{\mathrm{d}y}{\mathrm{d}t} + (k_{r1} + k_{r2}) y - S_{mp}\left(p_c^* - p_1^*\right)\frac{l_6}{l_2} , \tag{78}$$

where S_{mp} is the transducer's membrane surface area. After linearization and Laplace transformer applying, its new non-dimensional form becomes

$$\left(T_y^2 \mathrm{s}^2 + 2\omega_0 T_y \mathrm{s} + 1\right)\bar{y} = k_{yR}\overline{p_R} - k_{yp}\overline{p_p} - \left(k_{tf}\overline{p_f^*} - k_{1t}\overline{p_1^*}\right), \tag{79}$$

where

$$k_{fn} = \left(\frac{\partial p_f^*}{\partial n}\right)_0 \frac{n_0}{p_{f0}^*} , k_\pi = \frac{S_{28}}{S_{29}} \frac{S_{mp} p_{f0}^*}{x_{f0} k_{r26}} \pi_{fr}^* , k_{tf} = \frac{S_{mp} p_{f0}^* k_{sf}}{(k_{r1} + k_{r2}) y_0} \frac{l_6}{l_2} , k_{1t} = \frac{S_{mp} p_{10}^* k_{s\pi}}{(k_{r1} + k_{r2}) k_{1p} y_0} \frac{l_6}{l_2} . \tag{80}$$

For a controller with both of the correctors, the (13)-lever equation results overlapping (73) and (79)-equations, which leads to a new form

$$k_{yR}\overline{p_R} - k_{yp}\overline{p_p} - k_{yH}\overline{p_1^*} - \left(k_{tf}\overline{p_f^*} - k_{1t}\overline{p_1^*}\right) = \left(T_y^2 \mathrm{s}^2 + 2\omega_0 T_y \mathrm{s} + 1\right)\bar{y} \cong \bar{y} , \tag{81}$$

which should replace the (65)-equation in the mathematical model (equations (63) to (69)). The new block diagram with transfer functions is depicted in figure 11.

5.3.4 System's quality

System's behavior was studied comparing the step responses of a basic controller and the step response (same conditions) of a controller with correctors. Fig. 15.a presents the step responses for a controller with barometric corrector, when the engine's regime is kept constant and the flight regime receives a step modifying. The differential pressure $\overline{p_d}$ becomes non-periodic, but its static error grows, from -0.1% to 0.77% and changes its sign. The profiled needle position y behavior is clearly periodic, with a significant override, more pulsations and a much bigger static error (1.85%, than 0.2%). Engine's most important output parameter, the speed n, presents the most significant changes: it becomes non-periodic (or remains periodic but has a short time smaller override), its static error decreases, from 1.1% to 0.21% and it becomes negative.

However, in spite of the above described output parameter behavior changes, the barometric corrector has realized its purpose: to keep (nearly) constant the engine's speed when the throttle has the same position, even if the flight regime (flight altitude or/and airspeed) significantly changes.

Figure 15.b presents system's behavior when an air flow-rate corrector assists the controller's operation. The differential pressure keeps its periodic behavior, but the profiled needle's displacement tends to stabilize non-periodic, which is an important improvement. The main output non-dimensional parameters, the engine's speed $\overline{n}$ and the combustor's temperature $\overline{T_3^*}$ have suffered significant changes, comparing to figure 14.b; both of them tend to become non-periodic, their static errors (absolute values) being smaller (especially for $\overline{n}$). System's time of stabilization became smaller (nearly half of the basic controller initial value). So, the flow rate corrector has improved the system, eliminating the overrides (potential engine's overheat and/or overspeed), resulting a non-periodic stable system, with acceptable static errors (5.5% for $\overline{n}$, 3% for $\overline{T_3^*}$) and acceptable response times (5 to 12 sec).

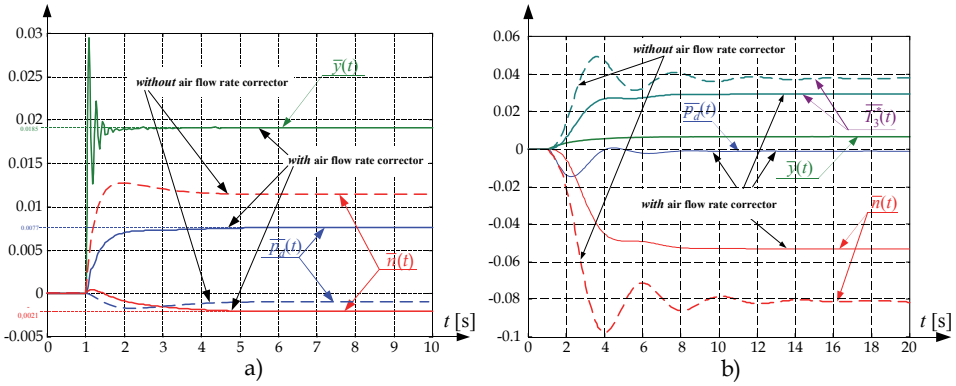


Fig. 15. Compared step response between a basic controller and a controller with
a) barometric corrector, b) air flow rate corrector

The barometric corrector is simply built, consisting of two capsules; its integration into the controller's ensemble is also accessible and its using results, from the engine's speed point of view, are definitely positives; new system's step response shows an improvement, the engine's speed having smaller static errors and a faster stabilization, when the flight regime changes. However, an inconvenience occurs, short time vibrations of the profiled needle (see figure 15.b, curve $\overline{y}(t)$), without any negative effects above the other output parameters, but with a possible accelerated actuator piston's wearing out.
The air flow rate corrector, in fact the pressure ratio corrector, is not so simply built, because of the drossels diameter's choice, correlated to its membrane and its spring elastic properties. However, it has a simple shape, consisting of simple and reliable parts and its operating is safe, as long as the drossels and the mobile parts are not damaged.
Air flow-rate corrector's using is more spectacular, especially for the unstable engines and/or for the periodic-stable controller assisted engines; system's dynamic quality changes (its step response becomes non-periodic, its response time becomes significantly smaller).

6. Conclusions

Fuel injection is the most powerful mean to control an engine, particularly an aircraft jet engine, the fuel flow rate being the most important input parameter of a control system.
Nowadays hydro-mechanical and/or electro-hydro-mechanical injection controllers are designed and manufactured according to the fuel injection principles; they are accomplishing the fuel flow rate control by controlling the injection pressure (or differential pressure) or/and the dosage valve effective dimension.
Studied controllers, similar to some in use aircraft engine fuel controllers, even if they operate properly at their design regime, flight regime's modification, as well as transient engine's regimes, induce them significant errors; therefore, one can improve them by adding properly some corrector systems (barometric and/or pneumatic), which gives more stability an reliability for the whole system (engine-fuel pump-controller).
Both of above-presented correctors could be used for other fuel injection controllers and/or engine speed controllers (for example for the controller with constant pressure chamber), if one chooses an appropriate integration mode and appropriate design parameters.

7. References

Abraham, R. H. (1986). *Complex dynamical systems,* Aerial Press, ISBN #0-942344-27-8, Santa Cruz, California, USA

Aron, I.; Tudosie, A. (2001). Jet Engine Exhaust Nozzle's Automatic Control System, *Proceedings of the 17th International Symposium on Naval and Marine Education*, pp. 36-45, section III, Constanta, Romania, May 24-26, 2001

Jaw, L. C.; Mattingly, J. D. (2009). *Aircraft Engine Controls:Design System Analysis and Health Monitoring,* Published by AIAA, ISBN-13: 978-1-60086-705-7, USA

Lungu, R.; Tudosie, A. (1997). Single Jet Engine Speed Control System Based on Fuel Flow Rate Control, *Proceedings of the XXVIIth International Conference of Technical Military Academy in Bucuresti*, pp. 74-80, section 4, Bucuresti, Romania, Nov. 13-14, 1997

Lungu, R. (2000). *Flying Vehicles Automation,* Universitaria, ISBN 973-8043-11-5, Craiova, Romania

Mattingly, J. D. (1996). *Elements of Gas Turbine Propulsion,* McGraw Hill, ISBN 1-56347-779-3, New York, USA

Moir, I.; Seabridge, A. (2008). *Aircraft Systems. Mechanical, Electrical and Avionics Subsystems Integration,* Professional Engineering Publication, ISBN-13: 978-1-56347-952-6, USA

Stoenciu, D. (1986). *Aircraft Engine Automation. Catalog of Automation Schemes,* Military Technical Academy Printing House, Bucuresti, Romania

Stoicescu, M.; Rotaru, C. (1999). *Turbo-Jet Engines. Characteristics and Control Methods,* Military Technical Academy Printing House, ISBN 973-98940-5-4, Bucuresti, Romania

Tudosie, A. N. (2009). Fuel Injection Controller with Barometric and Air Flow Rate Correctors, *Proceedings of the WSEAS International Conference on System Science and Simulation in Engineering* (ICOSSSE'09), pp. 113-118, ISBN 978-960-474-131-1, ISSN 1790-2769, Genova, Italy, October 17-19, 2009.

12

Plasma-Assisted Ignition and Combustion

Andrey Starikovskiy[1] and Nickolay Aleksandrov[2]
[1]*Princeton University*
[2]*Moscow Institute of Physics and Technology*
[1]*USA*
[2]*Russia*

1. Introduction

The history of application of thermally-equilibrium plasma for combustion control started more than hundred years ago with IC engines and spark ignition systems. The same principles still demonstrate high efficiency in different applications. Recently particular interest appears in non-equilibrium plasma for ignition and combustion control [Starikovskii, 2005; Starikovskaia, 2006]. The reason of the interest rise is new possibilities for ignition and flame stabilization which are proposed by plasma-assisted approach. Over the last decade, significant progress has been made in understanding the mechanisms of plasma-chemistry interaction, energy redistribution in discharge plasma and non-equilibrium initiation of combustion. Wide range of different fuels has been examined using different types of discharges.

There are several mechanisms to affect a gas when using a discharge to initiate combustion or stabilize a flame. There are two thermal mechanisms: 1) gas heating due to energy release leads to an increase in the rates of chemical reactions; 2) inhomogeneous gas heating generates flow perturbations and provokes turbulization and mixing. Non-thermal mechanisms include 3) the effect of ionic wind (momentum transfer from electric field to the gas due to space charge-electric field interaction); 4) the ion and electron drift in the electric fields can lead to the additional fluxes of active radicals in the gradient flows; and 5) excitation, dissociation and ionization of the gas by electron impact leads to the non-equilibrium radical production and changes the kinetic mechanisms of ignition and combustion. These mechanisms combined together or separately can provide an additional authority to combustion control which is necessary for ultra-lean flames, high-speed flows, cold low-pressure conditions of high-altitude GTE relight, detonation initiation in pulsed detonation engines, distributed ignition control in HCCI engines and so on.

But, of course, the main interest in plasma technologies is connected with extreme conditions like hypersonic aviation. A hypersonic aircraft that cruises at Mach 6 speed and ~ 30 km altitude could travel 13,000 km in slightly over 2 hours, including the time needed for takeoff, ascent, descent and landing [Bowcutt, 2009]. This is significant because almost all major world city-pairs are separated by no more than 10,000-12,000 km. It is assumed that the ability to travel such long distances in a short time would be valuable for the delivery of time-critical cargo and for the relatively small percentage of passengers.

The advantage of short flight time, high speed and altitude would be promising to military aircraft and missiles. An air-breathing hypersonic vehicle operates in multiple engine cycles

and modes to reach scramjet operating speeds and could be used for the development of a space transportation system that would allow cheap and flexible access to space. A typical air-breathing hypersonic flight corridor with operation limits is presented in Figure 1 [Andreadis, 2005].

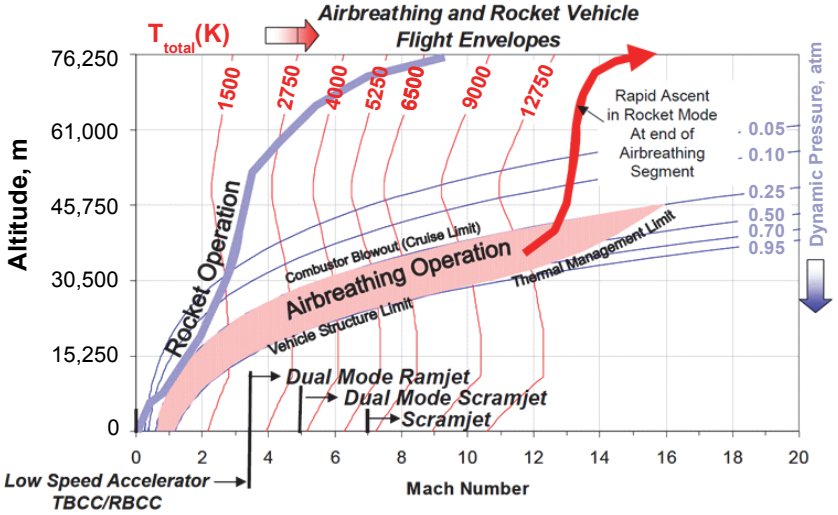


Fig. 1. Air-breathing hypersonic vehicle flight trajectory and operational limits [Andreadis, 2005].

Design challenges are dictated by flight conditions that become increasingly severe due to the combination of internal duct pressure, skin temperature, and dynamic pressure loading [Fry, 2004]. These constraints combination creates a narrow corridor of possible conditions suitable for flight based on ram air compression. The lower boundary of this envelope is set by thermal and structural limitations and is typically limited by a dynamic pressure of about 1 atm.

Relatively high dynamic pressure q is required, compared to a rocket, to provide adequate static pressure in the combustor. The upper boundary is characterized as a region of low combustion efficiency and narrow fuel/air ratio ranges thereby is restricted by combustion stability in the engines (dynamic pressure limit is 0.25-0.5 atm).

The high Mach number (M > 15) edge of the envelope is a region of strong leading shock waves, with strong dissociation and ionization of the gas in the shock layer. Here, nonequilibrium flow can influence compression ramp flow, induce large leading-edge heating rates, and reduce the efficiency of fuel injection and mixing and combustion. This lead to dramatic performance decrease and puts another limitation on the possible flight envelope for air-breathing hypersonic vehicles. On the contrary, at very low Mach numbers (M < 3) the compression ratio due to flow deceleration is not enough for efficient ramjet operation.

That is why in the low-speed regime (M = 0-3) the vehicle may utilize one of several possible propulsion cycles such as a Turbine Based Combined Cycle (TBCC) with a bank of gas turbine engines in the vehicle, or Rocket Based Combined Cycle (RBCC), with integrated

rockets, internal or external to the engine, to accelerate the vehicle from takeoff to M ~ 3 (Figure 2). Above this point only ramjet/scramjet cycle can propose efficient operation.

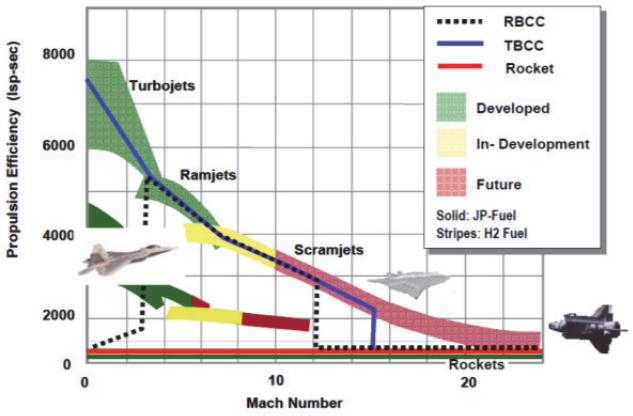


Fig. 2. Propulsion efficiency and operating regimes [Andreadis, 2005]

The concept of scramjet is rather simple; (i) it has a shaped duct with an air inlet at its front end and (ii) a constant-area or slightly divergent section where fuel is injected, mixed, and burned. The section sharply diverges at its aft end to form internal and external nozzles. Expansion of combustion products through the nozzle creates the thrust (Figure 3). The proper shaping of the duct to provide efficient air compression, fuel-air mixing and combustion, and gas expansion is a challenging task. To maximize the overall scramjet performance, the duct should operate in concert with the vehicle upon which it is integrated [Bowcutt, 2009]. This integration should take into account material limitations, fuel characteristics, range of speed, flight attitude, and atmospheric pressure.

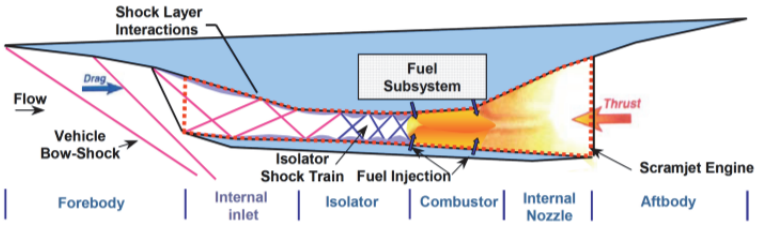


Fig. 3. Air-breathing supersonic combustion ramjet (scramjet), which provides a means of efficient hypersonic propulsion [Bowcutt, 2009].

The scramjet engine can operate as a dual-mode ramjet in the Mach 3 to 6 regime along the isolator capability limit to avoid inlet unstart and to remain within the structural limits (Figure 2) [Andreadis, 2005]. As the vehicle continues to accelerate beyond Mach 7, the combustion process is unable to separate the flow and the engine operates in scramjet mode

with a pre-combustion shock-free isolator. The inlet shocks propagate through the entire engine.

The efficiency of scramjet operation depends crucially on the efficiency and completeness of combustion process; low efficiency of combustion restricts the upper boundary of the flight envelope. Plasma assisted combustion is a promising way to stabilize ignition at high altitudes and at low dynamic pressures and temperatures (Figure 1). On the other hand, plasma assisted ignition can stabilize ultra-fast combustion that is required for M > 10 operations (Figure 2). Thus, the ultimate goal of plasma methods in hypersonic applications is an extension of the engine stability region to low dynamic pressures (0.2-0.05 atm) and to high Mach numbers (M = 10-20).

A review of recent experimental studies of non-equilibrium plasma assisted ignition and combustion is given in [Starikovskaia&Starikovskii, 2011]. There is also a detailed review of experimental and theoretical work in this field for H_2-O_2 mixtures [Popov, 2007]. Over the last years, considerable progress has been made in the understanding of the mechanisms of plasma assisted combustion in various mixtures including hydrocarbon-containing ones. The mechanisms were validated by making experiments under controlled conditions and by comparing the results with numerical simulation of discharge and combustion processes. The purpose of this work is to review the modern state-of-art in this area. We consider possibilities to generate chemically active discharge plasma with desirable parameters and mechanisms of the effect of non-equilibrium plasma on ignition and combustion. The difficulties in studies and unsolved problems are discussed. The number of works devoted to plasma assisted combustion is too large. Therefore, this survey does not tend to cite all available papers. Instead, we attempt to present only new results and to show the main obstacles for further progress in the use of non-equilibrium plasma in favoring combustion processes.

1.1 Demonstration of efficiency of plasma assisted combustion

There have been many impressive demonstrations of plasma-assisted ignition (PAI) advantages over the last decade [Starikovskii, 2005; Starikovskaia, 2006]. Control of ignition and combustion in fast gas flows is important since it is applicable for aircraft jet engines, where the problems of reduction of ignition delay time, of high-altitude flameholding and of extension of flame blow-off limits are of crucial importance. Plasma assisted ignition and flame holding in aerospace applications should cover an extremely wide range of parameters. The temperature range corresponds to regimes from high-altitude re-light (~220 K) to a full-load mode after gas compression (~800 K) and flame stabilization in afterburner (~1200 K). The pressure range is 0.1-10 atm, whereas the flow velocity is in the range 100-1000 m/s. Mixture composition varies from lean to rich in different locations. Here, we will mention some recent results only to demonstrate the progress under different flow conditions.

1.1.1 Supersonic flows

The paper [Leonov et al, 2011] studied particular mechanisms of the plasma effect on high-speed combustion. The TDLAS (tunable diode laser absorption) technique was applied for detection of temperature and H_2O concentration in the combustion zone of a supersonic (M = 2) air-fuel flow. Hydrogen and ethylene injected directly through the wall were used as a fuel. The combustion process was initiated and sustained by a near-surface electric discharge.

The discharge appeared in the form of oscillating plasma filaments as it is shown in Fig.4a, b. Initially, electrical breakdown occured near the electrodes. The individual filaments were blown down due to the main flow with velocity that was a bit less than the core velocity. The frequency of oscillations depended on flow speed, inter-electrode gap, and parameters of the power supply. In most cases this frequency was f=10-30 kHz. Plasma power was in a range of W_{pl} = 3-17 kW. Special experimental series showed a generation of sequence of active zones of reacting gas moving downstream from the place of immediate plasma-fuel interaction (see Figure 4).

Ignition in a supersonic propane-air flow at M=2 by a microwave discharge was reported in [Shibkov1,2]. The ignition was detected as a strong emission in the visible region of spectrum in the aerodynamic channel downstream of the discharge section. The induction time was simultaneously derived from different measurements, such as the control of CH emission at 431 nm, the measurements of the electric current through a plane capacitor at the outlet of the channel, and the control of output signal with an acoustic noise meter. The authors analyzed a direct current discharge (E/N=10-30 Td), a pulse-periodic electrode transversal discharge (30-70 Td), a freely localized microwave discharge (70-120 Td), and a surface microwave discharge (100-200 Td) and concluded that the discharges with higher reduced electric fields give shorter induction times for combustion.

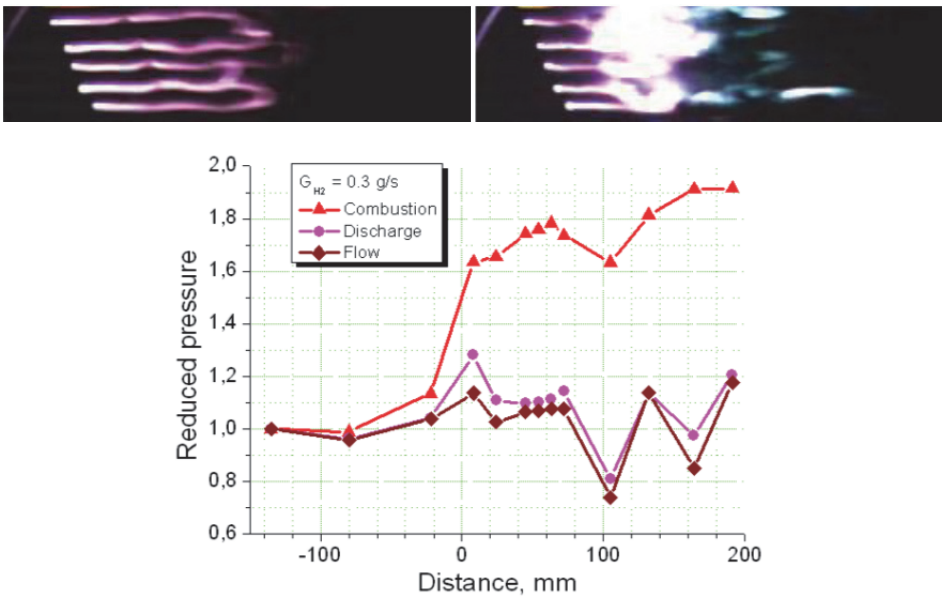


Fig. 4. Discharge appearance: (a) instant photo, flow from left to right; (b) discharge with ethylene injection; (c) evolution of wall pressure at hydrogen injection. Discharge power is W_{pl} = 8 kW. X = 0 is the electrode line. M = 2 [Leonov et al, 2011].

Another typical scheme of the experimental study of PAI in supersonic gas flows [Esakov et al, 2006] is given by Figure 5. The air flowed to the chamber through an M=2 supersonic

nozzle. Fuel or fuel-air mixture was injected through the pylon input. On the basis of temperature measurements, the authors [Esakov et al, 2006] made a conclusion that the combustion of propane-air mixtures occurs in an M=2 gas flow under the action of a microwave discharge with a power up to 200 W.

A nanosecond pulsed discharge located between two fuel jets was used to ignite and hold jet flames in supersonic crossflows, without the use of additional devices (e.g., cavities or backsteps) for flame holding [Do et al, 2010]. The fuel injection nozzles and discharge electrodes were mounted flush with the surface of the flat wall adjacent to the freestream flow. The nonequilibrium plasma was produced by repetitive pulses of 15 kV peak voltage, 10 ns pulse width and 50 kHz repetition rate. Sonic or subsonic fuel jets (hydrogen and ethylene) were injected into a pure oxygen freestream of Mach numbers M = 1.7–2.4. The shockwave/flow structures induced by the fuel jets and the OH radical distribution resulting from combustion were characterized by Schlieren photography and planar laser induced fluorescence imaging, respectively. A configuration combining an upstream subsonic oblique jet and a downstream sonic transverse jet was shown to provide adequate flow conditions for jet flame ignition assisted by the plasma discharge. The experimental results were interpreted using a simple model in which the pulsed plasma served as a source of reactive radicals added to a flammable gas mixture.

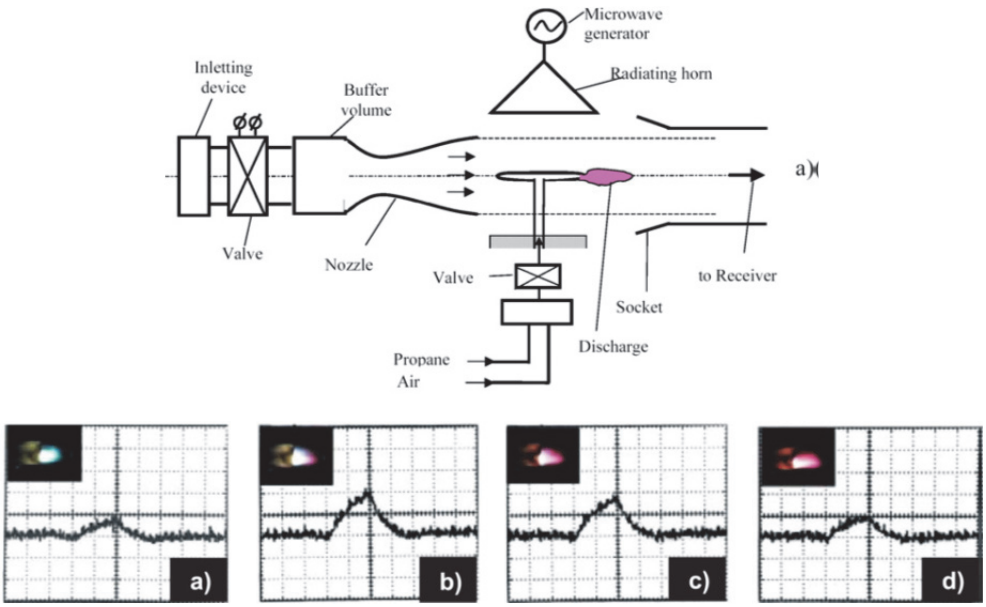


Fig. 5. Scheme of experimental setup for hypersonic microwave ignition (top) and the timeline of operation of experimental setup (bottom). Pressures and temperatures as well as times of discharge and fuel injection are marked in the scheme. Waveforms of stagnation temperature and emission from the discharge during combustion of a propane–air mixture for various air–fuel ratios, *r*. (*a*) $r = 0$, (*b*) $r = 5.9$, (*c*) $r = 14.4$, and (*d*) air flow without propane [Esakov et al, 2006].

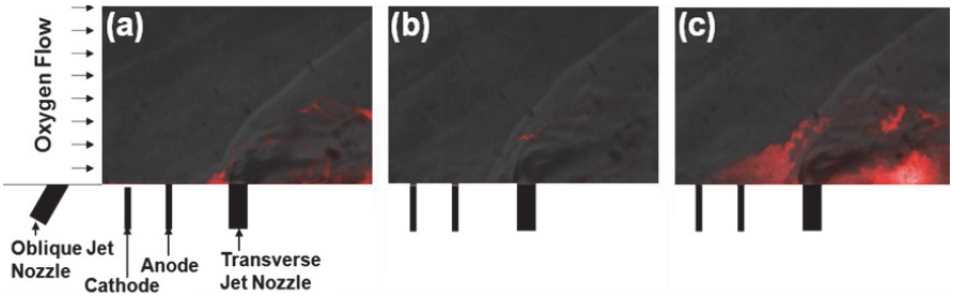


Fig. 6. The OH PLIF images overlapped with a typical Schlieren image of the same flow condition [Do et al, 2010]: (a) without the upstream jet, (b) with both jets active without plasma and (c) in the presence of the plasma with both jets.

Figure 6 depicts OH PLIF images superimposed onto Schlieren photographs with a pure oxygen crossflow. The left panel (Fig. 6a) corresponds to the case without the upstream jet and without the discharge, and the center panel (Fig. 6b) corresponds to the case with both the subsonic upstream and sonic downstream jet active, but without the discharge. In the absence of both the upstream jet and discharge, there is a moderate OH signal on the windward side of the main sonic jet in the recirculation region near the surface (Fig. 6a), induced by a strong bow shock. However, a rather weak fragmented flame detached from the surface was observed with both the jets activated (Fig. 6b), while the jet flame was significantly enhanced by the discharge plasma as shown in Fig. 6c [Do et al, 2010].

1.1.2 Subsonic flows

Experimental studies of the working processes in a low emission gas turbine combustor with plasma assistance have been conducted in [Serbin et al, 2011]. It was shown that the proposed concept of a gas turbine combustor with pilot flame stabilization by nonequilibrium plasma can provide higher performance, wider turn down ratios, and lower emission of carbon and nitrogen oxides. Simulation of the processes in a modified gas turbine combustor showed significant improvements of the combustor temperature distributions and ecological characteristics. Maximum temperature level of the combustion products (Figure 7) were decreased by 190 K (from 2340 up to 2150 K). As the result, the NO_x emission was reduced from 16 to 1 ppm without increase of carbon monoxide emission.

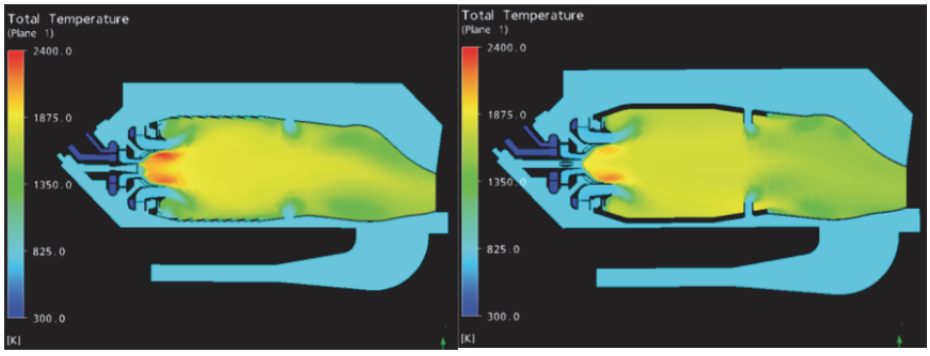


Fig. 7. Temperature field in the combustor: (*a*) – basic case; (*b*) – modified [Serbin et al, 2011].

A nanosecond repetitively pulsed plasma produced by electric pulses of amplitude 10 kV, duration 10 ns and frequency up to 30 kHz has been used to stabilize and improve the efficiency of a 25 kW lean turbulent premixed propane/air flame (Re_D = 30 000) at atmospheric pressure in [Pilla et al, 2006]. It was show that plasma generation in the recirculation zone of the flow led to a significant increase in heat release and in combustion efficiency., This allowed stabilization of the flame under lean conditions under which it would not exist without the plasma. Stabilization was obtained with a very low level of plasma power, about 75 W that accounts for 0.3% of the maximum power of the flame (Figure 8).
The domain of stability of the V-shaped flame without plasma is represented by region 1 in Figure 8. Region 2 represents the extended domain of stability of the V-shaped flame when the plasma is generated. Region 3 corresponds to a regime of intermittent flame, still with a relatively high level of heat release. Region 4 corresponds to the domain of the pilot flame. This pilot flame, being useful in reigniting the combustor, can exist over a much wider range of fuel equivalence ratios and flow rates than that without plasma [Pilla et al, 2006].

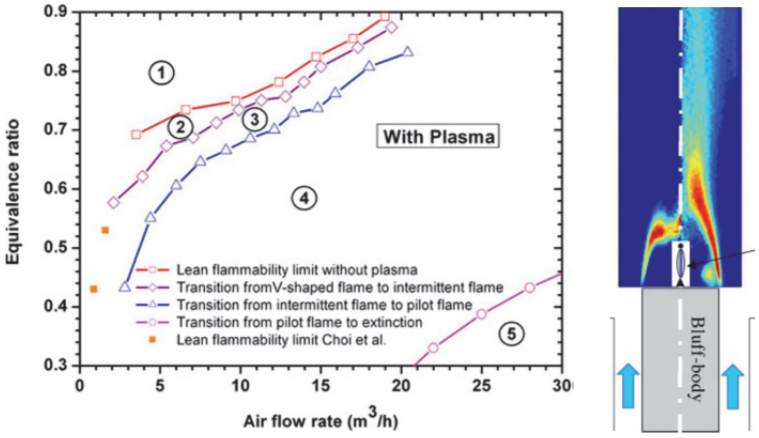


Fig. 8. (a) Burner regimes with plasma in filamentary regime. (b) Comparison of OH (A-X) emission with and without discharge. Emission of the plasma was blocked (mask location is shown by arrow). Flame power: 12.5 kW, discharge power: 75 W; air: 14.7 m^3h^{-1}, propane: 0.5 m^3h^{-1} (φ = 0.83) [Pilla et al, 2006].

1.1.3 Ignition control in quiescent gas

Ignition in methane/air mixtures has been achieved using low energy seed laser pulses and an overlapping subcritical microwave pulse [Michael et al, 2010]. It was shown that an extremely weak ionization by a laser localizes the microwave energy deposition and leads to rapid heating, high temperatures, and ignition. Multiple simultaneous localized regions of ignition were also achieved using the same microwave pulse. Interactions of the seed laser pulse and microwave heating pulse were observed using schlieren and shadowgraph to record the intensity of heating and the scale of the interaction and to confirm ignition. In addition, a coupled one-dimensional gasdynamic-plasma dynamic model has been developed to follow the rapidly evolving plasma properties and the gas properties achieved through this interaction. An increase in the femtosecond seed laser pulse energy resulted in

a significant change in the character of the ignition process. By increasing the pulse energy by a factor of 3, ignition along the seed laser path was observed. In this case, a much larger region of strong coupling with the microwave can be seen in the shadowgraph images, as shown in Figure 9 [Michael et al, 2010].

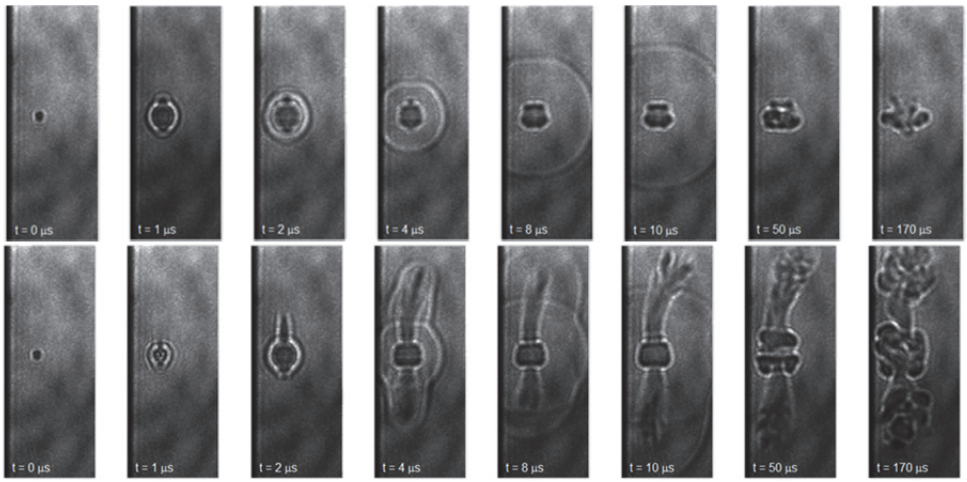


Fig. 9. Ps Laser-MW Evolution in Air. a) Laser spot evolution b) Laser + MW evolution [Michael et al, 2010].

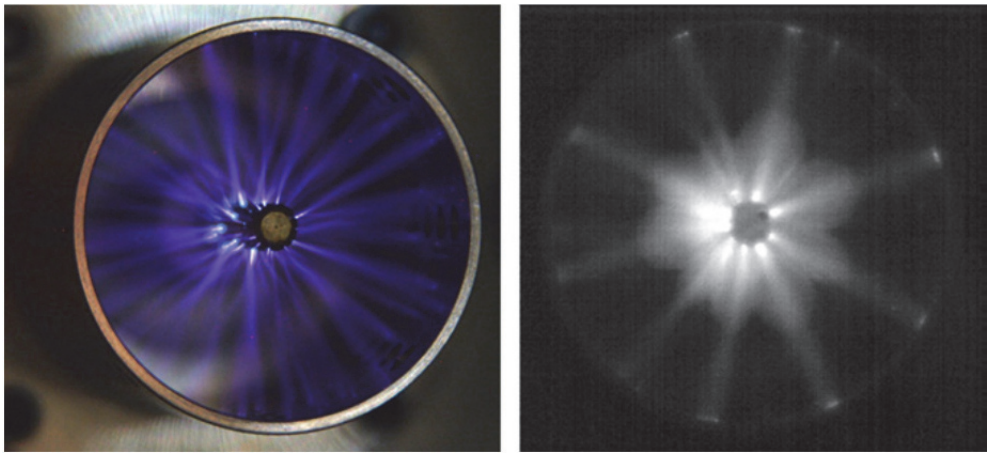


Fig. 10. Left: streamers generated by a single 370 mJ, 56 kV, 54 ns pulse (maximum $E/n \approx 400$ Td) in air (10 s gate time); Right: flame propagation from multiple ignition sites at the base of the streamers after a single pulse in $\varphi = 1.1$ C_2H_4–air mixture (1 ms gate time). Both images were captured from the same off-axis angle [Singleton et al, 2011].

Transient plasma ignition, involving short ignition pulses (typically 10–50 ns), has been shown to effectively reduce ignition delays and improve engine performance for a wide

range of combustion-driven engines relative to conventional spark ignition. Two distinct phases of the plasma-ignition process were demonstrated in [Singleton et al, 2011]: an initial non-equilibrium plasma phase, wherein energetic electrons transfer energy into electronically excited species that accelerate reaction rates, and a spatially distributed thermal phase, that produces exothermic fuel oxidation reactions that result in ignition. It is shown that ignition kernels are formed at the ends of the spatially separated streamer channels, at the cathode and/or anode depending on the local electric field strength, and that the temperature in the streamer channel is close to room temperature up to 100 ns after the discharge [Singleton et al, 2011].

The results presented in Figure 10 show that after a transient plasma discharge in a φ = 1.1 C_2H_4–air mixture at 1 atm, (i) ignition occurred within the streamer channel, (ii) flame initiation occurred within 1 ms of the discharge and (iii) flame propagation was faster than that initiated by conventional spark ignition. Therefore, when considering the mechanism of ignition via a transient plasma discharge, it is relevant to consider how the energy is deposited and over what timescale, and how that impacts ignition. In these studies, the electrical pulse delivered approximately 70 mJ to the mixture in 12 ns, distributed across many streamer channels (typically ≈ 50). Therefore, ~1.5 mJ was applied to the fuel–air mixture in each streamer, which was sufficient for ignition [Singleton et al, 2011].

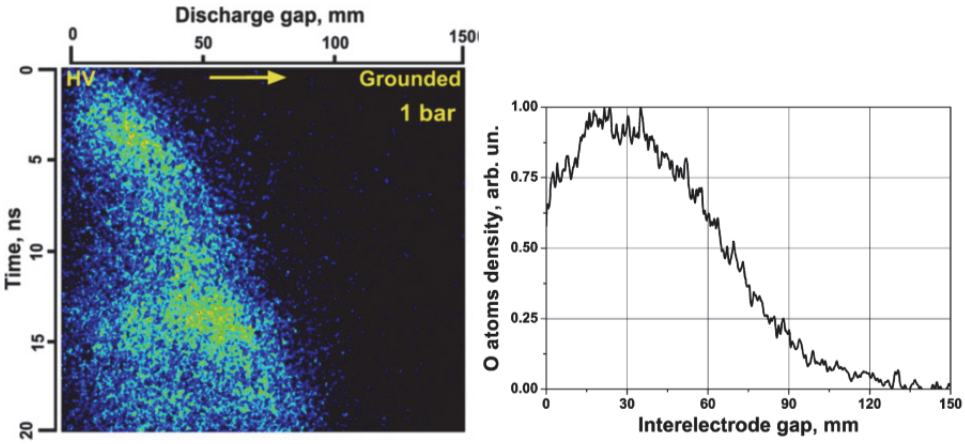


Fig. 11. A) Streak image of discharge development in air at 1 bar. Streamer velocity: 3.5 mm/ns. B) O atom density distribution over the interelectrode gap [Rakitin et al, 2011].

An experimental study of detonation initiation by high–voltage nanosecond gas discharge has been performed in smooth detonation tubes in [Rakitin&Starikosvkii, 2008], [Rakitin et al, 2011]. A gradient mechanism was used to initiate detonations in stoichiometric propane–oxygen mixtures with different nitrogen dilution and in propane–air mixtures. Initial pressures from 0.2 to 1 bar have been tested. Detonation was formed within 4 transverse tube sizes at initial pressures higher than 0.2 bar for the propane–oxygen mixture and higher than 0.8 bar for the diluted mixture with 40% of nitrogen. The discharge energy inputs were 0.2—0.3 J (Figure 11). The gradient mechanism of

detonation formation similar to the one suggested by Zeldovich has been shown to be the governing process.

The streamer mode of discharge propagation without the formation of a hot channel resulted in a fast DDT. The results are presented in terms of a time-resolved frame sequence captured with the UltraSpeedStar16 camera in Fig. 12,a and in terms of an x–t diagram in Fig. 12,b. From the IR sensor data, it is seen that the flame front already propagated at the CJ velocity between sensors 1 and 2 (40 and 120 mm from the discharge tube). The DDT time was amounted to 150 μs. At the same time, ICCD imaging showed that the ignition did not occur simultaneously over the channel.

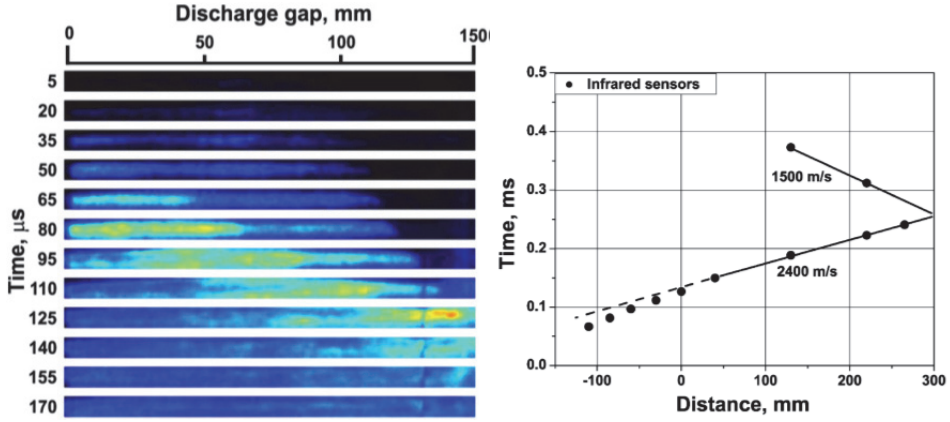
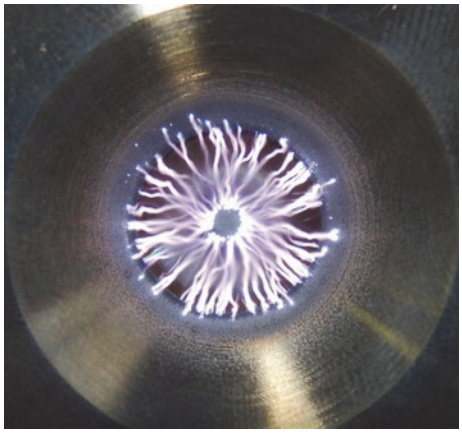


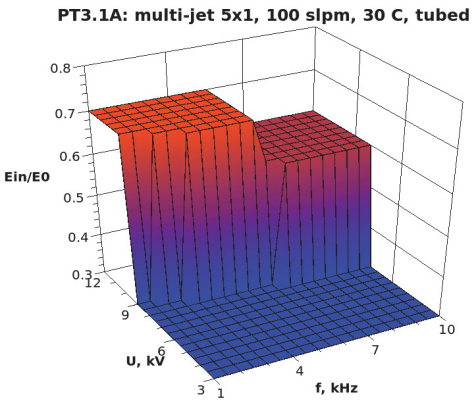
Fig. 12. a) Time-resolved ICCD imaging of fuel mixture ignition inside the discharge chamber. DDT at an initial pressure of 1 bar, streamer mode. b) x–t diagram of DDT at an initial pressure of 1 bar, streamer mode [Rakitin et al, 2011].

A study of pulsed high–voltage nanosecond discharge development in a series of plasmatrons has been conducted in [Nikipelov et al, 2011]. The discharge exhibited three modes of development depending on frequency, voltage and mass flow rate: a surface streamer, a localized spark, and a distributed nonequilibrium transient spark. The plasmatrons developed have been used to stabilize ultra–lean (f = 0.06-0.3) flames in a wide range of equivalence ratios and temperatures for methane and diesel vapour at 1 bar. The plasmatrons demonstrated exceptional flame stability with an average discharge power less than 20 W for a total power of the burner higher than 1 kW for ultra–lean flame conditions in methane–air mixtures.

Figure 13 shows the discharge development map. Three distinct modes were observed depending on the frequency, voltage and mass flow rate. In a "streamer" mode, the discharge did not bridge the gap mainly due to insufficient pulse voltage. Instead, it developed as a streamer corona starting from the high voltage electrode tip. Under higher voltage, the discharge was either localized at one or two points across the nozzle ("localized spark"), or was distributed over the cross–section ("distributed spark", Figure 13,a), developing as a nonequilibrium transient spark. Each mode was associated with a certain typical value of specific energy input into the discharge, which was measured with the back–current shunt technique. In the streamer mode, a larger part of the pulse was reflected off the discharge gap, while 30% of the incident energy was deposited into the gas. The

distributed spark deposited 60% of the energy on average, whereas the localized spark did 70%. The switching of the modes is shown in Figure 13,b in terms of voltage–frequency ($f - U$) maps. Different modes corresponded to different values of specific energy input E_{in}/E_0.





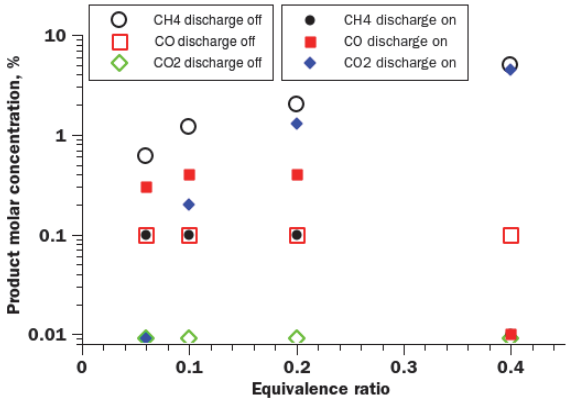


Fig. 13. a) Plasmatron with distributed ns spark. b) Voltage–frequency map of repetitive nanosecond discharge development mode. Inlet air temperature: 300 K. Mass flow rate 100 slpm. c) Plasma–assisted stabilization of premixed methane–air flame. Inlet temperature: 850 K; flow rate: 60 slpm [Nikipelov et al, 2011].

Results for premixed methane–air flow ignition initiation with inlet temperature of 580° C in terms of product composition downstream from the burner [Nikipelov et al, 2011] is shown in Figure 13,c. The gas flow rate was 60 slpm, the equivalence ratios ranged from 0.06 to 0.4. For 850 K, it was show that the discharge initiates complete fuel oxidation even for f = 0.06 (90—95% of methane was oxidized) into CO and CO_2.

A nanosecond surface barrier discharge was proposed as an igniter at high pressures [Kosarev et al, 2009]. ICCD imaging of a nanosecond DBD in dry 1-5 atm air has been taken.

Peculiarities of discharge uniformity at different gas pressures were discussed. The ability of the pulsed nanosecond discharge to produce a thin quasi-uniform plasma layer at elevated pressures was demonstrated (Figure 14,a).

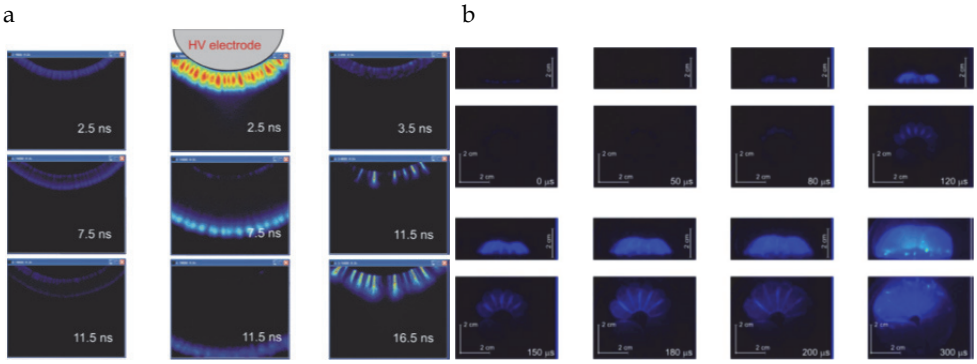


Fig. 14. a) ICCD images of the discharge at 1 atm dry air. Negative polarity of the high-voltage electrode, 22 kV, 25 ns duration, ϕ = 40 Hz [Kosarev et al, 2009]. b) Camera gate is 0.5 ns. Single pulse sliding DBD ignition of C_2H_6:O_2=2:7 mixture at 1 bar and ambient temperature [Sagulenko et al, 2009].

Successful demonstration of high-pressure distributed ignition in a single-pulse SDBD was done in [Sagulenko et al, 2009]. Mixture C_2H_6:O_2=2:7 at 1 bar and ambient initial temperature was successfully ignited in ~100 μs in relatively large volume (Figure 14b).
Thus there are successful demonstrations of plasma-assisted ignition and combustion control over very wide range of parameters – T = 300-1500 K, P = 0.1-40 atm, from ultra-lean to reach fuel conditions, flow speed up to M = 2. Nevertheless, to scale-up experimental installations we need to know the physical and chemistry underlying. Next section is devoted to mechanisms description which control the plasma assisted ignition and combustion.

2. Physics of plasma-assisted combustion

2.1 Energy branching in discharge plasma

The major difference between common combustion and plasma assisted combustion is an extremely non-equilibrium excitation of the gas in discharges. In this case, electrons gain energy from an external electric field and transfer this energy through collisions into the various degrees of freedom of other particles. Energy exchange between electrons and translational degrees of freedom of molecules is very slow because of a great difference in masses of the collading particles. This means that electron energy is spent predominantly on the excitation of internal degrees of freedom of molecules. If the rate of internal energy relaxation is not high, the population of the excited states of the molecules is far from the equilibrium energy distribution. Overpopulation of excited states leads to an increase in the system reactivity and facilitates ignition and flame propagation. From this point of view, the most important question for plasma-stimulated chemistry is the discharge energy branching ratios for different degrees of freedom of molecules, rate of system relaxation (thermalization) and chemically-active system response to this non-equilibrium excitation.
The rate of molecule excitation by electron impact in discharge plasma depends on the electron energy. The lowest energy is needed for excitation of rotational degrees of freedom. Typical

spacing between rotational levels of simple molecules is ~10-100 K and electron energy ~300 K (~0.03 eV) is enough for efficient rotational excitation. Typical quantum of vibrational degrees of freedom is in the range 1-3 kK. This means that, for efficient vibrational excitation, the average electron energy should be comparable with or higher than this magnitude (in air it should be in the range of 0.2-2 eV). Excitation of electronical degrees of freedom and molecule dissociation require energies of 3-10 eV. When the average electron energy exceeds 10 eV, ionization becomes one of the dominant electron impact processes. Thus the possibility to control the electron energy means the possibility to control the direction of energy deposition and selective excitation of different degrees of freedom of molecules.

Average electron energy in gas discharge is determined by a reduced electric field E/n, where E is the electric field and n is the gas number density [Raizer, 1991]. Figure 15 illustrates the deviation of the characteristic electron energy, D/μ, from the temperature of molecules, T, in different gases. Here, D is the diffusion coefficient of electrons and μ is their mobility. The critical E/n for the noticeable difference between D/μ and T and, consequently, for the formation of a non-equilibrium electron energy distribution is close to $E/n \sim 0.1$ Td for atomic gases and to $E/n \sim 1$ Td for molecular gases (1 Td = 10^{-17} V × cm^2).

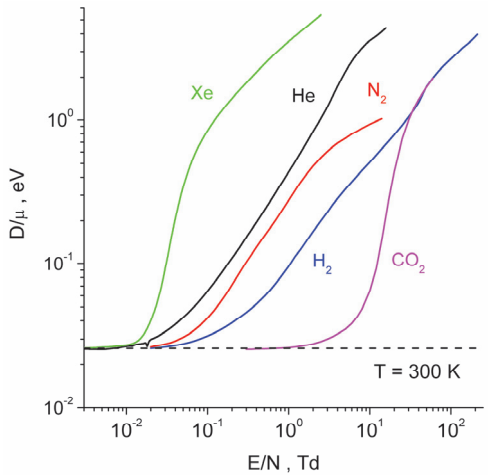


Fig. 15. Characteristic electron energy. He and Ar – [Dutton, 1975]; H_2, N_2 and CO_2 – [Huxley&Crompton, 1974].

Nonequilibrium electron energy distribution function (EEDF) can be found from a solution of the Boltzmann equation. In the simplest case, the EEDF is time- and space-independent and depends only on the local reduced electrical field, E/n, and gas composition [Babich, 2005]. Further simplification is possible using the so-called two-term approximation in which the EEDF is presented in the form $f(v) = f_0(v) + f_1(v)\cos\theta$ [Raizer, 1991]), where v is the electron velocity and θ is the angle between the electron velocity and the ambient electric field.The input data for a solution of the electron Boltzmann equation are the cross-sections of elastic and inelastic collisions between electrons and neutral particles. Figure 16 shows self-consistent sets of electron cross-sections for CH_4, O_2 and Ar; electron transport and rate coefficients calculated using these data agree well in the pure gases with available measurements.

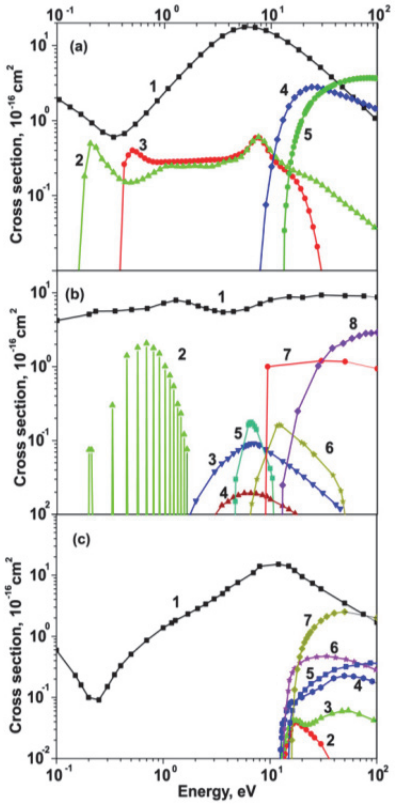


Fig. 16. Electron collision cross sections for (a) CH_4 [Hayashi, 1987], (b) O_2[Ionin et al, 2007], and (c) Ar [Tachibana, 1989] as a function of electron energy.
CH_4: (1) momentum transfer in elastic collisions, (2)–(3) vibrational excitation,
(4) electronic excitation, and (5) ionization.
O_2: (1) momentum transfer in elastic collisions, (2) vibrational excitation, (3) excitation of the $a^1\Delta_g$ electronic state, (4) excitation of the $b^1\Sigma_g^+$ electronic state, (5) excitation of the electronic states with ΔE = 4.5 eV, (6) excitation of the electronic state with ΔE = 6 eV,
(7) excitation of the electronic state with ΔE = 8.4 eV, and (8) ionization.
Ar: (1) momentum transfer in elastic collisions, (2) excitation of the $^1s_5$ electronic state,
(3) excitation of the $^1s_4$ electronic state, (4) excitation of the $^1s_2$ electronic state,
(5) excitation of other allowed electronic states, (6) excitation of other forbidden electronic states, and (7) ionization.

Using the EEDF, it is possible to calculate the energy branching ratios for the different degrees of freedom of neutral particles. Figure 17 demonstrates the calculated fractional power dissipated by electrons into different internal degrees of freedom in various gas mixtures as a function of *E/n* in discharge plasma. At very low *E/n* (~ 0.1 Td) rotational excitation is dominated; here up to 85% of the discharge energy goes to this excitation. Fast energy exchange between rotational and translational degrees of freedom leads to equilibrium gas heating in this case. An increase of the *E/n* value to 0.4 Td changes the

priorities. Above this point the main channel of electron energy dissipation becomes vibrational excitation of oxygen. For reduced electric fileds 4 Td < *E/n* < 110 Td, the most efficient channel of electron energy dissipation in air is vibrational excitation of nitrogen (Figure 17,a). Vibrational-translational (VT) relaxation is a rather slow process at low gas temperatures and the vibrational temperature in discharge plasma could be much higher than the translational one. In the same region of *E/n* ,excitation of the lowest electronic level of oxygen, $O(a^1\Delta)$ state, takes place. The efficiency of this electronic excitation is very small (~2%) in the presence of nitrogen. But a low rate of $O(a^1\Delta)$ state oxygen quenching leads to its concentration increase under some conditions. It should be noted that an electric field value *E/n* ~ 120 Td is a very important threshold in air. Above this point the electric filed is enough to ionize the gas and discharge can propagate in self-sustained regime. Below this value discharge can exist only in the presence of external source of ionization. At *E/n* from 140 to 500 Td the main channel of energy losses is excitation of electronic triplet states of nitrogen. Because of high electrons energy in this range of *E/n* (from 3 to 10 eV, correspondingly) ionization of the gas in discharge gap is very fast. Between 500 and 1000 Td the excitation of nitrogen singlet states becomes most important, and above 1000 Td the main portion of the electron's energy goes to ionization of the gas.

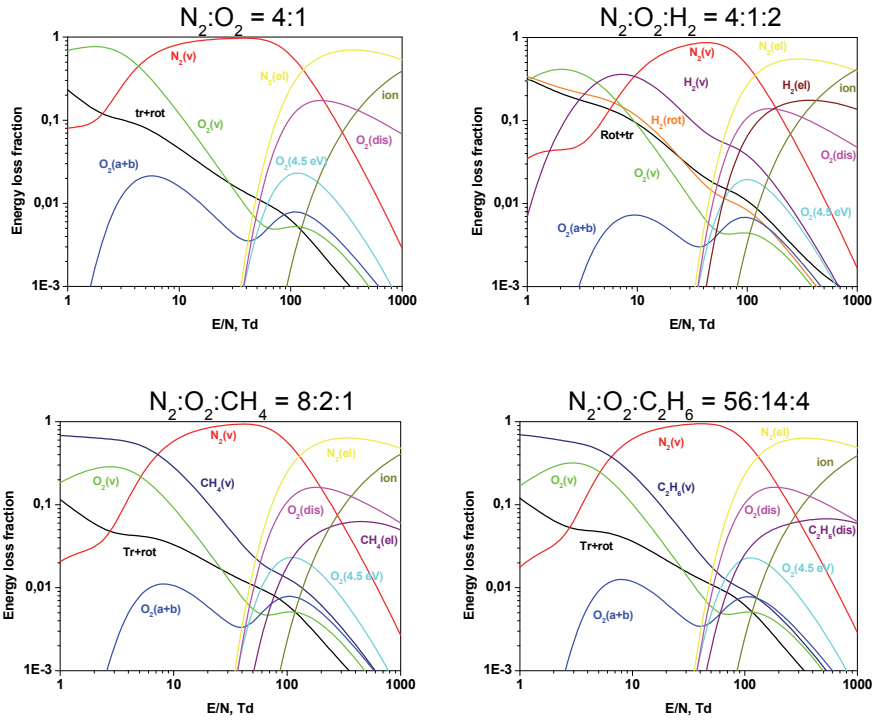


Fig. 17. Fractional power dissipated by electrons into molecular degrees on freedom as a function of *E/n*. (a) Air; (b) H_2-Air; (c) Methane-Air; (d) Ethane-Air stoichiometric mixtures.

Fuel addition in stochiometric ratio (Figures 17 (b,c,d)) do not change this picture dramatically. The main reason is relatively small concentration of fuel molecules in the mixtures under typical combustion conditions. Figure 17(b) demonstrates the effect of hydrogen on the electron energy branching. The stoichiometric mixture H_2-air (H_2:O_2:N_2 = 29:14.5:56.5) contains about 30% of hydrogen. It is follows from the calculations that these additives only slightly change the energy distribution at moderate and high E/n > 20 Td. Excitation of molecular nitrogen remains the main process. Vibrational and electronic excitation and ionization of hydrogen only slightly change the electron energy branching (Figure 17(b)). At low E/n < 10 Td, the effect of hydrogen addition becomes more significant. Vibrational excitation of hydrogen is the main channel of energy dissipation for E/n = 5-10 Td. The role of rotational excitation of hydrogen is also important and increases the energy flux into rotational and translational degrees of freedom at low E/n (Figure 17(b)). Almost the same picture is obtained in the stochiometric mixtures with hydrocarbon fuels, methane (Fig. 17(c)) and ethane (Fig. 17(d)).

Thus, the reduced electric field is an important parameter for plasma assisted combustion. The E/n value controls the direction of energy deposition in discharge and governs the composition of active particles produced in the discharge plasma.

In the simplest case, E/n is a function of the gas density, applied voltage and geometry of the discharge cell. Ignition and flame stabilization in different engines require different initial conditions. Gas temperature can vary from T~220 K (high-altitude re-light of the engine) to the value at the end of compression cycle (T~700-800 K) and temperatures corresponded to the combustion stage (1000-2500 K). Typical gas density varies from 10^{17}−10^{18} cm^{-3}, which corresponds to conditions of hypersonic flight, and may be as high as (1-3)×10^{20} cm^{-3} for high pressure turbines or in perspective automobile engines. At ambient temperature, this corresponds to a pressure range from a few Torr to tens of bar.

Figure 18 schematically demonstrate the variety of gas discharges. It was mentioned above that the E/n value in the discharge dictates the type of gas excitation. Electron density determines the power density in plasma:

$$P = q * E = n_e \times v_{dr} \times E = n_e \times n^2 \times \mu_e(E/n) \times (E/n)^2$$

where v_{dr} – electrons drift velocity in plasma; $\mu_e(E/n)$ – electron's mobility at given reduced electric field; n_e – electron's concentration. Thus the energy power density in plasma depends on gas density, E/n value, and electrons concentration.

That is why we choose for discharge classification two variables: E/n and n_e (Figure 18). It should be noted, however, that the same type of excitation under different conditions (for example, different pressures, initial gas temperatures) could lead to different results. Some types are "mix" of several types of plasma and different zones of these discharges are characterized by different E/n values. Nevertheless we can trace some common discharge features using this classification. Arc discharges develop at relatively low E/n. Main direction of energy deposition is rotational excitation of the gas [Raizer, 1991]. Steady-state MW discharges under high-pressure conditions consist of number of filaments of almost equilibrium relatively hot plasma [Raizer, 1991].

Pulsed gliding arks combines during different stages of their development some features of glow discharges and arks. Reduced electric field value changes from 10 Td to 100 Td and gas excitation is not very selective. Both gas heating through fast relaxation of rotational degrees of freedom and molecules dissociation by e-impact are effective. That is why this type of discharge sometimes is called as "warm" plasma [Fridman, 2008].

Typical *E/n* value in DC glow discharge is close to breakdown threshold (~100 Td in air) and main channels of gas excitation are vibrational and electronic levels population by e-impact [Raizer, 1991]. Streamer discharges (with or without barrier on the electrodes) demonstrate very high *E/n* value in the front of ionization wave in the streamer head (*E/n* ~ 500 Td) while in the streamer's channel *E/n* value is close to 20-30 Td [Pancheshnyi et al, 2005]. That is why in streamer discharge the energy distribution depends on the geometry of the discharge gap, length of the streamer channel etc. Streamer discharge could be formed if we have a significant (1.1-2 times) overvoltage in the discharge gap. Surface dielectric barrier discharge (SDBD) is characterized by even higher *E/n* because of very short effective inter-electrode gap. High overvoltage leads to significant energy release to gas ionization [Roupassov et al, 2008]. Extremely short high-voltage pulse rise time leads to formation of so-called fast ionization waves (FIW). For limited time (typically several nanoseconds at n ~ 10^{18} cm^{-3}, shorter at higher gas density) it is possible to maintain *E/n* on the level close or above the run-away threshold (*E/n* ~ 2 kTd in air). High-energy electrons are formed in the discharge region providing homogeneous pre-ionization and excitation of the gas. Discharge develops as a uniform ionization wave propagating at the velocity comparable to local speed of light [Pancheshnyi et al, 1999].

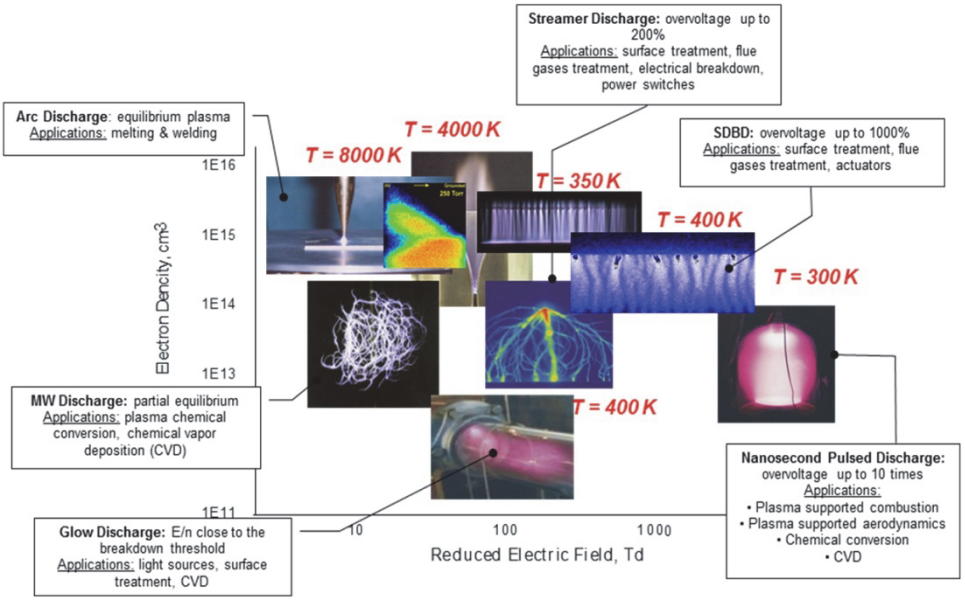


Fig. 18. Discharge development at different overvoltage and plasma generation.

Precise control of the direction of energy deposition in the plasma is possible in relatively short time scale when electrons multiplication (above breakdown threshold) or recombination (below threshold) have no enough time to change electrons concentration and plasma conductivity significantly. When plasma conductivity is not very high it is possible to maintain in inter-electrode gap desired value of *E/n*. This approach together with

combination of short high-voltage pulse and constant bias allow to provide selective and extremely nonequilibrium excitation of the gas. Critical high-voltage pulse duration depends on the gas parameters (density, composition) but for practically important range of parameters is restricted to few nanoseconds.

Thus the possibility of selective excitation of the gas by electric discharge critically depends on the possibility of ultra-short high-voltage pulses generation. Figure 19 demonstrate recent progress of solid-state generators based on "turn-on" FID and "turn-off" DRD switches according to FID GmbH data [Efanov et al, 2011]. In modern pulsers the pulse rise time goes down to 80 ps, voltage rise rate reaches 1 MV/ns, maximal voltage 2-10 MV, and maximal current up to 100 kA. Wide range of possibilities proposed by current progress in solid-state electronics will lead to the increase of our abilities of nonequilibrium plasma generation with predicted properties.

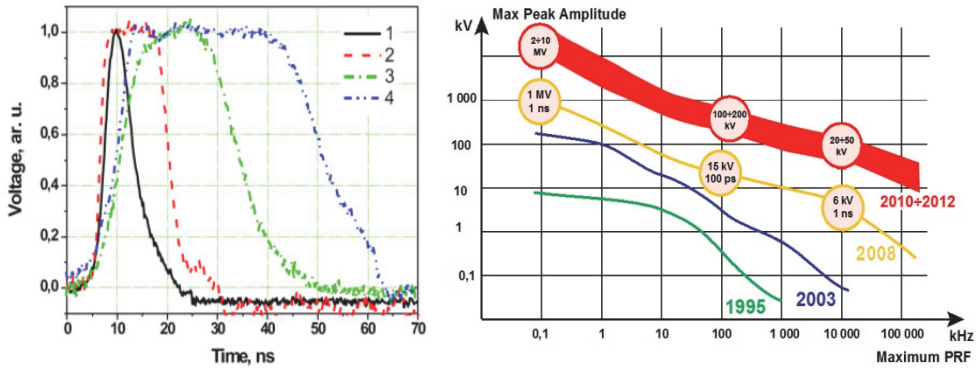


Fig. 19. Progress in ultra-short high-voltage pulse generators. A) typical nanosecond pulse shapes [Roupassov et al, 2008]; b) generators frequency-voltage map [Efanov, 2011].

2.2 Non-equilibrium plasma recombination and energy relaxation

For efficient production of large amount of active particles in the gas discharge it is necessary both efficient generation in the gas discharge plasma and slow recombination in collisions with major mixture components.

2.2.1 Rotational relaxation

Due to fast rotational-translational (RT) relaxation, rotational degrees of freedom of the molecules are quenched rapidly. This process requires few collisions only. For example, for rotational relaxation in air $O_2(rot) + M \rightarrow O_2 + M$ and $N_2(rot) + M \rightarrow N_2 + M$ typical relaxation time is comparable with gas-kinetics time. This means that typical time of rotational states thermalization is $\tau \sim 0.5$ ns under normal conditions. That is why rotationally-excited molecules cannot be considered as active particles for non-thermal acceleration of chemical reactions. Another important point is that the energy of excitation of rotational states is very small (roughly equal to translational temperature) and is significantly lower than typical chemical reaction's thresholds. From the other hand, it is possible to heat the gas through the rotational degrees of freedom excitation.

2.2.2 Vibrational relaxation

Opposite to rotational states relaxation, quenching of vibrationally excites states N_2 and O_2 (vibrational-translational (VT) relaxation) is very slow process. Time of VT relaxation usually is longer than typical time of plasma-assisted ignition (~10-100 μs). These times become comparable when significant amount of H_2 or hydrocarbons is presented in te mixture. This means that the vibrationally-excited N_2 and O_2 molecules can be accumulated in the discharge with intermediate E/n values.

VT relaxation leads to slow thermalization of vibrational energy of the molecules. This process becomes faster if the mixtures contain hydrocarbons. For example, VT relaxation of molecular oxygen on methane in stoichiometric methane-air mixture at T = 1000 K and pressure 1 atm has a characteristic time t ~ 1.3 μs. Fast relaxation does not allow to maintain a significant deviation of vibrational temperature from translational on the long time scale. From the other hand, VT relaxation of oxygen in H_2-air mixture lasts ten times longer and reaches t ~ 15 μs for T = 1000 K and P = 1 atm (29% H_2 in the mixture). VT-relaxation of hydrogen in the same mixture takes approximately 380 μs. Thus, vibrational excitation of hydrogen molecules can be very far from equilibrium during the ignition delay time and can effect significantly the radical's production.

Under uncompleted vibrational relaxation conditions chemical reactions between vibrationally excited molecules play an important role. There are several theoretical models for rate coefficients of reactions between excited reagents. Almost all these models were developed as an engineering substitution of time-consuming *ab initio* calculations [Kovach et al, 2010; Adamovich et al, 1996; Macheret et al, 1994; Park, 1988]. A model of vibrational energy usage was developed in [Losev et al, 1996]. The model assumes the decrease of the reaction threshold by αE_{vib}. The efficiency of vibrational excitation α can be estimated using activation energy and thermal effect of the reaction. A model proposed by Macheret [Macheret et al, 1994] allows to estimate the rate constant of simple exchange endothermic reaction. The model requires the fraction of energy release in the reverse reaction directed to vibrational excitation and it is applicable only to a certain type of reactions [Kovach et al, 2010].

It should be noted that almost all analytical models available estimate reaction rate constants using "vibrational temperature". This assumes that we have Boltzmann distribution over vibrational levels. Such an approach cannot be used at non-equilibrium conditions when the population over vibrational levels has non-Boltzann shape [Capitelli, 1996]. State-to-state model was considered in [Starikovskii, 2003]. Reactions between excited hydrogen molecules $H_2(v)$ and radicals are extremely important for ignition and combustion. As an example of reaction rate dependence on the vibrational excitation of reagents let us consider the process

$$H_2(v) + O \rightarrow H + OH(w)$$

It is shown in [Light& Matsumoto, 1978] that ratio of specific constants of the reaction rates at v=1 and v=0 is k(v=1)/k(v=0) = 2600 at T=300 K. The process at v=1 leads to formation of radical OH in vibrationaly excited state [Light& Matsumoto, 1978]

$$H_2(v=1) + O(^3P) \rightarrow H + OH(w=1) \qquad (k = (1.0^{+0.4}_{-0.6}) \cdot 10^{-14}\ \mathrm{cm^3/s})$$

$$H_2(v=1) + O(^3P) \rightarrow H + OH(w=0) \qquad (k \leq 4.7 \cdot 10^{-15}\ \mathrm{cm^3/s})$$

Experimental measurements show that the averaged factor of vibration energy usage in this reaction is α = 0.31 [Rusanov&Fridman, 1985]. Figure 20,a shows results of calculation of the

reaction rate constant at translation temperature T_{tr} = 300 K for various vibration temperatures with the Boltzmann distribution of molecules over vibrational levels. The dependence calculated using model [Starikovskii, 2003] is in good agreement with calculation by α-model with experimentally found α = 0.31 at overheating degree T_{vib}/T_{tr} < 5 (Fig. 20,a). Model [Starikovskii, 2003] predicts the ratio k(v=1)/k(v=0) = 2795, which is in perfect agreement with experiments [Light& Matsumoto, 1978] (2600). Ratio of channels to OH(w=1) and OH(w=0) at T = 300 K estimated in [Starikovskii, 2003] is equal to k(w=1)/k(w=0) = 7.9, which also is in good agreement with experiments [Light& Matsumoto, 1978] (>2).

The analysis of the reaction rate constant dependence on the vibrational excitation degree for reaction OH + H_2(v) → H_2O + H is shown on Figure 20,b. The predictions of model [Starikovskii, 2003] are in a good agreement with calculations based on experimentally measured value of α = 0.24 [Fridman&Rusanov, 1985]. Work [Light& Matsumoto, 1978] gives an experimental estimation for ratio of the rate constants of processes OH + H_2(v=0) → H_2O + H and OH + H_2(v=1) → H_2O + H: $k_{v=1}/k_{v=0} \le 1000$ at T=298 K, which is in good agreement with the estimation by model [Starikovskii, 2003] $k_{v=0}$ = $2.9 \cdot 10^{-15}$ $cm^{-3}s^{-1}$, $k_{v=1}$ = $1.8 \cdot 10^{-12}$ $cm^{-3}s^{-1}$ ($k_{v=1}/k_{v=0}$ = 620).

Thus, vibrational excitation of reagents can significantly accelerate chemical reactions. The influence of vibrational excitation is limited by VT-relaxation of the molecules. This process becomes extremely fast in the presence of hydrocarbons. In mixtures with hydrogen the efficiency of vibrational excitation increases because of relatively slow vibrational relaxation of H_2. Analysis of [Zatsepin et al, 2001] shows the oxidation rate increase in H_2-air mixture at T = 300K in 3-5 times.

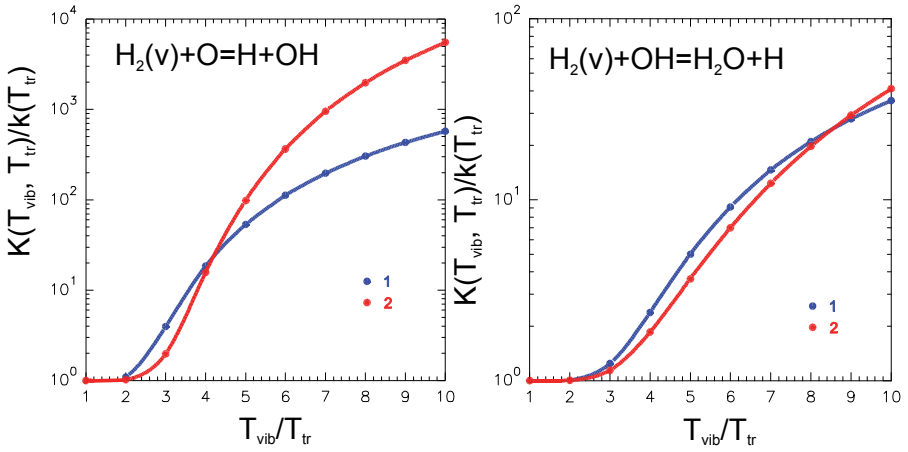


Fig. 20. Dependence of the rate constant of reaction on non-equilibrium excitation degree T_{vib}/T_{tr} at T_{tr} = 300 K. 1 – model [Starikovskii, 2003]; 2 – α-model [Macheret et al, 1994].

a) H_2(v) + O → H + OH, α = 0.31; b) H_2(v) + OH → H_2O + H, α = 0.24.

As an example of possible applications of vibrational excitation of the flow we will mention the paper [Bezgin et al, 2006]. Peculiarities of an oblique detonation wave formation in a supersonic hydrogen–oxygen mixture flow over a plane wedge were numerically analyzed.

Preliminary excitation of molecular vibrations of H_2 was shown to lead to a noticeable decrease in the induction zone length and the distance at which the detonation wave was formed. It was demonstrated that the reason for these effects was an intensification of chain reactions in the H_2–O_2 (air) mixture owing to the presence of vibrationally excited hydrogen molecules in the flow [Bezgin et al, 2006].

2.2.3 Electronic levels excitation and relaxation

At E/n ~ 100-500 Td the main channel of gas excitation is population of electronic degrees of freedom by electron impact and by energy exchange between vibrationally-excited states. An important exception from this rule is singlet state of molecular oxygen $O_2(a)$. This state has a low excitation threshold and the maximum efficiency of its population corresponds to E/n ~ 3-10 Td.

There are number of different electronically-excited particles in low-temperature plasma. Unfortunately, reaction rate constants, quenching rates and products are known only for limited number of them. That is why we will mention here the most important levels from the point of view of plasma assisted combustion only.

The efficiency of energy of electronic state usage for plasmachemistry depends on the ratio between channels of depopulation of the state. For example, if the radiative life-time of the state is too short we will have some photon flux but no reactions with this state. Collisional quenching efficiency depends on the products of the reaction. Quenching of triplet states of nitrogen molecules by molecular oxygen lead to oxygen dissociation and atomic oxygen production. Another example – quenching of singlet oxygen molecules by hydrogen or hydrocarbons mostly leads to heat release without formation of active radicals.

The most important reactions with electronically-excited molecules from the point of view of plasma assisted combustion are channels which lead to radicals formation. There are four different ways to produce radicals through excitation of electronically-excited states:

1. Excitation of the molecular electronic state and radicals production in chemical chains:
   a. $O_2 + e \rightarrow O_2(a^1\Delta_g)[0.98\ eV] + H \rightarrow O(^3P) + OH$
   b. $O_2 + e \rightarrow O_2(b^1\Sigma_g^+)\ [1.64\ eV] + H_2 \rightarrow OH + OH$
   c. $N_2 + e \rightarrow N_2(A^3\Sigma_u^+)\ [6.2\ eV] + O_2 \rightarrow N_2O + O(^3P)$
2. Excitation of the molecule to repulsive or pre-dissociative term leads to molecule dissociation and formation of two radicals:
   a. $O_2 + e \rightarrow O_2(B^3\Sigma_u^-)[8.4\ eV] \rightarrow O(^3P) + O(^1D) + e$
   b. $O_2 + e \rightarrow O_2(C^3\Delta_u)[6.87\ eV] \rightarrow O(^3P) + O(^3P) + e$
   c. $H_2 + e \rightarrow H_2(a^3\Sigma_g^+)[11.8\ eV] \rightarrow H(^1S) + H(^1S) + e$
3. Excitation of the molecule and dissociative quenching of excited state by another molecule:
   a. $N_2 + e \rightarrow N_2(C^3\Pi_u)[11.02\ eV] + O_2 \rightarrow N_2 + O(^3P) + O(^1D)$
   b. $N_2 + e \rightarrow N_2(C^3\Pi_u)[11.02\ eV] + H_2 \rightarrow N_2 + H(^1S) + H(^1S)$
   c. $O_2 + e \rightarrow O_2(A^3\Sigma_g^+)[4.5\ eV] + CH_4 \rightarrow O_2 + CH_3 + H(^1S)$
4. Excitation of the molecular electronic state with radiative depopulation, high-energy photon flux generation and dissociation (ionization) of gas molecules by this radiation:
   a. $N_2 + e \rightarrow N_2(B^1\Pi_u)\ [12.5\ eV] \rightarrow N_2 + h\nu \rightarrow O_2 + h\nu \rightarrow O_2^+ + e$
   b. $N_2 + e \rightarrow N_2(B^1\Pi_u)\ [12.5\ eV] \rightarrow N_2 + h\nu \rightarrow CH_4 + h\nu \rightarrow CH_3 + H$
   c. $H_2 + e \rightarrow H_2(a^3\Sigma_g^+)[11.8\ eV] \rightarrow H_2(b^3\Sigma_g) + h\nu \rightarrow O_2 + h\nu \rightarrow O + O$

Comprehensive detailed kinetic models were discussed, for example, in [Kossyi et al, 1992] for N_2-O_2 mixtures, in [Zatsepin et al, 2001] for H_2-O_2-N_2 mixtures and in [Anikin et al, 2006] for C_xH_y-O_2 mixtures. It should be noted however, that channel branching, rate coefficients and even products of such reactions are not very well known. The first group of processes was investigated much better, than second and third. Simultaneous presence in the plasma of all sorts of excited particles and radicals makes detailed kinetic analysis an extremely challenging and resource-consuming task. As an example we just mention that mixture composition variation, very popular approach in combustion chemistry, will not work in plasma chemistry because simultaneously with afterglow kinetics variation we will change electron energy distribution function in the discharge phase and kinetics of gas excitation.
Mechanism (I) requires very low electric field to increase the efficiency of the excitation process because of low energy threshold for oxygen singlet states population. On the contrary, mechanisms (II)-(IV) require high *E/n* value and high electron energy for upper electronic states excitation.

2.3 Low-energy electronic states excitation

Singled oxygen molecules as a tool for ignition and combustion control were proposed by group of Starik [Smirnov et al, 2008]. The effect of the excitation of oxygen molecules to the $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ electronic states in the electrical discharge on the velocity of laminar flame propagation in the H_2–O_2 mixture was analyzed. The calculations showed that the excitation of O_2 molecules to the $a^1\Delta_g$ and $b^1\Sigma_g^+$ electronic states allows one to increase significantly (by a factor of 2.5) the velocity of flame propagation for the fuel lean hydrogen–oxygen mixture. For stoichiometric and fuel rich mixtures the increase in flame velocity due to an abundance of singlet oxygen molecules in the mixture was found to be significantly smaller (about a factor of 1.1). Later the same team proposed to use a laser radiation at λ = 762.346 nm for O_2 molecules excitation to the $b^1\Sigma_g^+$ electronic state. Experimental observation of the shortening of the induction zone length in a premixed mode of combustion in a subsonic H_2–O_2 low pressure flow due to the presence of oxygen molecules excited to the singlet $a^1\Delta_g$ electronic state was reported in [Smirnov et al, 2008]. The low pressure electric glow discharge was used to produce singlet oxygen molecules. The analysis showed that ~1% of $O_2(a^1\Delta_g)$ molecules in the H_2–O_2 mixture allows to noticeably reduce the ignition delay length and to ignite the mixture at a lower temperature. Authors conclude that the results obtained demonstrate the possibility to intensify the combustion of a hydrogen–oxygen mixture by means of excitation of O_2 molecules by electrical discharge at low pressure (P = 10–20 Torr).
A numerical study of the plasma assisted ignition of hydrogen-oxygen mixtures at different *E/n* has been performed in [Wu et al, 2010]. Results at low *E/n* values are compared with experimental data [Smirnov et al, 2008] and good agreement between experimental and numerical data was demonstrated. It was shown that the efficiency of radicals production through the oxygen singlet states excitation is limited by collisional quenching of SDO molecules in oxygen-fuel mixtures; in oxygen-nitrogen mixtures main efficiency limitation comes from discharge energy flux alternation by vibrational excitation of nitrogen [Wu et al, 2010] (see also Figure 17,b). In paper [Wu et al, 2010] two different mechanisms of radical formation were analyzed: 1) at low *E/n* - through oxygen singlet states excitation with subsequent quenching and conversion into radicals in reactions with fuel molecules, and 2) at high *E/n* – through direct dissociation of molecular oxygen by electron impact and

quenching of nitrogen triplet states in collisions with molecular oxygen. It was shown that the first channel is more efficient in pure oxygen, while the second is much more efficient for mixtures containing more than 10% of nitrogen.

2.4 High electronic states excitation

In papers [Kof&Starikovskii, 1996-1, 1996-2] authors proposed to use pulsed nanosecond discharges for plasma assisted ignition and flame stabilization. The idea was to maintain an extremely high electrical field for a short period of time. This approach allows to generate highly-excited nonequilibrium plasma with the energy distribution shifted to the electronic excitation and dissociation. Short pulse duration restricts the plasma conductivity increase and keeps the energy density in the gas on the relatively low level (equivalent gas heating is in the range of 10-100 K). Paper [Starikovskiy et al, 2011] summarizes the requirements to the pulse discharges to maintain the high efficiency of excitation:

1. High-voltage pulse amplitude is limited to set the value of the reduced electric field E/n > 200-300 Td in the discharge gap which provides optimal conditions for dissociation of molecular oxygen by electron impact and quenching of nitrogen excited states (in air and lean fuel-air mixtures).
2. High-voltage rise dU/dt > 300÷1000 kV/(ns×atm) to obtain the field intensity sufficient for homogeneous ionization wave formation. This condition allows to achieve the homogeneous gas excitation in the gap and simplifies the analysis of the kinetic data. It shold be mentioned, however, that for practical applications inhomogeneous excitation may have specific advantages in some cases (for example, reduction of energy consumption).

This type of the discharge was used in [Zatsepin et al, 2001] to investigate low-temperature kinetics in plasma of pulsed nanosecond discharge. Oxidation of molecular hydrogen in stoichiometric hydrogen-air mixture in the Fast Ionization Wave (FIW) was studied at total pressures p = 1-8 Torr, and the detailed kinetics of the process has been numerically investigated. The excitation of the gas in FIW and dynamics of molecular hydrogen concentration were monitored with the use of measurements of absolute H_2 radiation intensity (transition $a^3\Sigma_g^+ \rightarrow b^3\Sigma_u^+$). Comparison of calculation and experimental results allows to make a conclusion that the gas is predominantly excited behind the FIW front in relatively low electric fields E/n ~ 300-600 Td at electron concentration n_e ~ (1-2)×10^{12} cm^{-3} during approximately 10 ns and the excitation can be described with a good accuracy using the two-term approximation of Boltzmann's equation. In the subsequent processes the reactions including electron-excited particles play a dominant role for the time up to 100 ns, ion-molecular reactions – for the time of microsecond range, reactions including radicals mostly contribute for the time interval of several milliseconds. The most critical processes have been separated for each time interval. The principal role of processes with formation of excited components that support the development of the chain mechanism of oxidation has been shown.

Detailed state-to-state kinetic mechanism [Zatsepin et al, 2001] includes 750 chemical and 8700 vibrational exchange processes with participation of 254 particles including electron-excited and charged atoms and molecules, electrons, radicals, non-excited components, and vibrational-excited molecules H_2, O_2, N_2, H_2O and OH-radical. The most important processes in each time interval in plasma afterglow and radicals recombination were identified. Because the overall picture observed in [Zatsepin et al, 2001] is very typical for plasma assisted ignition by pulsed discharges, we will analyze it in more details.

2.5 Kinetics of plasma assisted combustion below self-ignition threshold

The mixture compression in the engine before the ignition leads to temperature increase. For example, in IC engines initial temperature is close to 600 K, in GTEs – 600-700 K, in SCRAMjets 650-800 K. In these cases the initial temperature of the mixture is below or close to self-ignition threshold. That is why this range of parameters attracts in attention of researchers. From the other hand, this temperature interval is poor investigated from the point of view of chemical kinetic mechanisms. The problem is the lack of data for low-temperature mechanisms validation. As an example, methane combustion GRIMech-3.0 model was validated in the range 1250 – 2500 K. C1-C4 Konnov's mechanism was validated down to ~910 K, hydrogen Popov's mechanism [Popov, 2008] – to 880 K. Direct extrapolation of these models down to room temperature conditions or even to intermediate temperature range below self-ignition threshold, of course, is very questionable (see, for example, analysis in [Uddi et al, 2011]). Thus the task of kinetics investigations in low temperature region becomes extremely difficult and complex. We have to take into account kinetics in gas discharge and plasma afterglow and almost unknown mechanisms of chemical chains initiation under low temperature conditions.

Another problem of investigations of kinetics in plasma is gas discharge inhomogeneity. Under low pressure conditions homogeneous gas ionization and excitation can be achieved even with rather slow voltage increase across the discharge gap. Pressure increase requires a sharp decrease of the voltage rise time (relations 1)-2) above suggest to keep the voltage rise rate on the level of ~ 1 MV/ns/atm for room temperature air to achieve homogeneous excitation). For low pressure conditions this leads to critical voltage rise time about 8 ns and correlate with homogeneous picture of plasma formation in the reaction chamber of 5 cm diameter.

Uncontrollable inhomogeneous excitation significantly compromises the kinetic analysis. That is why some authors prefer to use controlled inhomogeneous excitation instead. For example, in papers [Bak et al, 2011; Stancu et al, 2010; Grisch et al, 2009; Wu et al, 2010, 2011] the point-to-point electrodes geometry was used. This geometry generates non-uniform streamer-like discharge but because of its high reproducibility allows to reconstruct the spatial distribution of excitation and kinetics in plasma.

In [Grisch et al, 2009], detailed experimental investigation of a non-equilibrium nanosecond pulsed discharge in premixed CH_4/air mixtures at atmospheric pressure has been carried out. The electron temperature and density properties were measured using laser Thomson scattering (LTS). Temperature measurements were performed using N_2 CARS thermometry to quantify the energy transfer in the gas mixture. Effect of the discharge on the local temperature shows the existence of the ignition of the gas mixture for equivalence ratio between 0.7 and 1.3. The experiments demonstrated significant reductions in ignition delay and increased lean burn capability relative to conventional spark ignition. Fast development of a flame kernel is then observed. OH and CH PLIF experiments were performed to confirm the large OH and CH streamer-induced production over the discharge volume.

Papers [Bak et al, 2011; Stancu et al, 2010] discuss an important question on the channels of molecular oxygen dissociation in pulsed discharges. In [Bak et al, 2011] time-resolved emission measurements for N_2(C-B) and N_2(B-A) transitions were carried out in nanosecond pulsed discharges in air and pure nitrogen. 0-D kinetic simulations coupled with energy equation are conducted to predict quenching rate coefficients of quenching of N_2* by N_2 and dissociative quenching of N_2* by O_2 by matching the simulated emission curves to the

corresponding measurements. The dissociative quenching was found to be responsible for 82 % of O production whereas the electron-impact dissociation was ~5%.
Papers [Pai et al, 2009, Stancu et al, 2010] reports the results of investigations of nanosecond repetitively pulsed discharge in atmospheric pressure discharge in air or nitrogen preheated at 1000 K. The ground state of atomic oxygen was measured by two-photon absorption laser induced fluorescence, the density of $N_2(A)$ was measured by cavity ring down spectroscopy and the densities of $N_2(B)$ and $N_2(C)$ were measured by optical emission spectroscopy. Measurements of O, $N_2(B)$ and $N_2(C)$ densities have confirmed that the formation of atomic oxygen occurs through the fast two-step mechanism through excitation and quenching of nitrogen triplet states [Stancu et al, 2009].
Papers [Wu et al, 2010, 2011] present measurements of time evolution of hydroxyl radicals in premixed hydrocarbon-air flow in the afterglow of a nanosecond pulsed discharge at atmospheric pressure. The temperature ranged from 300 to 800 K. The fuels were methane, ethane, propane and butane, at an equivalence ratio of 0.1. The plasma was generated by 20 kV pulses of 10 ns duration with < 1 ns rise time at repetition rate of 10 Hz. The tip electrode shape ensured a stable streamer discharge. The reactant flow rate was set at ~20 cm/s so that each discharge pulse occurred in a fresh gas mixture. Laser induced fluorescence was used to measure the concentration of OH radicals after the discharge. The energy of the excitation laser was adjusted to insure that the measurements were made under saturation conditions for all experiments. The time evolution of OH radicals was tracked by adjusting the delay time between the high-voltage pulse and the concentration measurement. It was shown that the OH concentration demonstrates three maxima: immediately after discharge, on time scale ~100 μs, and the third ~2-5 ms after the initiation. This behavior demonstrates relatively long chains development under low temperature conditions below self-ignition threshold.
The important conclusion was made in [Uddi et al, 2011; Wu et al, 2010; Wu et al, 2011] that a new, validated mechanism for low temperature hydrocarbon combustion is required for qualitative description of plasma assisted combustion below self-ignition threshold. This problem is still unsolved at require a lot of new efforts.

3. Plasma assisted combustion above self-ignition threshold

Kinetics above self-ignition threshold is relatively good understood for hydrogen and small hydrocarbons. Verified kinetic models exist for all saturated hydrocarbons from methane to n-decane at temperatures T > 1000-1200 K and pressures from several Torr to several atmospheres. Presence of detailed chemical models simplifies the analysis of plasma assisted combustion experiments in this range of parameters. The only difference between auto-ignition and plasma assisted ignition is a high concentration of radicals from the very beginning of the process and potential influence of non-equilibrium mechanisms with participation of vibrationally- and electronically- excited particles and ions.
The challenge of high-temperature experiments is the controllable heating of the mixture in combination with the homogeneous non-equilibrium excitation by gas discharge. The problem was solved in [Kof&Starikovskiy, 1996-1; 1996-2] where the combined excitation of the combustible mixture by shock wave and fast ionization wave was proposed. Experimental installation was based on the shock tube coupled with the discharge section. Discharge was generated by Marks-type high-voltage pulse generator. The generator consisted of 10 steps and operated at U = 80-250 kV. Ferrite line with non-linear response

and an impedance of 40 Ohm allowed to decrease the pulse leading front down to 500 ps. The voltage increase rate on a high voltage electrode was up to 500 kV/ns and allowed a fast ionization wave formation in the discharge section.

Ignition delay time was analyzed for oxygen-hydrogen mixtures and numerical analysis of chemical kinetics was performed for simultaneous mixture excitation by shock wave and high voltage ionization wave. Ionization wave influence (U ~ 250 kV, t_{pulse} ~ 40 ns) on the ignition delay time of the mixture H_2:O_2:N_2 = 5:19:76 at p = 1 atm was investigated. High efficiency of the fast ionization wave for spatially-uniform excitation of the chemically-reacting systems has been found [Kof&Starikovskiy, 1996-1; 1996-2]. The experimental works using this installation show the high efficiency of this methodology for high-temperature plasma assisted combustion investigation (see, for example, [Aleksandrov et al, 2009-1; 2009-2; Kosarev et al, 2009; 2008; 2008-2; Starikovskii et al, 2006; Starikovskii, 2005; Starikovskaia et al, 2004; Bozhenkov et al, 2003; 2002]).

The kinetics of ignition in C_nH_{2n}+2O_2:Ar mixtures for n = 2 to 5 has been studied experimentally and numerically after a high-voltage nanosecond discharge [Kosarev et al, 2008]. The ignition delay time behind a reflected shock wave was measured with and without the discharge. It was shown that the initiation of the discharge with a specific deposited energy of 10–30 mJ/cm^3 leads to an order of magnitude decrease in the ignition delay time. Discharge processes and following chain chemical reactions with energy release were simulated. The generation of atoms, radicals and excited and charged particles was numerically simulated using the measured time-resolved discharge current and electric field in the discharge phase. The calculated densities of the active particles were used as input data to simulate plasma-assisted ignition. The sensitivity of the results to variation in electron cross sections, reaction rates and radical composition was investigated. Good agreement was obtained between the calculated ignition delay times and the experimental data. The analysis of the simulation results showed that the effect of nonequilibrium plasma on the ignition delay is associated with faster development of chain reactions, due to atoms and radicals produced by the electron impact dissociation of molecules in the discharge phase. Finally, we studied the role of various hydrocarbon radicals in the plasma-assisted ignition of the mixtures under consideration.

Figures 21,a-d show the delay times measured and calculated in [Kosarev et al, 2008] in C_2H_6- to C_5H_{12}-containing stoichiometric mixtures with oxygen with 90% Ar dilution as a function of the gas temperature for autoignition and plasma-assisted ignition. The effect of gas discharge leads to a drastic decrease in the ignition delay and to ignition of the mixtures at noticeably lower temperatures and gas number densities.

Good agreement between the measured and calculated ignition delay time after the discharge in most cases studied shows that the developed kinetic model adequately describes PAI under the conditions considered. Simulation of discharge processes was also validated by comparison between calculated and measured temporal evolution in the discharge current and in the specific energy deposited in the discharge phase.

It should be mentioned that kinetic model of active particles formation in the discharge used in [Kosarev et al, 2009; Aleksandrov et al, 2009] is significantly simplified and rate coefficients of number of processes are not very well known. It is related to ions composition and in part to composition of hydrocarbon radicals. Fortunately under conditions of typical lean mixtures combustion the atomic oxygen always plays a major role. Atomic hydrogen and hydrocarbon radicals are less important but processes of their formation are also well investigated and could be modeled with rather high accuracy. Uncertainty in the radical's

relative composition is not of critical importance under such conditions because the ignition delay time and rate of chemical energy release at high temperatures does not significantly depend on the radical's nature [Aleksandrov et al, 2009]. Thus even the plasma-chemical systems are very complex and some processes are not investigated in details at the moment, plasma assisted ignition and combustion at high temperatures are controlled by rather simple and well-understood mechanisms and radicals.

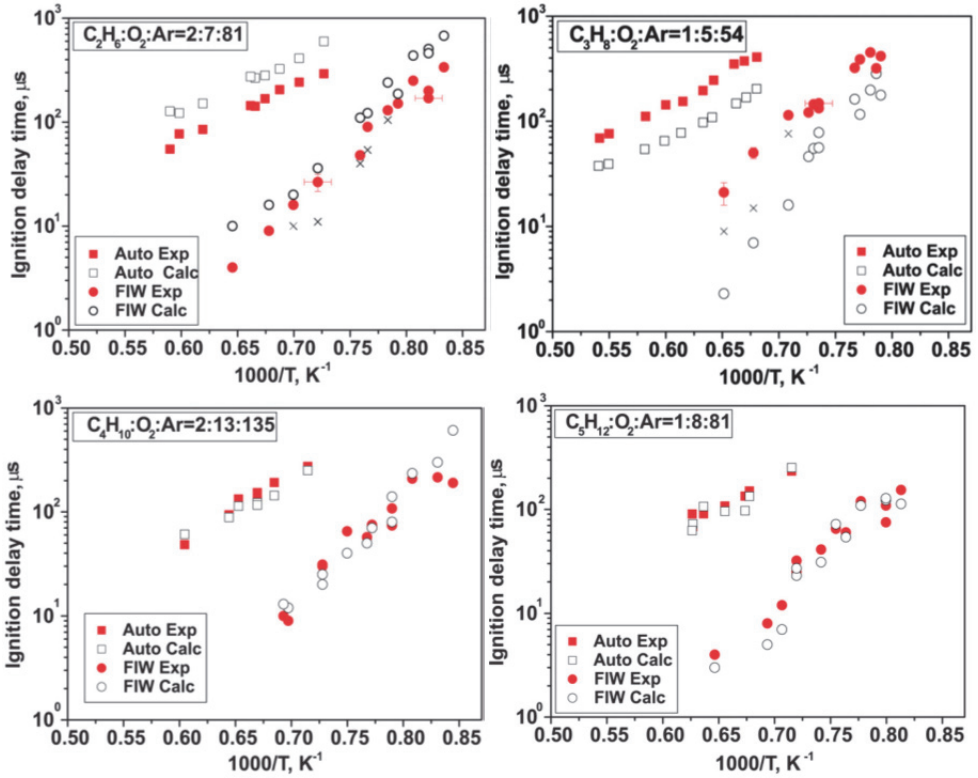


Fig. 21. The delay time for autoignition and ignition with the discharge as a function of temperature. Closed symbols correspond to measurements and open symbols correspond to calculations. a) C_2H_6:O_2:Ar mixture; b) C_3H_8:O_2:Ar mixture; c) C_4H_{10}:O_2:Ar mixture; d) C_5H_{12}:O_2:Ar mixture [Kosarev et al, 2008].

3.1 Vacuum Ultraviolet Emission of the discharge

It is well known that both equilibrium and non-equilibrium plasma are strong sources of vacuum ultraviolet radiation (VUV). Absorption of VUV radiation by oxygen leads to molecular oxygen dissociation with quantum efficiency close to one. Thus ultraviolet sources potentially can generate high concentration of active radicals in the gas.

In [Berezhetskaya et al, 2005] the possibility to utilize the discharge VUV self-emission for ignition stimulation has been considered. The pulsed microwave radiation was generated by

a MI-389 magnetron. The radiation parameters are the following: peak power $P_i \leq 400$ kW, pulse duration $\tau_f \leq 50$ ms, wavelength $\lambda_f \sim 2.5$ cm, repetition frequency f ~ 10 Hz. It was shown that, both in hydrogen-oxygen and in methane—oxygen media, the non-self-sustained discharge initiates the primary combustion wave with relatively low temperature and low glow intensity. In contradiction to the literature data, the experiments did not show a significant difference between the propagation velocities of combustion waves in hydrogen- and methane-containing media, and considerable (above 1000 K) temperature jumps were observed behind the front of the primary wave. After the primary wave started from the initiator passed some distance, a bright burst occurred rapidly and almost simultaneously throughout the region under observation. This burst is characteristic of the transition to the explosive combustion of a gas mixture. The observed specifics of the initiation and propagation of the combustion is attributable to both "gasdynamic" and "kinetic" mechanisms (in particular, to the action of ultraviolet radiation on the gas medium; the source of radiation is numerous sparks arising at metal—dielectric contacts at the target surface irradiated with the microwave beam).
It should be mentioned that excitation of upper states with typical thresholds of 10-15 eV require significant energy. Energy price per radical in this case becomes comparable or even higher than for ionization channel. That is why the VUV radiation cannot be considered as a primary channel of mixture excitation. From the other side radiation can propagate through the gas and de-localize the discharge's excitation.

3.2 Role of gas ionization and plasma recombination in PAC

Ionization is very expensive process from the point of view of energy consumption. Best results can be achieved using e-beam with the energy above 1 keV (~34 eV/ion in air). Gas discharges at extremely high overvoltage operate in the same mode, generating flux of run-away electrons with the energy close to voltage applied [Vasilyak et al, 1994]. Gas discharge in the form of fast ionization wave, developing at E/n~1000 Td, has approximately two times lower efficiency of electron-ion pair's production (~65 eV/ion). Gas discharges with lower overvoltage (E/n~100 Td for streamers, glow discharges) have very low efficiency of gas ionization.
Another important point about gas ionization is fast plasma recombination. This process is discussed in details in [Aleksandrov et al, 2009; 2010]. Even for electron's concentration $n_e \sim 10^{12}$ cm^{-3} recombination time at T ~ 300 K is less than 1 μs because of cluster ions formation, and for T ~ 3000 K is order of magnitude longer (~ 10 μs, molecular ion – electron recombination mechanism). Fast termalization of ionization energy leads to effective gas heating in microsecond time scale [Aleksandrov et al, 2009]. Combination of these two factors – very high energy price of ionization and very high rate of recombination – makes the ionization ineffective from the point of view of plasma assisted combustion. Authors of [Anikin et al, 2001; Starikovskiy, 2003] showed that the efficiency of radicals production in air-fuel mixtures has a maximum at E/n ~200-400 Td. Further increase of electrical field value leads to shift of discharge energy distribution to gas ionization and increases the price of radical's production.
Of course detailed analysis of the efficiency of gas ionization on the ignition process should take into account the gas composition, temperature, pressure and plasma density. For example for high concentration of electrons the main process of plasma recombination is electron-ion dissociative recombination [Aleksandrov et al, 2009]:

$$N_2^+ + e \rightarrow N(^4S) + N(^4S, ^2D)$$

$$O_2^+ + e \rightarrow O(^3P) + O(^3P, ^1D)$$

For intermediate concentration of electrons (less than 10^{12} cm^{-3}) binary recombination becomes slower than charge transfer reactions and reactions of cluster ions formation:

$$N_2^+ + N_2 + M \rightarrow N_4^+ + M$$

$$N_4^+ + O_2 \rightarrow O_2^+ + N_2 + N_2$$

$$O_2^+ + O_2 + M \rightarrow O_4^+ + M$$

Recombination of cluster ions O_4^+ and N_4^+ is order of magnitude faster process than recombination of molecular ions and at low temperatures (~300 K) and electron concentration $n_e \sim 10^{12}$-10^{13} cm^{-3} is the main process:

$$N_4^+ + e \rightarrow N_2 + N_2$$

$$O_4^+ + e \rightarrow O_2 + O_2$$

$$O_4^+ + e \rightarrow O_2 + O + O$$

Products of recombination of cluster ions are not very well known. Recent measurements of ozone formation in pulsed SDBD demonstrate very low oxygen dissociation degree in the discharge at high *E/n* (~ 1000 Td) and high pressure (P = 1 atm). It demonstrates that the channel $O_4^+ + e = O_2 + O_2$ prevails and suppresses the atomic oxygen formation in plasma afterglow, increasing simultaneously the energy release to translational degrees of freedom.
At low plasma density ($n_e \sim 10^{10}$ cm^{-3}) and high oxygen concentration the main channel of electron losses is an attachment [Aleksandrov et al, 2009]:

$$O_2 + e + M \rightarrow O_2^- + M$$

Ion-ion recombination becomes the most important process:

$$O_2^- + O_2^+ + M \rightarrow O + O + O_2 + M$$

$$O_2^- + O_2^+ + M \rightarrow O_2 + O_2 + M$$

However, if oxygen dissociation degree is high, negative ions will be effectively destroyed in collisions with atomic oxygen:

$$O_2^- + O \rightarrow O_3 + e$$

and recombination will take place in electron-ion mode. Another scenario of ions influence on the oxidation and combustion processes is ionic oxidation chains formation [Zatsepin et al, 2003; Kosarev&Starikovskii, 2000].
Ionic mechanism, connected with charge transfer in H_2-O_2 system was proposed in [Kosarev&Starikovskii, 2000]:

$$O_2 + e^- + M \rightarrow O_2^- + M$$

$$H_2 + O_2^- \rightarrow OH^- + OH$$

$$OH + H_2 \rightarrow H_2O + H$$

$$OH^- + H \rightarrow H_2O + e^-$$

$$H + O_2 + M \rightarrow HO_2 + M$$

$$OH^- + HO_2 \rightarrow H_2O + O_2 + e^-$$

The same mechanism can be constructed for hydrocarbons. The efficiency of such mechanisms is limited by plasma decay (electron-ion and ion-ion recombination), but the mechanism remains very effective at low temperatures because of very low activation energy of ion-molecular reactions.

In [Jiao et al, 2007] absolute cross sections for electron impact ionization of n-butane (n-C_4H_{10}) as functions of electron energy from 10 to 200 eV have been measured by Fourier transform mass spectrometry. The major ions including the parent ion and eight fragment ions $C_2H_{3-5}^+$, $C_3H_{3,5-7}^+$ and $C_4H_9^+$ are observed, with the total cross section reaching a maximum of ~1.2×10^{-15} cm^2 at ~ 80 eV. It was clearly demonstrated that each ion produces additional radical in the process of charge transfer reactions. The most important process is H^- transfer. This process leads to effective generation of additional radicals and unsaturated hydrocarbons.

Thus gas ionization plays two ways during plasma decay. Recombination can lead to formation of molecules and significant heat release to translational degrees of freedom. Competing mechanism is recombination with radicals (atoms) formation or excited particles formation. This mechanism produces less heat but more active radicals in the discharge afterglow. Overall, ionization produces more thermal heat and less radicals than excitation of electronic degrees of freedom of nitrogen and direct dissociation by electron impact [Aleksandrov et al, 2010] and not very effective from the point of view of active radicals formation because of relatively high energy price per radical.

3.3 Fast gas heating by discharge

Energy release during plasma decay increases the gas temperature and helps to initiate chemical reaction. The key issue is the rate of energy release. Long relaxation time leads to energy "freezing" in the chemical or internal degrees of freedom.

The model [Aleksandrov et al, 2009] of fast gas heating takes into account the mechanisms of energy release suggested in [Popov, 2001] to describe observations at moderate values of E/n. In addition, it considers the channels associated with the excitation of higher excited states of the molecules and with the formation, transformation and recombination of charged particles, which are the processes that become important at high values of E/n. The experiments and calculations show that the fraction of the discharge power spent on "fast" gas heating increased from 10% at E/n = 100 Td to 30–55% at E/n = 1000 Td. This noticeably depended on gas pressure and only slightly on the electron density at $\tau = \tau_d$, n_{ef} . The effect of pressure was negligible at E/n = 100 Td and became more profound at high E/n, at which most of the deposited energy was spent on ionization [Starikovskiy&Aleksandrov, 2011].

4. Summary

Nonequilibrium plasma demonstrates great abilities to control ultra-lean, ultra-fast, low-temperature flames and becomes an extremely promising technology for a wide range of applications, including aviation GTEs, piston engines, RAMjets, SCRAMjets and detonation

initiation for pulsed detonation engines. To use nonequilibrium plasma for ignition and combustion, it is necessary to understand the mechanisms of plasma assisted ignition and combustion and to simulate numerically discharge and combustion processes under various conditions.

The analysis of discharge processes shows that the discharge energy can be deposited into desired internal degrees of freedom of molecules when varying the reduced electric field, E/n, at which the discharge is maintained. The amount of the deposited energy is controlled by other discharge and gas parameters including electric pulse duration, discharge current, gas number density, gas temperature, etc. As a rule, the dominant mechanism of the effect of nonequilibrium plasma on ignition and combustion is associated with the generation of active particles in the discharge plasma. Numerical simulation of discharge processes is based on the solution of the Boltzmann equation for electrons and of the balance equations for active particles. Here, input data are electron-molecule cross sections and rate constants for reactions with excited and charged particles. These data are available for simple molecules such as N_2, O_2, H_2, and, to a smaller extent, for simple hydrocarbons. However, little is known about cross sections and rates for complex hydrocarbon molecules. The lack of this information does not seem critical when considering lean and stechiometric mixtures; but this problem is serious for the simulation of ignition of rich mixtures.

For plasma assisted ignition and combustion in air-containing mixtures, the most promising active species are O atoms and, to a smaller extent, some other neutral atoms and radicals. These active particles are efficiently produced in high-voltage nanosecond pulse discharges due to electron impact dissociation of molecules and due to electron impact excitation of N_2 electronic states followed by collisional quenching of these states to dissociate molecules. This mechanism was validated above self-ignition threshold for lean and stoichiometric fuel-oxygen and fuel-air mixtures at pressures up to 2 atm. It was shown that in a wide range of conditions the optimal *E/n* from the point of view of atomic oxygen generation is approximately two times higher than the electrical breakdown threshold of the mixture. Excitation of low-energy singlet states of O_2 could be efficient in pure oxygen. However, fast quenching of oxygen singlet states by fuel molecules (hydrogen, hydrocarbons) significantly decreases the efficiency of this channel. An addition of N_2 leads to a smaller amount of energy spent on the excitation of these states because of the competition with vibrational excitation of N_2. Vibrationally-excited particles accelerate the reactions and rate of energy release. This mechanism becomes very important at low temperature conditions when reaction energy thresholds cannot be overcome using translational energy only. Competing process is VT relaxation of the gas, which reduces the chances of long energy-back-coupling chains development. Generation of charged particles in discharge plasmas seems to be inefficient for favoring ignition because the energy cost for ionization is too high and the lifetime of charged particles is too short. Despite of this the ionization and fast plasma recombination remains an important channel of fast volumetric gas heating and can be used for plasma assisted ignition.

The major problems of physical and chemical models of plasma assisted combustion are associated with the low-temperature regimes below self-ignition threshold. For this range of parameters there are no validated combustion mechanisms. Uncertainties connected with attempts to extrapolate high-temperature combustion mechanisms to low temperatures are too large. A new, validated mechanism for low temperature hydrocarbon combustion is required for qualitative description of plasma assisted combustion below self-ignition threshold. This problem is still unsolved and requires a lot of efforts.

5. Acknowledgements

The work was partially supported by Russian Foundation for Basic Research under the project "Nonequilibrium plasma thermalization", AFOSR under the project "Fundamental Mechanisms, Predictive Modeling, and Novel Aerospace Applications of Plasma Assisted Combustion", and DOE Combustion Energy Frontiers Research Center.

6. References

Adamovich I.V., Macheret S.O., Rich J.W., Treanor C.E., Fridman A.A. In: Molecular Physics and Hypersonic Flows. Ed by M. Capitelli. Kluwer Acad. Publ. 1996. P.85.

Aleksandrov N.L., Kindusheva S.V., Nudnova M.M. and Starikovskiy A.Yu. Mechanism of ultra-fast heating in a nonequilibrium weakly-ionized air discharge plasma in high electric fields. J. Phys. D: Appl. Phys. 43 (2010) 255201 (19pp)

Aleksandrov N.L., Kindysheva S.V., Kosarev I.N., Starikovskaia S.M., Starikovskii A.Yu. Mechanism of ignition by non-equilibrium plasma. Proceedings of the Combustion Institute 32 (2009) 205–212.

Aleksandrov N.L., Kindysheva S.V., Kukaev E.N., Starikovskaya S.M., Starikovskii A.Yu. Simulation of the Ignition of a Methane−Air Mixture by a High_Voltage Nanosecond Discharge. Plasma Physics Reports, 2009, Vol. 35, No. 10, pp. 867–882.

Andreadis D. Scramjet engines enabling the seamless integration of air & space operations. Pratt & Whitney Space Propulsion, Hypersonics, West Palm Beach, FL, 2005, http://www.pwrengineering.com/dataresources/TIP_Scramjet_Engines_final.pdf

Anikin N B, Starikovskaia S M and Starikovski A Yu. Oxidation of saturated hydrocarbons under the effect of nanosecond pulsed space discharge. J. Phys. D: Appl. Phys. 39 (2006) 3244–3252

Anikin N.B., Pancheshnyi S.V., Starikovskaia S.M., Starikovskii A.Yu. Air Plasma Production by High-Voltage Nanosecond Gas Discharge. 4 Weakly Ionized Gases Workshop. Anaheim. CA, 2001. AIAA 2001-3088.

Babich L.P. Analysis of a new electron runaway mechanism and record-high runaway electron currents supposedly achieved in dense gas discharges. Sov. Phys. Usp., vol. 48 (10), 2005.

Bak M.S., Kim W., Cappelli M.A. Quenching of Excited Electronic States of Molecular Nitrogen in Nanosecond Pulsed Discharges in Atmospheric Pressure Air. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida AIAA 2011-1150

Berezhetskaya N.K., Gritsinin S.I., Kop'ev V.A., Kossyi I.A., Popov N.A., Silakov V.P., and Van Wie D. Microwave Discharge As A Method For Igniting Combustion in Gas Mixtures. 43rd AIAA Aerospace Sciences Meeting and Exhibit 10 - 13 January 2005, Reno, Nevada. AIAA 2005-991.

Bezgin L.V., Kopchenov V.I., Starik A.M., and Titova N.S. Numerical Study of Formation of a Detonation Wave in a Supersonic Flow over a Wedge by an H_2–O_2 Mixture with Nonequilibrium Excitation of Molecular Vibrations of Reagents. Combustion, Explosion, and Shock Waves, Vol. 42, No. 1, pp. 68–75, 2006

Bowcutt K., "Hypersonic vehicles," In: AccessScience, McGraw-Hill Companies, 2009, http://www.accessscience.com

Bozhenkov S.A., Starikovskaia S.M., Starikovskii A.Yu. Low Temperature Nonequilibrium Plasma Production for Ignition and Combustion Control in Hypersonic Flow. 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibition, AIAA 2002-2065.

Bozhenkov S.A., Starikovskaia S.M., Starikovskii A.Yu. Nanosecond Gas Discharge Ignition of H_2 and CH_4 Containing Mixtures. Combustion and Flame, 2003. V 133, pp. 133-146.

Capitelli M. Molecular Physics and Hypersonic Flows. Kluwer Acad. Publ. 1996. P.85.

Do H., Im S., Cappelli M.A., Mungal M.G. Plasma assisted flame ignition of supersonic flows over a flat wall. Combustion and Flame 157 (2010) 2298–2305.

Dutton J. J. Phys. Chem. Ref. Data 4 577 (1975)

Efanov V.M. High-Voltage and High-PRF FID Pulse Generators. FID GmbH. 2011.

Esakov I.I., Grachev L.P., Khodataev K.V., Vinogradov V.A., Van Wie D.M. Efficiency of propane-air mixture combustion assisted by deeply undercritical MW discharge in cold high-speed airflow, AIAA-2006-1212, 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, USA, 9-12 January 2006.

Fridman A.A. Plasma Chemistry. Cambridge University Press. 2008.

Fry R.S. A Century of Ramjet Propulsion Technology Evolution. Journal of propulsion and power. Vol. 20, No. 1, January–February 2004.

Grisch F., Grandin G.-A., Messina D., Attal-Trétout B., Laser-based measurements of gas-phase chemistry in non-equilibrium pulsed nanosecond discharges Comptes Rendus Mecanique 337 (2009) 504–516

Hayashi M., in: Pitchford L.C., McCoy B.V., Chutjian A., Trajmar S. (Eds.), Swarm Studies and Inelastic Electron–Molecule Collisions, Springer-Verlag, New York, 1987, pp. 167–187.

Huxley L.G.H. and Crompton R.W. The Diffusion and Drift of Electrons in Gases (New York: Wiley-Interscience) 1974.

Ionin A.A., Kochetov I.V., Napartovich A.P., Yuryshev N.N., J. Phys. D Appl. Phys. 40 (2007) R25–R61.

Jiao C Q, DeJoseph Jr C A and Garscadden A. Electron impact ionization and ion reactions in n-butane. J. Phys. D: Appl. Phys. 40 (2007) 409–414 doi:10.1088/0022-3727/40/2/018

Kof L.M., Starikovskii A.Yu. Gas-phase ignition initiation by high-voltage ionization wave. XI Russian Combustion Symposium. Chernogolovka. 1996.

Kof L.M., Starikovskii A.Yu. Oxygen-Hydrogen Mixtures Ignition under the High-Voltage Ionization Wave Conditions at High Temperatures. 26th International Symposium on Combustion, 1996. WIP Poster Session: Ignition & Extinction. R.05-052.

Kosarev I., Aleksandrov N., Kindysheva S., Starikovskaia S., Starikovskii A. Kinetics of ignition of saturated hydrocarbons by nonequilibrium plasma: C_2H_6- to C_5H_{12}-containing mixtures, Combustion and Flame, 156, 221–233 (2009).

Kosarev I.N., Aleksandrov N.L., Kindysheva S.V., Starikovskaia S.M. and Starikovskii A.Yu. Kinetic mechanism of plasma-assisted ignition of hydrocarbons. J. Phys. D: Appl. Phys. 41 No 3 (7 February 2008) 032002.

Kosarev I.N., Sagulenko P.N., Khorunzhenko V.I., Starikovskaia S.M. Development of DBD surface nanosecond discharge. Proc. of Int. Symp. on Plasma Chemistry (ISPC), 2009.

Kosarev I.N., Starikovskii A.Yu. Mechanism for Electric Breakdown in a Chemically Nonequilibrium System and the Influence of the Chain Oxidation Reaction in an H_2–Air Mixture on the Breakdown Threshold. Plasma Physics Reports, 2000. V.26. N.8. P.701.

Kosarev I.N., Aleksandrov N.L., Kindysheva S.V., Starikovskaia S.M., Starikovskii A.Yu. Kinetics of ignition of saturated hydrocarbons by nonequilibrium plasma: CH_4-containing mixtures, Combustion and Flame, 154, 569-586 (2008).

Kossyi I.A., Kostinsky A.Yu., Matveyev A.A. and Silakov V.P. Kinetic scheme of the non-equilibrium discharge in nitrogen-oxygen mixtures. PSST. 1 207 1992

Kovach E.A., Losev S.A., Sergievskaia A.L., Khrapak N.A. Models for Physical-Chemical Processes. 3. Thermally-equilibrium and Nonequilibrium Chemical Reactions. Physical-Chemical Kinetics in Gasdynamics. MSU. 2010.

Leonov S.B., Firsov A. A., Yarantsev D. A., Bolshov M. A., Kuritsyn Yu. A., Liger V. V., Mironenko V. R. Dynamics of H_2O Temperature and Concentration in Zone of Plasma-Assisted High-Speed Combustion. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida AIAA 2011-972.

Light G.C., Matsumoto J.H. // Chem.Phys.Letters. 1978. V.58. N.4. P 578-581.

Losev S.A., Sergievskaia A.L., Rysanov V.D., Fridman A.A., Macheret S.O. Reports of Russian Academy of Science. 1996. V.346. N 2. P.192.

Macheret S.O., Fridman A.A., Adamovich I.V., Rich J.W., Treanor C.E.// AIAA-Paper 1994-1984. 1994.

Michael J.B., Dogariu A., Shneider M.N., and Miles R.B., "Subcritical microwave coupling to femtosecond and picosecond laser ionization for localized, multipoint ignition of methane/air mixtures" Journal of Applied Physics. Volume: 108. Issue: 9. Article Number: 093308. NOV 1 2010.

Nikipelov A.A., Rakitin A.E., Popov I.B., Correale G., Starikovskii A.Yu. Plasmatrons Powered By Pulsed High–Voltage Nanosecond Discharge For Ultra–Lean Flames Stabilization. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida. AIAA 2011-1214.

Pai, D.Z., Stancu, G.D., Lacoste, D.A., and Laux, C.O. Plasma Sources Sci. Technol., 18(4) 045030, 2009

Pancheshnyi S. , Nudnova M., Starikovskii A., Development of a cathode-directed streamer discharge in air at different pressures: experiment and comparison with direct numerical simulation, Phys. Rev. E, 71, 016407 (1-12) (2005).

Pancheshnyi S.V., Starikovskaia S.M., Starikovskii A.Yu. Dynamics of population of electronic levels of molecular nitrogen and structure of fast ionization wave. Plasma Physics Reports, 1999. V.25 N.4 P.326.

Park C. Two-Temperature Interpretation of Dissociation Rate Data for N2 and O2. AIAA-Paper 1988-0458. 1988.

Pilla G. , Galley D. , Lacoste D., Lacas F., Veynante D. , Laux C.O., Stabilization of a turbulent premixed flame using a nanosecond repetitively pulsed plasma, IEEE Trans. on Plasma Science, 34(6), 2471-2477 (2006).

Polak L.S. Plasma Chemistry. Cambridge International Science Publishing. 1998.

Popov N.A. Effect of a Pulsed High-Current Discharge on Hydrogen–Air Mixtures. Plasma Physics Reports, 2008, vol. 34, No. 5, pp. 376–391.

Popov N.A. The effect of nonequilibrium excitation on the ignition of hydrogen-oxygen mixtures. High Temp. 45 (2007) 261-279

Popov N.A., Investigation of the mechanism for rapid heating of nitrogen and air in gas discharges Plasma Phys.Rep. 27 886 (2001)

Raizer Y.P. Gas Discharge Physics. Springer; Corrected edition (September 19, 1991)

Rakitin A.E., Popov I.B., Starikovskii A.Yu. Detonation Initiation by a Gradient Mechanism in Propane–Oxygen and Propane–Air Mixtures. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida. AIAA 2011-586.

Rakitin A.E., Starikosvkii A.Yu. Mechanisms of Deflagration-to-Detonation Transition under Initiation by High-Voltage Nanosecond Discharges, Combustion and Flame, 155, 343-355 (2008).

Roupassov D., Nudnova M., Nikipelov A., Starikovskii A. Sliding DBD for Airflow Control: Structure and Dynamics. 46th AIAA Aerospace Sciences Meeting and Exhibit 7 - 10 Jan 2008 Grand Sierra Resort Hotel Reno, Nevada. Paper AIAA-2008-1367.

Rusanov V.D., Fridman A.A. Physics of Chemically-Active Plasma. Nayka, Moscow. 1985.

Sagulenko P.N., Khorunzhenko V.I., Kosarev I.N. Ignition by nanosecond surface discharge. 8th Int. Workshop on Magneto-Plasma Aerodynamics, March 31-April 2, 2009, Moscow, Russia

Serbin S., Mostipanenko A., Matveev I., Tropina A. Improvement of the Gas Turbine Plasma Assisted Combustor Characteristics. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida AIAA 2011-61.

Shibkov V M, Aleksandrov A F, Chernikov V A, Ershov A P, Konstantinovskij R S and Zlobin V V 2006 Combined MW-DC discharge in a high speed propane–butane–air stream *44th AIAA Aerospace Sciences Meeting and Exhibit (Reno, Nevada, USA, 2006)* AIAA-2006-1216

Singleton D., Pendleton S.J. and Gundersen M.A. The role of non-thermal transient plasma for enhanced flame ignition in C_2H_4–air. J. Phys. D: Appl. Phys. 44 (2011) 022001

Smirnov V.V., Stelmakh O.M., Fabelinsky V.I., Kozlov D.N., Starik A.M. and Titova N.S. On the influence of electronically excited oxygen molecules on combustion of hydrogen–oxygen mixture. J. Phys. D: Appl. Phys. 41 (2008) 192001 (6pp)

Stancu G.D., Kaddouri F., Lacoste D.A., and Laux C.O. Atmospheric pressure plasma diagnostics by OES, CRDS and TALIF. J. Phys. D: Appl. Phys. 43 (2010) 124002 (10pp)

Starikovskaia S M, Starikovskii A Yu. Plasma assisted ignition and combustion. In: Handbook of Combustion, by Maximilian Lackner (Ed.), Franz Winter (Ed.), Avinash K. Agarwal (Ed.), Wiley-VCH, ISBN: 978-3527324491 (2010).

Starikovskaia S.M., Plasma assisted ignition and combustion, J. Phys. D. Appl. Phys. 39 (2006) 265-299.

Starikovskaia S.M., Kukaev E.N., Kuksin A.Yu., Nudnova M.M., Starikovskii A.Yu. Analysis of the spatial uniformity of the combustion of a gaseous mixture initiated by a nanosecond discharge, Combustion and Flame 139 177-187 (2004).

Starikovskii A.Yu. Plasma Supported Combustion. Proceedings of the Combustion Institute, Volume 30, Issue 2, January 2005, Pages 2405-2417.

Starikovskii A.Yu. State-to-State Chemical Reaction Rate Constants: Vibronic Terms Approximation. Physical and Chemical Kinetics in Gas Dynamics. 2003. V 1. http://www.chemphys.edu.ru/pdf/2003-12-25-001.pdf

Starikovskii A.Yu., Anikin N.B., Kosarev I.N., Mintoussov E.I., Starikovskaia S.M., and Zhukov V.P. Plasma-assisted combustion. Pure Appl. Chem., Vol. 78, No. 6. 2006.

Starikovskii A.Yu., Plasma supported combustion, Proc. Comb. Inst. 30 (2005) 2405-2417.

Starikovskiy A.Yu., Aleksandrov N.L. Nonequilibrium Plasma Aerodynamics. In: "Aeronautics and Astronautics", ISBN 978-953-307-473-3. InTech, 2011

Starikovskiy A.Yu., Rakitin A.E., Pancheshnyi S.V. Periodic Pulse Discharge Self-Focusing and Streamer-to-Spark Transition in Under-Critical Electric Field. 49th AIAA Aerospace Sciences Meeting. Orlando, Florida. Jan 2011. Paper AIAA-2011-1271.

Tachibana K., Phys. Rev. A 34 (1989) 1007–1015.

Uddi M., Guo H., Sun W., Ju Y. Studies of C_2H_6/air and C_3H_8/air Plasma assisted combustion kinetics in a nanosecond discharge. 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida AIAA 2011-970

Wu L., Fridman A.A. and Starikovskiy A.Yu. Kinetics of Plasma Assisted Combustion At Low Reduced Electric Fields. 48th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2010, Orlando, Florida AIAA 2010-1593

Wu L., Lane J., Cernansky N., Miller D., Fridman A. and Starikovskiy A. Time resolved PLIF and CRD diagnostics of OH radicals in the afterglow of plasma discharge in hydrocarbon mixtures. 7th US National Combustion Meeting. March 20-23, 2011.

Wu L., Lane J., Cernansky N.P., Miller D.L., Fridman A.A. and Starikovskiy A.Yu. Plasma Assisted Ignition Below Self-Ignition Threshold In Methane, Ethane, Propane And Butane-Air Mixtures. Proceedings of the Combustion Institute 33 (2010)

Zatsepin D.V., Starikovskaia S.M., Starikovskii A.Yu. Hydrogen oxidation in a stoichiometric hydrogen-air mixtures in the fast ionization wave. Combust. Theory Modeling, 2001. V.5 pp.97-129.

Zatsepin D.V., Starikovskaia S.M., Starikovskii A.Yu. Non-Thermal Decomposition of N_2O in Pulsed High Current Discharge. Plasma Physics Reports, 2003. V 29, N 6, pp. 517-527.

13

O_2/CH_4 Kinetic Mechanisms for Aerospace Applications at Low Pressure and Temperature, Validity Ranges and Comparison

Angelo Minotti
"Sapienza" University of Rome, Italy

1. Introduction

In recent years, the propellant combination Ox/CH_4 has received attraction in Japan, USA and Europe as a propellant combination for attitude control, upper stage, booster engines and microcombustion systems. Moreover, this propellant pair is of interest for exploration missions (Stone et al., 2008; Hulbert et al., 2008; Arione, 2010; Kawashima et al., 2009) and for in-space propulsion systems. The reason of the exploration/in-space interest stays in the fact that all the missions with a reduced requirement of thermal management and propellant losses through evaporation will surely profit from a Ox/CH4 based propulsion system. Microcombustion, for space and terrestrial use, takes profit from the Ox/CH4 propellant combination thanks to its availability, easy to handle, and knowledge. Besides the interest in methane for space-terrestrial applications, this propellant being a renewable bio-fuel has seen rising interest for both economic and ecologic reasons.

Microthrusters were associated with the emergence of micro- and nano-satellite concepts, in which satellites are conceived capable of the same or similar performance of conventional satellites within a much smaller package/weight by using MEMS technology (Micro Electrical Mechanical System). This increasing interest in MEMS devices, in particular those based or including combustion/chemical propulsion, is also forcing new needs and problems to emerge (Janson, 1994; DeGroot & Oleson, 1996; Mueller, 1997; Bruno, 2001). One of these is the heat loss through combustor walls due to the much increased surface/volume ratio reducing the actual energy available for the cycle chosen: this explains the sometimes startlingly low temperatures observed experimentally (Minotti et al., 2009; Bruno, 2001; Cozzi, 2007; Cozzi & Caratti, 2007; Bruno et al., 2003; Cozzi et al., 2007). Even when equivalence ratios (Φ) are close to one, these call for kinetics capable of realistically predicting ignition delays times and combustion efficiency at a reasonable computational cost.

The requirement to predict with sufficient accuracy combustion performance and heat load to the chamber walls has lead, in the last decade, the numerical modelling to rapidly become an essential part of combustion research and development programs, and there has been an accelerating evolution from the use of single-step empirical kinetics, to the use of lumped semiglobal (multistep) models (Wesbrook & Dryer, 1981; Bowman, 1986), and finally to the inclusion of full detailed chemical kinetic mechanisms to better simulate chemistry interactions. In addition, detailed mechanisms have been developed and validated for the

simplest fuel molecules (Westbrook and Dryer, 1981) and are not available for most practical fuels. Finally there are many occasions where the great amount of chemical information produced by a detailed reaction mechanism is not necessary and a simple mechanism will suffice together with the fact that 3D combustors cannot easily include detailed kinetic mechanisms because the computational costs of such a treatment would be much too great.
Several works concerning hydrocarbon kinetics are present in literature (Paczko et al., 1988; Westbrook Dryer, 1981; Kee et al., 1985; Heffington, 1997; Hautman, 1981; Trevino & Mendez, 1992; Dagaut, 1991), and the work of Gardiner (1999) is important to understand the hydrocarbon oxidation chemistry, in particular for what concerns differences between methane and other hydrocarbons. The state of the art for methane reactions is by the Gas Research Institute, periodically releasing new updated versions of its detailed methane-air reaction mechanism (GRI-Mech, http://www.me.berkeley.edu/gri_mech/ or http://www.gri.org).
Said that, this work indicates two ways to "define/build" a reaction mechanism and presents five reaction mechanisms adopted in hydrocarbons simulations: one global, two 2-steps, one multisteps and one detailed reaction mechanism.
All of them are compared with the detailed GRI-Mech3.0 reaction mechanism (GRI-Mech, 1999) by means of the CHEMKin3.7 tool (the Aurora application) to figure out the ignition delay time and final temperature differences, in order to understand the problems, and limits, related to a delicate topic as the reaction mechanism modelling is.
Section 2 provides few important hints to define a reaction mechanism, section 3 shows the five reaction mechanisms which are studied, while section 4 and 5 report comparisons and their validity ranges.

2. Reaction mechanism definition

A reaction mechanism may be obtained following, in general, two different paths, depending on whether a reduced mechanism or a semiglobal mechanism is required.
If a reduced mechanism is the goal, the "recipe" might be summarized by:

1. definition of the starting detailed mechanism;
2. definition of the operating conditions;
3. sensitivity analysis to reduce the reactions number.

(The sensitivity $\left(Y'_{X_i}\right)$ analysis is the study of how the variation (uncertainty) in the output (Y) of a mathematical model can be apportioned, qualitatively or quantitatively, to different sources (X_i) of variation in the input of a model, that is $Y'_{X_i} = \frac{\partial Y}{\partial X_i}$; this measure tells how sensitive the output is to a perturbation of the input. If a measure independent from the units used for Y and X_i is needed, $S^r_{X_i} = \left(\frac{\bar{X}_i}{\bar{Y}}\right)\left(\frac{\partial Y}{\partial X_i}\right)$ can be used, where $\bar{X}_i$ is the nominal (or central, if a range is known) value of factor X_i and $\bar{Y}$ is the value taken by Y when all input factors are at their nominal value. In the reaction mechanisms the sensitivity analysis is carried out analysing the sensitivity of some species or of some reaction velocities on the overall mechanism).
On the other hand, if a semiglobal mechanism is the goal, the "recipe" might be summarized in the following way:

1. definition of species of interest (they affect the reaction enthalpy and then the final temperature);

2. definition of reactions (and their number);
3. definition of operating conditions;
4. modification of Arrhenius variables (A, n and E_a) to obtain the required ignition delay.

Any simplified reaction mechanism must be capable of reproducing experimental flame properties over the range of operating conditions under consideration. Hence, in both the paths the operating conditions definition plays a fundamental rule; they must be previously decided because the chemistry model, as every model, has a narrow range of validity and fits real data in a narrow range. It is not uncommon that models which fit data just in some points are adopted, by means of extrapolation laws, to figure out chemistry behaviours in ranges wider than their original validity without highlighting the errors percentage differences in these new ranges. Unfortunately this operation leads to big mistakes which are often neglected.

Experience shows, and this will be clear in the following sections, that most or almost all reduced mechanisms are tuned to predict data at *high* temperatures (where it is easy to obtain accurate data) but often at low temperatures, and low pressure, (i.e. 1000K-2000K and for pressures in the range between 1atm and 5atm, typical of non-adiabatic combustion) they are not accurate or do not predict ignition at all.

In general, for a semiglobal mechanism, the simplest overall reaction representing the oxidation of a conventional hydrocarbon fuel is:

$$Fuel \; + \; n_1 O_2 + \left(3.76 N_2\right) \rightarrow n_2 CO_2 + n_3 H_2 O + \left(n_1 3.76 N_2\right)$$

where n_i are determined by the choice of fuel.

This global reaction is often a convenient way of approximating the effects of the many elementary reactions which actually occur but it overestimates the final temperature and mispredicts the overall reaction rate.

The rate expression of the single reaction is usually expressed by:

$$k_{ov} = A T^n \exp\left(-\frac{E_a}{RT}\right)\left[Fuel\right]^a \left[Oxider\right]^b$$

where:

- A is the frequency factor which depends on how often molecules collide when all concentrations are 1mol/L and on whether the molecules are properly oriented when they collide;
- E_a is the energy that must be overcome for a chemical reaction to occur (kJ/mole);
- n defines the functionality rate law with temperature;
- a and b define the functionality rate law with fuel and oxider mass fractions;

This rate must therefore represent an appropriate average of all of the individual reaction rates involved during the reaction and this is obtained tuning the A, E_a, n, a and b variables.

3. Reaction mechanisms

The reaction mechanisms presented here are:

1. Westbrook and Dryer: 4 species and 1 reaction (Westbrook & Dryer, 1981);
2. Westbrook and Dryer: 5 species and 2 reactions (Westbrook & Dryer, 1981);
3. Minotti: 6 species and 2 reactions (Minotti et al., 2009);
4. Kee: 17 species and 58 reactions (Kee et al., 1985);

5. GRI-Mech 12: 32 species and 177 reactions (Gri-Mech 1.2, 1994; Heffington et al., 1997);

These mechanisms have been compared to the predictions given by the detailed GRI-Mech 3.0 (53 species and 325 reactions (GRI-Mech 3.0, 1999; Dagaut et al., 1991)), assumed as the "*reference model*" , for a wide range of equivalence ratio ($0.3 \leq \Phi \leq 1.9$), and at three different pressures (P=1, 3 and 5 atm).

In the following sections the ignition delay comparison and the final temperature comparison are respectively reported.

4. Comparisons – ignition delay times

The ignition delay time is the elapsed time to obtain a temperature increase, from the injection temperature, of 400K.

The ignition delay time has been compared among the five mechanisms, listed above, adopting reactants in the temperature range 1000K - 2000K and at pressure 1, 3 and 5atm. The equivalence ratio (Φ) range tested was from Φ=0.3 to Φ=1.9 ($\Delta\Phi$=0.2), plus Φ=1.

Table 1a and Table 1b provide the ignition delay times, $t_{id,}$ predicted by the reference detailed GRI-Mech 3.0 mechanism as function of temperature, for P=1atm, and at Φ previously indicated (Tables 12a-12b and 23a-23b report data respectively at P=3atm and P=5atm).

Reactants Temperature, K	**Φ=0.3**	**Φ=0.5**	**Φ=0.7**	**Φ=0.9**	**Φ=1**
1000	0.608	0.772	0.982	1.03	1.04
1100	0.111	0.143	0.131	0.192	0.202
1200	0.0244	0.0314	0.0331	0.0424	0.044
1300	0.00649	0.00815	0.000967	0.011	0.0114
1400	0.00211	0.00247	0.00289	0.00323	0.00336
1500	0.000881	0.000915	0.00103	0.00112	0.00114
1600	0.000422	0.000398	0.000423	0.000439	0.000456
1700	0.000246	0.000197	0.000203	0.000203	0.000211
1800	0.000174	0.000108	0.000106	0.000107	0.000109
1900	0.000148	0.000067	0.0000638	0.0000627	0.0000628
2000	0.000139	0.0000398	0.0000381	0.0000375	0.0000377

Table 1. a Ignition Delay, s, P=1atm

Reactants Temperature, K	**Φ=1.1**	**Φ=1.3**	**Φ=1.5**	**Φ=1.7**	**Φ=1.9**
1000	1.15	1.2	1.22	1.43	1.53
1100	0.216	0.23	0.232	0.233	0.292
1200	0.0473	0.0522	0.0524	0.0523	0.0642
1300	0.0122	0.0134	0.0142	0.0155	0.0163
1400	0.00355	0.0039	0.00425	0.00445	0.00433
1500	0.00121	0.00127	0.00133	0.00143	0.00152
1600	0.00044	0.000492	0.000525	0.000529	0.000533
1700	0.000213	0.000221	0.00023	0.000236	0.000244
1800	0.000111	0.000112	0.000112	0.000116	0.00012
1900	0.000062	0.0000622	0.0000623	0.0000627	0.0000638
2000	0.0000374	0.0000376	0.0000384	0.0000385	0.0000392

Table 1. b Ignition Delay, s, P=1atm

From Tables 1a-1b, Tables 12a-12b and Tables 23a-23b it is possible to define a reactants temperature range where reactions might be completed, that is a range in which the Damkoehler number (residence time/chemical time) is less than 1. For example these tables indicate that ignition delay times vary between 8.43×10^{-6}s and 1.54s. Figure 1 reports the ignition delay (t_{id}), at Φ = 1, as function of reactants temperature and for the different reaction mechanisms.

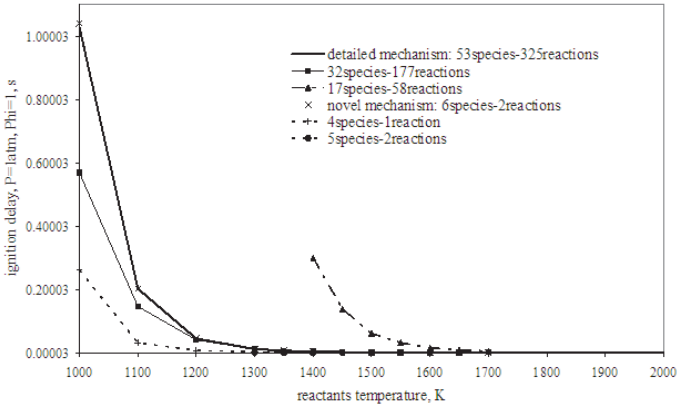


Fig. 1. Φ=1.0: t_{id} comparison

Tables 2 to 11 show the percent differences between the t_{id} predicted by GRI-Mech 3.0 and the reduced mechanisms tested (that is, GRI-Mech3.0 - Reduced Mechanism)/GRI-Mech3.0) at pressure equal to 1atm and for all the equivalence ratios mentioned above. Negative percentages mean that the reduced mechanism overpredicts the reference.
Blank spaces mean that no convergence or no ignition has been obtained at that temperature.
Tables, instead of figures, have been chosen for clarity (in some cases differences are too large).

	Mechanisms				
Reactants Temperature	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	42.93%		-4176.32%		
1100	25.86%		-2152.25%		
1200	10.25%		-1400.00%		
1300	0.62%		-1166.56%		
1400	-3.79%	-13075.36%	-1298.10%		
1500	0.00%	-6347.22%	-1534.51%		
1600	-3.79%	-3383.41%	-1817.06%		
1700	-6.50%	-1822.76%	-1900.00%		
1800	-9.77%	-974.71%	-1721.84%		
1900	-16.22%	-513.51%	-1352.70%		
2000	-34.53%	-308.63%	-1000.72%		

Table 2. Φ=0.3: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	45.34%		-576.17%		
1100	27.97%		-455.24%		
1200	12.74%		-355.41%		
1300	1.35%		-271.78%		
1400	-3.64%		-275.30%		
1500	-5.36%	-6468.31%	-331.69%		
1600	-6.53%	-3744.22%	-407.54%		
1700	-8.63%	-2331.47%	-488.83%		
1800	-8.33%	-1557.41%	-572.22%		
1900	-2.69%	-1050.75%	-614.93%		
2000	-7.79%	-847.24%	-736.68%		

Table 3. Φ=0.5: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	56.31%		-125.05%		
1100	6.87%		-206.11%		
1200	-0.30%		-145.92%		
1300	1.34%		-85.11%	84.59%	91.52%
1400	-4.50%		-72.66%	84.15%	91.28%
1500	-1.94%	-5919.42%	-90.29%	83.98%	91.21%
1600	-7.09%	-3611.58%	-134.75%	84.00%	91.61%
1700	-7.88%	-2303.94%	-181.28%	84.53%	93.89%
1800	-9.43%	-1616.98%	-233.96%		
1900	-7.05%	-1121.00%	-263.64%		
2000	-9.97%	-871.13%	-317.32%		

Table 4. Φ=0.7: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	49.13%		-18.45%		
1100	33.85%		-30.73%		
1200	12.03%		-28.07%		
1300	5.45%		-9.09%	84.82%	91.64%
1400	-5.88%	-9590.40%	-5.88%	84.43%	91.21%
1500	-5.36%	-5542.86%	-14.29%	83.13%	90.98%
1600	-10.48%	-3544.65%	-43.05%	82.44%	90.75%
1700	-12.32%	-2338.42%	-76.35%	82.32%	90.49%
1800	-13.08%	-1619.63%	-106.54%		
1900	-10.05%	-1161.56%	-129.67%		
2000	-12.00%	-908.00%	-162.93%		

Table 5. Φ=0.9: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	44.76%		-0.97%	74.66%	86.21%
1100	27.72%		0.00%	84.60%	90.54%
1200	9.09%		-5.00%	84.05%	91.36%
1300	1.75%		5.26%	84.74%	92.01%
1400	-2.72%		0.00%	84.88%	83.38%
1500	-1.40%	-8828.57%	3.57%	90.54%	84.35%
1600	-0.88%	-7087.50%	1.56%	89.84%	84.38%
1700	-0.53%	-5207.02%	-5.26%	82.98%	84.47%
1800	-0.19%	-4280.28%	-15.35%	85.21%	85.49%
1900	-0.17%	-3255.26%	-28.95%	82.46%	84.96%
2000	-0.28%	-2768.42%	-28.62%	83.55%	86.41%

Table 6. Φ=1.0: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	48.52%		29.30%	80.96%	
1100	29.63%		20.37%	85.28%	
1200	10.99%		15.01%	84.97%	
1300	0.00%		22.21%	84.75%	91.64%
1400	-6.76%	-8857.75%	26.76%	84.51%	91.15%
1500	-8.26%	-5197.52%	22.40%	82.98%	90.66%
1600	-18.18%	-3581.82%	-1.82%	80.41%	89.73%
1700	-12.21%	-2252.11%	-19.72%	81.22%	90.23%
1800	-9.91%	-1575.68%	-41.44%	81.89%	90.63%
1900	-12.68%	-1187.32%	-63.72%	82.66%	5.14%
2000	-13.37%	-926.74%	-86.36%	83.48%	91.28%

Table 7. Φ=1.1: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	50.56%		51.43%	75.24%	90.40%
1100	31.65%		44.73%	82.57%	90.84%
1200	11.69%		40.61%	84.54%	92.07%
1300	1.49%		43.81%	84.55%	91.72%
1400	-5.64%	-8156.41%	45.90%	84.38%	91.95%
1500	-11.81%	-5002.36%	41.65%	83.31%	90.24%
1600	-11.99%	-3213.01%	29.67%	80.67%	89.76%
1700	-13.12%	-2171.49%	11.31%	80.90%	90.41%
1800	-12.50%	-1578.57%	-7.14%	81.43%	90.00%
1900	-14.61%	-1203.37%	-25.52%	80.74%	89.70%
2000	-14.63%	-937.23%	-42.02%	81.73%	90.29%

Table 8. Φ=1.3: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	48.69%		58.28%	73.93%	90.00%
1100	21.98%		56.03%	82.03%	90.56%
1200	4.58%		51.53%	95.42%	92.02%
1300	-2.82%		53.87%	84.37%	91.62%
1400	-5.88%	-7547.06%	58.12%	83.72%	92.54%
1500	-13.53%	-4817.29%	53.61%	80.98%	89.92%
1600	-11.05%	-3023.81%	45.52%	80.19%	89.70%
1700	-13.04%	-2113.04%	30.87%	79.30%	90.35%
1800	-16.96%	-1596.43%	13.21%	80.63%	77.59%
1900	-17.98%	-1219.42%	-1.61%	81.86%	95.92%
2000	-14.84%	-931.25%	-12.50%	89.97%	93.46%

Table 9. Φ=1.5: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	49.30%		71.26%	77.62%	91.40%
1100	16.88%		60.69%	81.86%	90.43%
1200	-64.71%		31.58%	71.73%	84.33%
1300	-1.29%		63.48%	84.45%	91.68%
1400	-8.09%	-7270.79%	64.94%	83.96%	91.01%
1500	-12.59%	-4508.39%	63.08%	80.77%	89.86%
1600	-16.45%	-3019.09%	54.06%	78.64%	88.62%
1700	-14.83%	-2069.49%	42.37%	78.22%	88.98%
1800	-31.90%	-1555.17%	29.48%	79.05%	89.91%
1900	-20.41%	-1225.36%	15.15%	77.19%	87.93%
2000	-17.14%	-944.16%	5.45%	78.70%	88.83%

Table 10. Φ=1.7: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	52.22%		79.28%	78.95%	91.90%
1100	30.82%		75.45%	82.36%	92.40%
1200	11.06%		70.25%	84.27%	91.98%
1300	-3.07%		69.14%	83.99%	91.60%
1400	-18.24%	-7544.34%	68.13%	81.27%	90.12%
1500	-12.50%	-4268.42%	69.01%	80.53%	89.74%
1600	-17.82%	-3014.45%	60.23%	77.11%	88.16%
1700	-15.57%	-2002.46%	52.87%	77.17%	87.70%
1800	-17.50%	-1516.67%	40.92%	76.92%	90.00%
1900	-21.79%	-1216.61%	27.74%	75.71%	87.21%
2000	-18.37%	-938.27%	19.64%	77.32%	88.14%

Table 11. Φ=1.9: t_{id} % differences between reduced and reference mechanisms

Tables 2-11 above (P=1atm) show that:

- at Φ=0.3 and Φ=0.5 the Westbrook and Dryer mechanisms do not predict ignition at any temperature and ignite in a narrow T range at Φ=0.7 and Φ=0.9. This is worthily since many researchers use this mechanism out of context, e.g., to predict properties of *non-premixed* flames.
- when the Westbrook and Dryer mechanisms ignite, the single step behaves better than the 2-step. Percentage differences are almost constant, that is 80% for the mechanism with 1 reaction and 90% for the mechanism with 2 reactions;
- the Kee mechanism (17 species and 58 reactions) starts predicting ignition only at temperatures > 1400K and consistently overpredicts reference values (differences are 13000% at low Φ);
- even though differences are significant in some T range, the Minotti (6 species and 2 reactions) predicts the 'best' t_{id} at every equivalence ratio except Φ=0.7;
- Minotti predicts realistically t_{id} at Φ=1; in particular from 1000K to 1700K the difference is < 5%, sometimes even <1% and better than GRIMech 12 predictions which adopts 32 species and 177 reactions;
- the GRI-Mech 12 (32 species and 177 reactions) usually fits the reference values better than the other mechanisms;

Ignition delay times at P=3atm and P=5atm are also compared and provided.
Here tables 12a-12b and 13a-13b, report the ignition delay times of the GRIMech3.0 detailed mechanism at P=3atm and P=5atm.

Reactants Temperature, K	**Φ=0.3**	**Φ=0.5**	**Φ=0.7**	**Φ=0.9**	**Φ=1**
1000	0.206	0.238	0.273	0.302	0.302
1100	0.0385	0.0474	0.0562	0.0632	0.0641
1200	0.00957	0.0114	0.0133	0.0152	0.0152
1300	0.00266	0.00324	0.00383	0.00433	0.00438
1400	0.000878	0.00104	0.00122	0.00133	0.00133
1500	0.000341	0.000385	0.000433	0.000479	0.000481
1600	0.000156	0.000163	0.000174	0.000186	0.000187
1700	0.0000822	0.0000767	0.0000808	0.0000836	0.0000838
1800	0.0000503	0.0000404	0.0000408	0.0000416	0.0000417
1900	0.0000356	0.0000234	0.0000227	0.0000228	0.0000228
2000	0.0000291	0.0000144	0.0000136	0.0000134	0.0000134

Table 12. a Ignition Delay, s, P=3atm

Reactants Temperature, K	**Φ=1.1**	**Φ=1.3**	**Φ=1.5**	**Φ=1.7**	**Φ=1.9**
1000	0.332	0.225	0.401	0.432	0.462
1100	0.071	0.0641	0.0826	0.0902	0.0962
1200	0.0172	0.0191	0.0202	0.0221	0.0233
1300	0.00426	0.00591	0.00527	0.00622	0.00631
1400	0.00151	0.00191	0.00172	0.00185	0.00201
1500	0.000521	0.00066	0.000601	0.000636	0.000631
1600	0.000201	0.000248	0.000222	0.000235	0.000248
1700	0.0000874	0.000103	0.0000924	0.0000972	0.000102
1800	0.0000427	0.000051	0.0000451	0.0000465	0.0000479
1900	0.0000231	0.0000272	0.0000239	0.0000244	0.0000251
2000	0.0000135	0.0000159	0.0000139	0.0000141	0.0000143

Table 12. b Ignition Delay, s, P=3atm

Reactants Temperature, K	Φ=0.3	Φ=0.5	Φ=0.7	Φ=0.9	Φ=1
1000	0.125	0.134	0.153	0.17	0.172
1100	0.0234	0.0274	0.0322	0.0362	0.0382
1200	0.00577	0.00682	0.00827	0.00923	0.00982
1300	0.0017	0.00204	0.00242	0.00272	0.00282
1400	0.000578	0.000685	0.000738	0.000886	0.000921
1500	0.000228	0.000255	0.000285	0.000316	0.000331
1600	0.000101	0.000106	0.000114	0.000123	0.000131
1700	0.0000519	0.0000504	0.0000529	0.0000555	0.0000571
1800	0.0000303	0.0000262	0.0000264	0.0000272	0.0000275
1900	0.0000202	0.0000147	0.0000145	0.0000146	0.0000147
2000	0.0000154	0.00000893	0.00000868	0.00000843	0.00000861

Table 13. a Ignition Delay, s, P=5atm

Reactants Temperature, K	Φ=1.1	Φ=1.3	Φ=1.5	Φ=1.7	Φ=1.9
1000	0.182	0.202	0.213	0.233	0.248
1100	0.0402	0.044	0.0472	0.0512	0.0543
1200	0.0102	0.0112	0.0122	0.0132	0.0141
1300	0.00301	0.00328	0.00352	0.00381	0.00402
1400	0.000975	0.00102	0.00112	0.00122	0.00128
1500	0.000342	0.000371	0.000391	0.000422	0.00043
1600	0.000132	0.000142	0.000131	0.00016	0.000166
1700	0.0000582	0.0000612	0.0000623	0.0000665	0.000068
1800	0.000028	0.0000289	0.0000298	0.0000308	0.0000318
1900	0.0000149	0.0000151	0.0000154	0.0000159	0.0000163
2000	0.00000845	0.00000872	0.00000877	0.00000902	0.00000916

Table 13. b Ignition Delay, s, P=5atm

Tables A1-A10 and A11-A20 (in the appendix) report, respectively, the percentage differences at P=3atm and at P=5atm.

Tables A1-A20 show that reaction mechanisms worse their accuracy increasing the operating pressure to P=3atm and P=5 atm, in fact:

- the 1-step and 2-step Westbrook and Dryer mechanisms start igniting only at Φ= 1.0 and Φ=1.1, respectively, both for P=3atm and P=5atm (at P=1atm ignition occurs for Φ>0.7).
- the Kee mechanism predicts ignition only at T > 1500K, both at P=3atm and P=5atm, (at P=1atm the minimum temperature is 1400K), and it consistently overpredicts reference values;
- the Minotti 'ignites' at all Φ, both for P=3atm and P=5atm, behaves worse than at 1atm;
- the single-step Westbrook and Dryer mechanism when it ignites, behaves better than the two-step and the Minotti mechanism;
- the GRI-Mech 12 usually fits the reference values better than the other mechanisms.

5. Comparison – equilibrium temperatures

A parallel comparison was carried out for the equilibrium temperatures predictions.

Table 14a-14b report the temperature predicted by the GRI-Mech 3.0, when equilibrium is reached, as a function of **Φ** and reactants temperature, at P=1atm.

Figures 2 to 11 report equilibrium temperature at P= 1, from Φ=0.3 to Φ=1.9 as function of reactants temperature.
Simulations were carried out up to one hundred times the ignition delay.

Reactants Temperature, K	Φ=0.3	Φ=0.5	Φ=0.7	Φ=0.9	Φ=1
1000	1682	2043	2325	2497	2517
1100	1774	2123	2387	2541	2560
1200	1866	2201	2445	2584	2601
1300	1956	2276	2499	2626	2641
1400	2046	2347	2551	2665	2680
1500	2135	2417	2602	2704	2717
1600	2218	2482	2650	2742	2753
1700	2298	2542	2689	2777	2789
1800	2373	2591	2730	2812	2823
1900	2445	2643	2770	2845	2856
2000	2507	2692	2808	2878	2888

Table 14. a Temperature at Steady State, K, P=1atm

Reactants Temperature, K	Φ=1.1	Φ=1.3	Φ=1.5	Φ=1.7	Φ=1.9
1000	2559	2516	2516	2287	2165
1100	2600	2570	2570	2364	2247
1200	2640	2621	2621	2438	2327
1300	2499	2626	2602	2508	2405
1400	2551	2665	2658	2574	2479
1500	2602	2704	2709	2636	2549
1600	2650	2742	2756	2693	2614
1700	2689	2777	2799	2745	2675
1800	2853	2861	2861	2793	2731
1900	2885	2894	2877	2838	2783
2000	2915	2926	2913	2879	2831

Table 14. b Temperature at Steady State, K, P=1atm

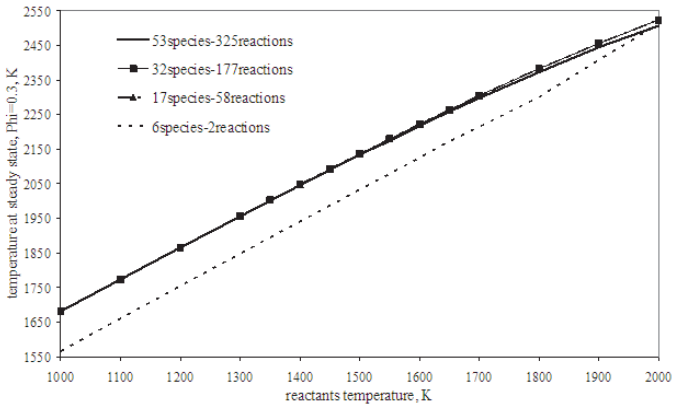


Fig. 2. Φ=0.3, P=1atm, temperature

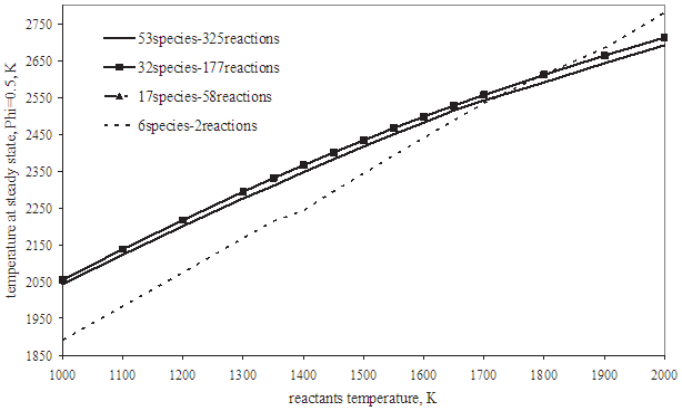


Fig. 3. Φ=0.5, P=1atm, temperature

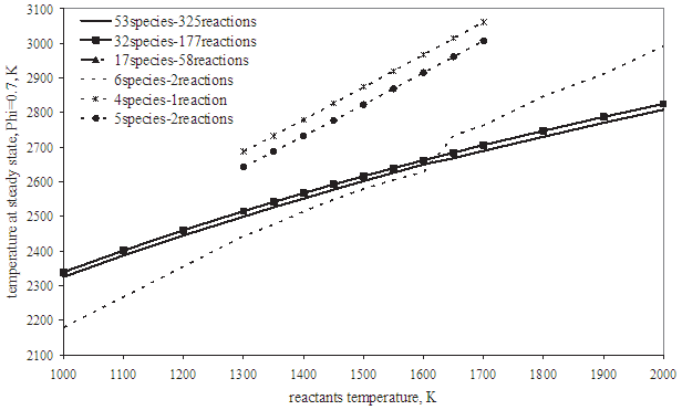


Fig. 4. Φ=0.7, P=1atm, temperature

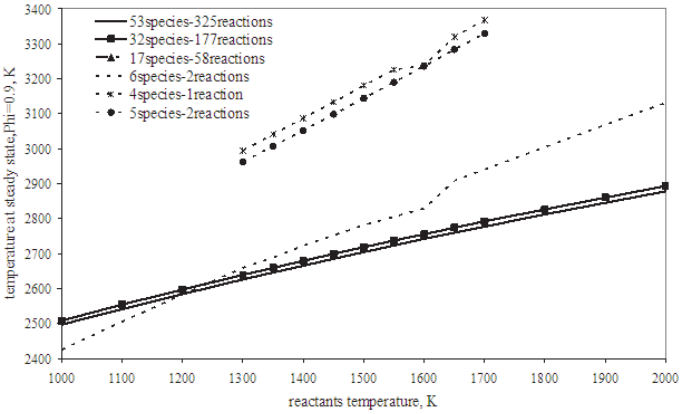


Fig. 5. Φ=0.9, P=1atm, temperature

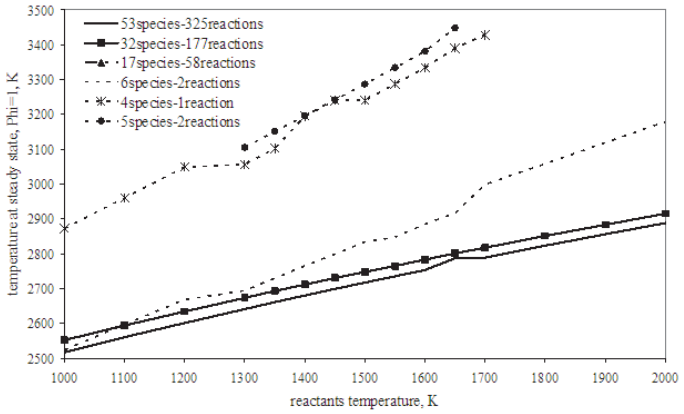


Fig. 6. Φ=1.0, P=1atm, temperature

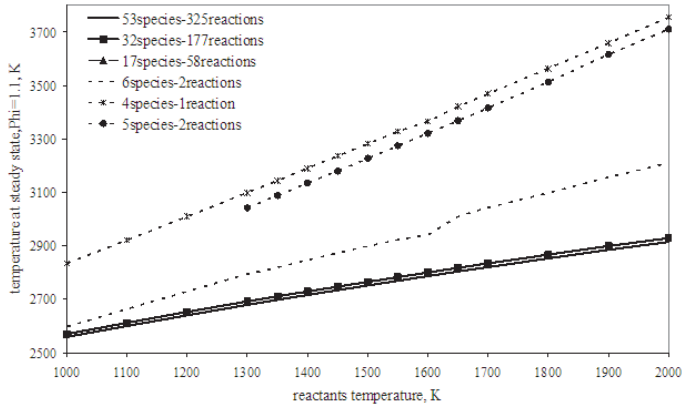


Fig. 7. Φ=1.1, P=1atm, temperature

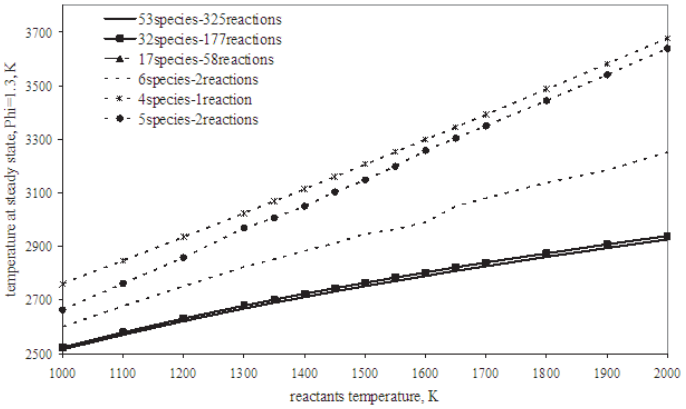


Fig. 8. Φ=1.3, P=1atm, temperature

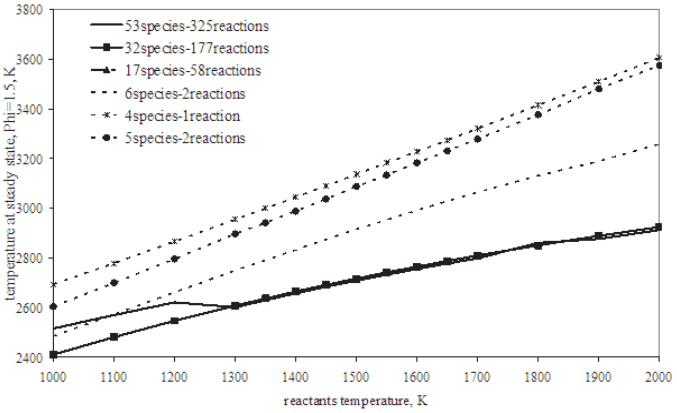


Fig. 9. Φ=1.5, P=1atm, temperature

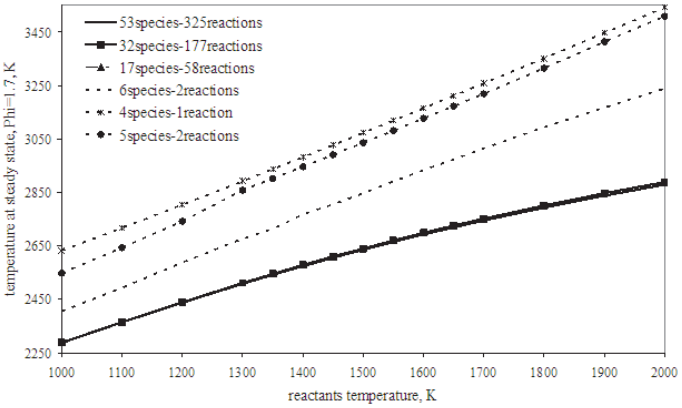


Fig. 10. Φ=1.7, P=1atm, temperature

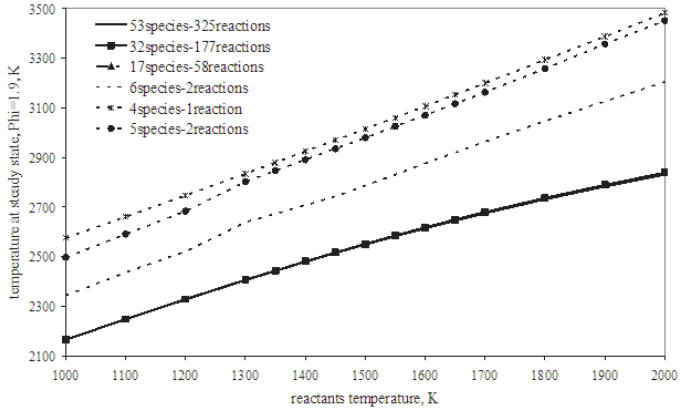


Fig. 11. Φ=1.9, P=1atm, temperature

Above results indicate that:

- Minotti predicted temperatures are always much closer to reference than using the two Westbrook and Dryer mechanisms (when they ignite);
- Minotti mechanism is much more accurate than the other semiglobal W&D mechanisms; accuracy (in percentage difference) is at least of the double, but for some temperature and equivalence ration ranges it reaches grade of accuracy of 4 times better;
- unlike the ignition delay times, mechanisms with large number of reactions and species predict temperatures, *when they ignite*, with always higher accuracy than reduced mechanisms (those predicted with the Kee mechanism, differently from the ignition delay, are close to reference);
- the GRI-Mech 12 usually fits the reference values better than the other mechanisms.

Tables 15a-15b and 16a-16b provide the Temperature at Steady State at pressures equal to 3 and 5atm predicted by the GRIMech3.0 (detailed mechanism). Figures A1-A10 and A11-A20 (in the Appendix) report equilibrium temperature, respectively, at P=3atm and at P=5atm, from Φ=0.3 to Φ=1.9 as function of reactants temperature.

Reactants Temperature, K	**Φ=0.3**	**Φ=0.5**	**Φ=0.7**	**Φ=0.9**	**Φ=1**
1000	1682	2047	2346	2543	2549
1100	1774	2130	2413	2593	2598
1200	1866	2210	2477	2641	2646
1300	1958	2289	2538	2687	2692
1400	2051	2366	2596	2732	2736
1500	2140	2438	2651	2775	2779
1600	2230	2511	2703	2817	2821
1700	2315	2578	2753	2858	2861
1800	2397	2651	2800	2898	2901
1900	2475	2705	2847	2936	2939
2000	2545	2756	2890	2973	2976

Table 15. a Temperature at Steady State, K, P=3atm

Reactants Temperature, K	**Φ=1.1**	**Φ=1.3**	**Φ=1.5**	**Φ=1.7**	**Φ=1.9**
1000	2611	2554	2426	2296	2170
1100	2659	2619	2501	2377	2255
1200	2704	2680	2573	2457	2339
1300	2748	2737	2642	2534	2422
1400	2790	2790	2706	2608	2502
1500	2831	2839	2766	2678	2580
1600	2870	2885	2822	2744	2654
1700	2909	2928	2873	2806	2723
1800	2946	2969	2922	2863	2789
1900	2982	3008	2967	2916	2850
2000	3017	3045	3009	2966	2907

Table 15. b Temperature at Steady State, K, P=3atm

Reactants Temperature, K	Φ=0.3	Φ=0.5	Φ=0.7	Φ=0.9	Φ=1
1000	1682	2048	2353	2562	2618
1100	1774	2131	2423	2615	2666
1200	1867	2214	2490	2665	2713
1300	1960	2294	2554	2714	2758
1400	2052	2372	2615	2761	2802
1500	2144	2447	2673	2807	2845
1600	2234	2521	2728	2851	2887
1700	2321	2590	2780	2894	2928
1800	2404	2655	2831	2936	2968
1900	2483	2722	2878	2977	3007
2000	2561	2778	2924	3016	3045

Table 16. a Temperature at Steady State, K, P=5atm

Reactants Temperature, K	Φ=1.1	Φ=1.3	Φ=1.5	Φ=1.7	Φ=1.9
1000	2634	2560	2431	2299	2171
1100	2684	2626	2509	2380	2257
1200	2732	2690	2584	2463	2343
1300	2778	2749	2656	2543	2427
1400	2823	2804	2725	2620	2511
1500	2866	2855	2789	2649	2591
1600	2908	2904	2849	2764	2668
1700	2949	2950	2905	2830	2742
1800	2988	2993	2957	2892	2812
1900	3026	3034	3006	2949	2877
2000	3064	3074	3051	3003	2938

Table 16. b Temperature at Steady State, K, P=5atm

Figures A1-A20 (in the appendix) show that all reaction mechanisms, but the Minotti one, worse their accuracy increasing the operating pressure to P=3atm and P=5 atm, in fact:

- the 1-step and 2-step Westbrook and Dryer mechanisms do not predict ignition for wide ranges of equivalence ratio both for P=3atm and P=5atm (at P=1atm ignition occurs for Φ>0.7); moreover the 2-step mechanism is slightly more accurate than the 1-step mechanism;
- the Minotti 'ignites' at all Φ, and it holds its overall level of accuracy and it behaves better than at P=1atm for some ranges of temperature and equivalence ratio;
- Minotti mechanism is much more accurate than the other semiglobal W&D mechanisms; accuracy (in percentage difference) is at least of the double, but, for some temperature and equivalence ratio ranges, it reaches grade of accuracy of 100 times better.
- the Kee mechanism predicts ignition only at T > 1500K, both at P=3atm and P=5atm, (at P=1atm the minimum temperature is 1400K), and it predicts values with high level of accuracy;
- the GRI-Mech 12 usually fits the reference values better than the other mechanisms.

6. Conclusions

Current studies on space missions and on micro-combustion, also for Micro-Rockets applications, gave the cue for the present work.
This chapter focuses its attention on some of the most important points concerning with reaction mechanisms and, at same time, five mechanisms are presented, analysed and compared; they deal with hydrocarbon oxidation, in particular methane. Comparison is carried out for a wide range of equivalence ratios ($0.3 \leq \Phi \leq 1.9$), temperatures (1000K-2000K) and pressure ($1 \leq P$ [atm] ≤ 5), adopting as benchmark the detailed GRIMech3.0 reference mechanism, which adopts 55 species and 325 reactions.
This study is important because analyses the mechanisms effectiveness in predicting ignition in good agreement with detailed kinetics calculations with low temperature and low pressures ignition ranges.
These ranges are typical of non-adiabatic combustion and, unfortunately, reduced mechanisms found in literature often fail to predict realistic delay times and equilibrium flame temperatures under these conditions but are usually adopted without a previous validity study. In applications where the flame temperature is lower or much lower than adiabatic, realism and accuracy are indeed critical.
Some general results may be summarised.
Among the semiglobal mechanisms, the Minotti 2-step reduced mechanism is well suited for low temperature flames, that is, in devices where heat losses, e.g., through non-adiabatic walls, are not negligible. Results are in some cases surprising, showing this mechanism predicts, in some ranges, ignition delay times and equilibrium temperatures better than other reduced and even detailed mechanisms, showing also that it 'ignites' at all P and Φ.
Thus it is not always true that including larger numbers of species and reactions predict ignition delays better, in fact there are wide ranges of T and Φ in which the Kee mechanism (17 species and 58 reactions) does not ignite and/or is less accurate than other simplified mechanisms. The Westbrook and Dryer mechanisms ignite only at particular equivalence ratios and at particular reactants temperature.
As for the effect of pressure, at 1atm the Minotti mechanism predicts ignition delays times always more accurately than the other reduced mechanisms, except at $\Phi=0.7$ and for $1500<T$ [K]<1700, where the two Westbrook and Dryer mechanisms behave better. At 3 and 5 atm the behavior is in general the same.
The comparison among predicted equilibrium temperature has shown that mechanisms with high number of species, predict final (stationary) temperatures, when they ignite, better than mechanism with a lower number of species, as it is expected since simpler mechanism cannot include radicals. Comparisons are then meaningful only between mechanisms with similar numbers of species and reactions.
This said, Minotti one is always much more accurate than the two Westbrook and Dryer mechanisms.
To conclude the Minotti mechanism is appropriate for preliminary combustion studies in all the devices that operate at low temperature and pressure, as it combines high accuracy and reliability with ease of implementation and a modest computational effort.

7. Appendix

Ignition delay, P=3atm and P=5atm, comparison

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	29.61%		-2919.42%		
1100	7.01%		-858.44%		
1200	0.00%		-377.53%		
1300	-10.90%		-217.29%		
1400	-11.50%		-172.21%		
1500	-9.97%	-6996.77%	-190.91%		
1600	-8.33%	-3592.31%	-237.82%		
1700	-8.03%	-2004.62%	-284.43%		
1800	-9.34%	-1237.97%	-303.58%		
1900	-12.08%	-686.52%	-287.64%		
2000	-17.18%	-405.15%	-233.68%		

Table A1. P=3atm, Φ=0.3: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	31.51%		-144.54%		
1100	10.97%		-71.31%		
1200	-7.02%		-34.21%		
1300	-12.04%		-1.85%		
1400	-11.54%		18.46%		
1500	-12.47%	-6731.17%	24.16%		
1600	-7.98%	-3703.68%	17.18%		
1700	-9.13%	-2272.88%	1.43%		
1800	-8.91%	-1503.96%	-15.59%		
1900	-8.12%	-1053.85%	-31.20%		
2000	-7.64%	-781.94%	-46.53%		

Table A2. P=3atm, Φ=0.5: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	52.75%		31.14%		
1100	23.67%		37.54%		
1200	-6.02%		45.26%		
1300	-12.53%		50.39%		
1400	-14.75%		59.26%		
1500	-13.86%	-6251.04%	62.36%		
1600	-11.49%	-3624.14%	59.66%		
1700	-6.68%	-2226.73%	53.47%		
1800	-10.78%	-1529.90%	43.87%		
1900	-11.45%	-1120.26%	34.36%		
2000	-11.76%	-848.53%	11.76%		

Table A3. P=3atm, Φ=0.7: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	58.28%		69.83%		
1100	17.25%		68.20%		
1200	0.00%		71.91%		
1300	-11.55%		71.36%		
1400	-15.04%		74.59%		
1500	-15.24%	-5829.02%	76.62%		
1600	-14.52%	-3491.40%	75.32%		
1700	-11.48%	-2220.57%	71.53%		
1800	-12.98%	-1544.23%	65.63%		
1900	-13.16%	-1145.61%	59.34%		
2000	-13.43%	-892.54%	52.84%		

Table A4. P=3atm, Φ=0.9: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	36.42%		90.66%		
1100	15.60%		67.86%		
1200	0.00%		72.17%		
1300	-12.10%		71.92%		
1400	-19.55%		73.91%		
1500	-14.76%	-5804.37%	77.13%		
1600	-13.90%	-3477.54%	75.83%		
1700	-11.34%	-2215.04%	72.08%		
1800	-12.95%	-1542.69%	66.19%		
1900	-13.16%	-1145.61%	60.04%		
2000	-13.43%	-892.54%	53.66%		

Table A5. P=3atm, Φ=1.0: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	36.14%		81.72%	15.66%	
1100	26.34%		81.55%	55.21%	83.10%
1200	0.00%		81.57%	58.66%	81.57%
1300	-22.54%		78.36%	72.07%	78.64%
1400	-13.91%		84.24%	63.58%	81.13%
1500	-17.85%	-5466.22%	88.96%	59.50%	79.08%
1600	-12.94%	-3302.99%	84.38%	57.11%	79.55%
1700	-14.42%	-2165.45%	80.78%	54.12%	76.77%
1800	-14.75%	-1527.63%	76.11%	52.93%	75.88%
1900	-14.29%	-1151.08%	71.56%	53.25%	76.36%
2000	-14.07%	-900.00%	66.89%	54.22%	77.04%

Table A6. P=3atm, Φ=1.1: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	37.85%		88.65%	16.85%	58.29%
1100	11.59%		86.07%	43.03%	70.90%
1200	0.00%		88.59%	57.75%	78.59%
1300	-13.00%		86.25%	60.99%	80.69%
1400	-17.90%		86.42%	62.41%	80.86%
1500	-17.65%	-5176.29%	87.65%	58.65%	79.14%
1600	-16.98%	-3187.74%	87.26%	55.14%	77.36%
1700	-12.94%	-2114.91%	85.75%	53.73%	75.33%
1800	-16.17%	-1517.31%	82.07%	49.43%	74.26%
1900	-16.24%	-1156.41%	78.46%	48.72%	74.02%
2000	-16.06%	-907.30%	75.04%	49.85%	75.33%

Table A7. P=3atm, Φ=1.3: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	43.64%		92.19%	17.46%	59.85%
1100	24.33%		90.24%	49.52%	73.61%
1200	0.00%		89.46%	58.12%	79.41%
1300	-18.22%		88.35%	57.87%	78.94%
1400	-17.44%		91.92%	59.77%	81.69%
1500	-18.80%	-4908.32%	90.25%	64.23%	80.70%
1600	-18.47%	-3093.69%	89.95%	53.15%	77.07%
1700	-20.13%	-2118.61%	88.10%	48.48%	76.84%
1800	-17.74%	-1496.45%	85.83%	46.12%	72.73%
1900	-17.57%	-1151.05%	82.89%	45.19%	72.43%
2000	-17.27%	-907.19%	80.14%	45.97%	72.88%

Table A8. P=3atm, Φ=1.5: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	41.67%		94.93%	18.75%	60.19%
1100	19.84%		97.02%	53.55%	75.61%
1200	4.07%		94.48%	58.69%	81.09%
1300	-9.49%		91.72%	65.11%	80.71%
1400	-19.46%		93.19%	61.41%	82.49%
1500	-19.50%	-4711.32%	91.71%	64.47%	78.30%
1600	-19.57%	-2959.57%	91.70%	51.91%	76.21%
1700	-20.37%	-2039.92%	90.77%	47.12%	75.93%
1800	-18.92%	-1469.89%	88.39%	43.01%	71.40%
1900	-19.26%	-1141.80%	85.94%	41.39%	70.66%
2000	-18.44%	-914.18%	83.48%	41.84%	70.92%

Table A9. P=3atm, Φ=1.7: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	41.13%		95.32%	17.53%	59.31%
1100	19.85%		94.66%	46.47%	72.87%
1200	0.43%		94.81%	56.65%	78.11%
1300	-14.90%		93.34%	58.80%	79.40%
1400	-14.93%	-8009.45%	93.83%	60.05%	79.90%
1500	-28.68%	-4812.84%	95.28%	53.09%	76.55%
1600	-19.76%	-2839.52%	93.35%	50.81%	75.56%
1700	-19.61%	-1958.82%	93.17%	45.39%	73.53%
1800	-20.25%	-1444.89%	90.21%	40.29%	70.15%
1900	-19.92%	-1127.09%	88.09%	38.25%	69.28%
2000	-20.28%	-913.99%	85.87%	37.83%	69.37%

Table A10. P=3atm, Φ=1.9: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	23.60%		-3020.00%		
1100	-0.43%		-652.14%		
1200	-13.69%		-225.82%		
1300	-20.00%		-87.65%		
1400	-17.47%		-39.45%		
1500	-13.16%	-7838.60%	-32.02%		
1600	-9.90%	-3919.80%	-49.50%		
1700	-8.67%	-2115.80%	-71.10%		
1800	-9.24%	-1236.63%	-87.46%		
1900	-10.89%	-746.53%	-89.60%		
2000	-13.64%	-452.60%	-75.32%		

Table A11. P=5atm, Φ=0.3: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	23.13%		-80.60%		
1100	0.73%		-9.85%		
1200	-15.98%		20.53%		
1300	-19.12%		40.20%		
1400	-20.15%		55.47%		
1500	-16.86%	-7664.71%	62.31%		
1600	-16.04%	-4098.11%	61.60%		
1700	-10.71%	-2360.32%	57.14%		
1800	-9.54%	-1518.32%	50.00%		
1900	-10.88%	-1063.27%	41.63%		
2000	-9.07%	-780.18%	34.15%		

Table A12. P=5atm, Φ=0.5: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	26.80%		60.13%		
1100	0.31%		63.35%		
1200	-10.16%		72.91%		
1300	-16.94%		73.97%		
1400	-30.08%		79.95%		
1500	-20.00%	-7198.25%	85.79%		
1600	-17.54%	-4022.81%	80.70%		
1700	-13.42%	-2357.47%	79.21%		
1800	-12.88%	-1574.24%	75.49%		
1900	-12.41%	-1127.59%	71.24%		
2000	-9.79%	-836.64%	67.28%		

Table A13. P=5atm, Φ=0.7: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	29.65%		82.33%		
1100	8.29%		83.12%		
1200	-10.51%		86.89%		
1300	-18.38%		84.60%		
1400	-16.25%		86.00%		
1500	-20.89%	-6703.80%	87.85%		
1600	-17.89%	-3859.35%	87.97%		
1700	-15.68%	-2332.43%	87.19%		
1800	-14.34%	-1576.47%	85.00%		
1900	-14.38%	-1153.42%	82.19%		
2000	-15.90%	-892.88%	79.12%		

Table A14. P=5atm, Φ=0.9: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	29.07%		86.34%		
1100	8.12%		86.52%		
1200	-3.87%		86.66%		
1300	-20.92%		87.16%	37.23%	
1400	-21.61%		88.27%	44.63%	
1500	-21.45%	-6486.10%	89.73%	39.58%	
1600	-16.03%	-3678.63%	90.31%		
1700	-16.29%	-2299.30%	89.47%	33.45%	
1800	-16.00%	-1580.00%	87.60%	30.55%	
1900	-15.65%	-1158.50%	85.24%	30.61%	
2000	-14.63%	-884.90%	82.93%		

Table A15. P=5atm, Φ=1.0: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	32.97%		93.46%	-53.85%	27.47%
1100	9.95%		89.85%	5.22%	52.74%
1200	-9.80%		88.92%	30.29%	68.92%
1300	-16.94%		91.33%	38.21%	69.83%
1400	-21.03%		95.11%	43.59%	73.95%
1500	-23.10%	-6332.75%	90.99%	38.30%	69.88%
1600	-21.97%	-3695.45%	91.21%	34.70%	69.09%
1700	-10.48%	-2288.32%	91.07%	31.10%	65.46%
1800	-16.79%	-1571.43%	89.57%	28.21%	63.93%
1900	4.03%	-1155.03%	87.65%	27.52%	64.23%
2000	-17.87%	-915.38%	85.09%	26.86%	62.72%

Table A16. P=5atm, Φ=1.1: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	34.65%		93.37%	-51.49%	25.25%
1100	11.14%		92.95%	4.55%	51.82%
1200	-8.04%		92.58%	27.95%	63.84%
1300	-18.60%		92.50%	37.80%	69.21%
1400	-28.43%		92.53%	40.29%	69.71%
1500	-14.56%	-5964.69%	93.34%	37.47%	69.27%
1600	8.45%	-3505.63%	93.66%	33.03%	67.46%
1700	-17.65%	-2220.26%	93.28%	31.05%	63.89%
1800	-17.99%	-1553.98%	92.18%	22.49%	61.59%
1900	-17.88%	-1164.90%	90.40%	20.53%	60.40%
2000	-16.97%	-905.73%	89.12%	21.22%	62.04%

Table A17. P=5atm, Φ=1.3: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	33.33%		95.12%	-55.40%	24.41%
1100	10.81%		95.49%	4.24%	53.18%
1200	-6.56%		97.86%	30.66%	65.98%
1300	-19.32%		94.18%	36.93%	68.18%
1400	-25.89%		97.54%	38.21%	71.96%
1500	-25.58%	-5731.20%	94.42%	35.29%	68.29%
1600	-39.69%	-3884.73%	96.45%	20.61%	61.37%
1700	-23.43%	-2227.45%	94.56%	23.92%	61.48%
1800	-19.80%	-1537.58%	93.83%	18.46%	59.40%
1900	-19.48%	-1166.23%	92.47%	14.94%	58.12%
2000	-18.59%	-920.52%	91.61%	14.37%	58.15%

Table A18. P=5atm, Φ=1.5: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	34.76%		96.42%	-50.64%	90.60%
1100	12.30%		96.11%	4.30%	52.93%
1200	0.00%		95.83%	30.83%	68.33%
1300	-17.59%		95.54%	36.75%	68.77%
1400	-23.77%		95.48%	41.48%	74.02%
1500	-24.41%	-5397.63%	95.64%	46.45%	68.25%
1600	-21.25%	-3218.75%	95.80%	29.38%	66.06%
1700	-21.65%	-2110.53%	95.58%	22.71%	60.90%
1800	-21.10%	-1510.39%	95.00%	13.96%	57.47%
1900	-20.13%	-1151.57%	93.91%	10.06%	55.66%
2000	-19.73%	-911.09%	93.28%	9.09%	55.21%

Table A19. P=5atm, Φ=1.7: t_{id} % differences between reduced and reference mechanisms

Reactants Temperature	*Mechanisms*				
	32species 177reactions	17species 58reactions	6species 2reactions	4species 1reaction	5species 2reactions
1000	34.68%		98.90%	-53.63%	26.61%
1100	13.08%		97.77%	5.16%	53.22%
1200	-0.71%		98.01%	28.37%	64.18%
1300	-17.91%		96.99%	35.32%	68.66%
1400	-25.78%	-9821.88%	97.76%	37.27%	69.06%
1500	-30.23%	-5365.12%	96.98%	32.56%	66.28%
1600	-24.70%	-3146.99%	97.99%	26.51%	63.73%
1700	-25.44%	-2105.88%	96.16%	18.09%	58.82%
1800	-22.64%	-1484.91%	95.75%	10.06%	55.66%
1900	-22.09%	-1139.26%	95.11%	4.91%	53.68%
2000	-21.18%	-913.10%	94.26%	2.95%	53.06%

Table A20. P=5atm, Φ=1.9: t_{id} % differences between reduced and reference mechanisms

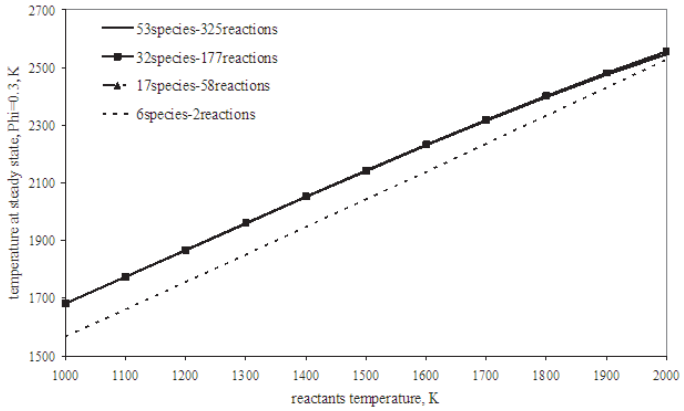


Fig. A1. Φ=0.3, P=3atm, temperature

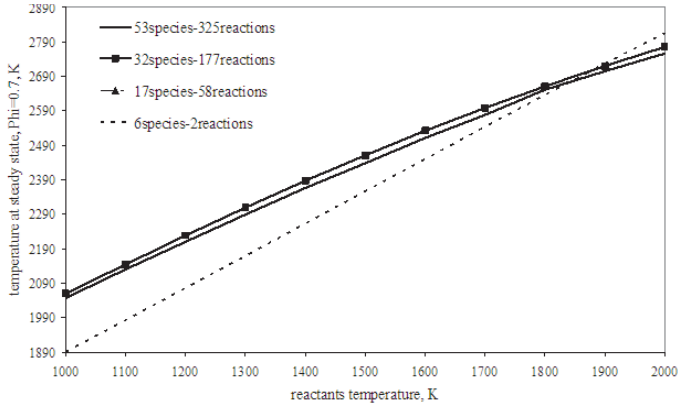


Fig. A2. Φ=0.5, P=3atm, temperature

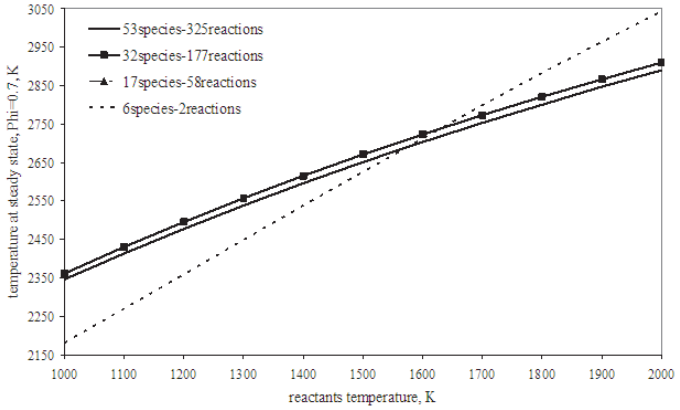


Fig. A3. Φ=0.7, P=3atm, temperature

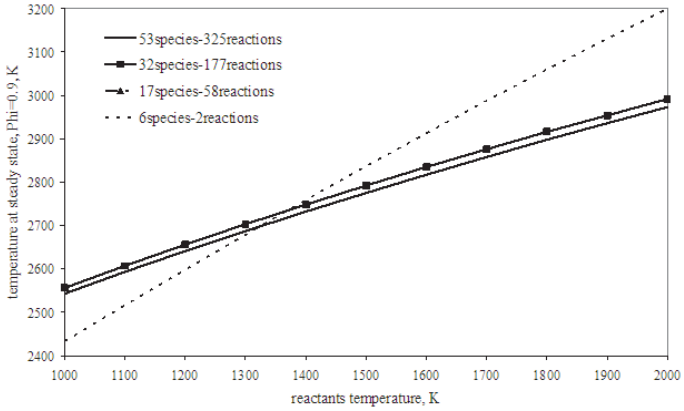


Fig. A4. Φ=0.9, P=3atm, temperature

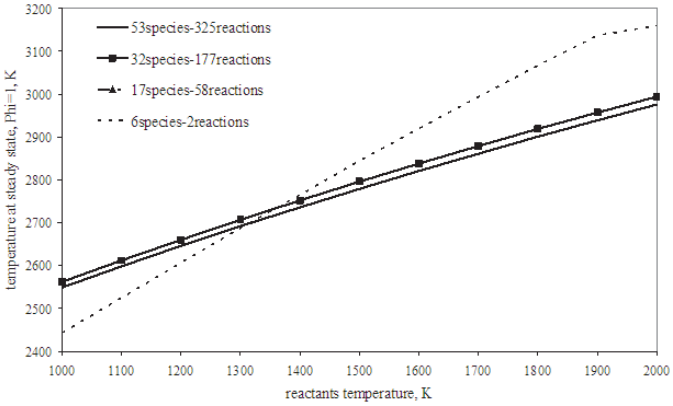


Fig. A5. Φ=1, P=3atm, temperature

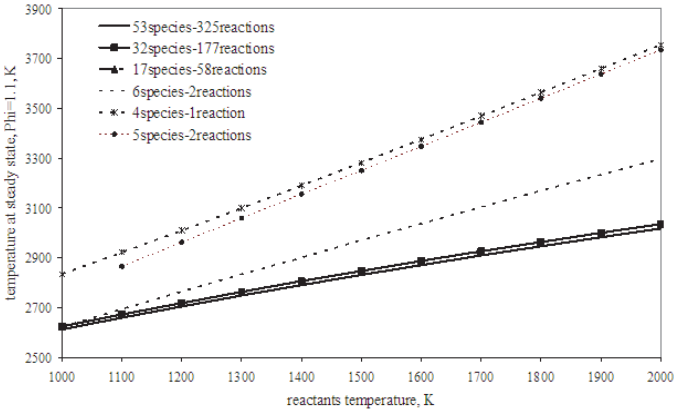


Fig. A6. Φ=1.1, P=3atm, temperature

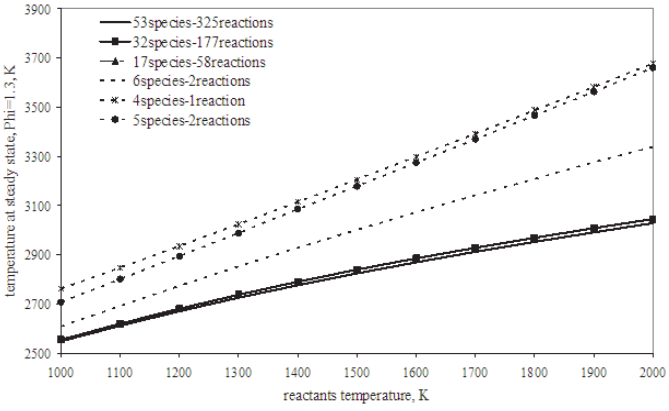


Fig. A7. Φ=1.3, P=3atm, temperature

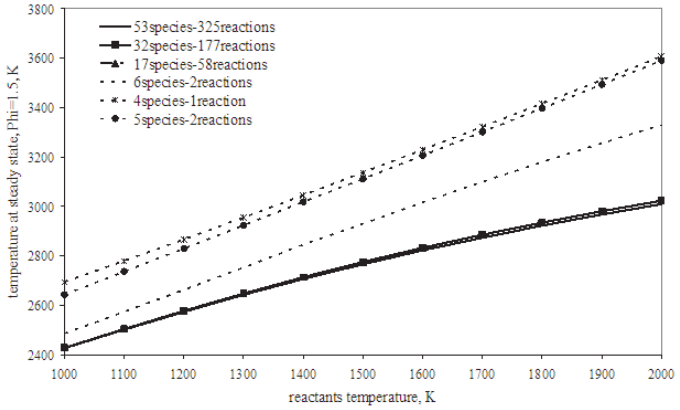


Fig. A8. Φ=1.5, P=3atm, temperature

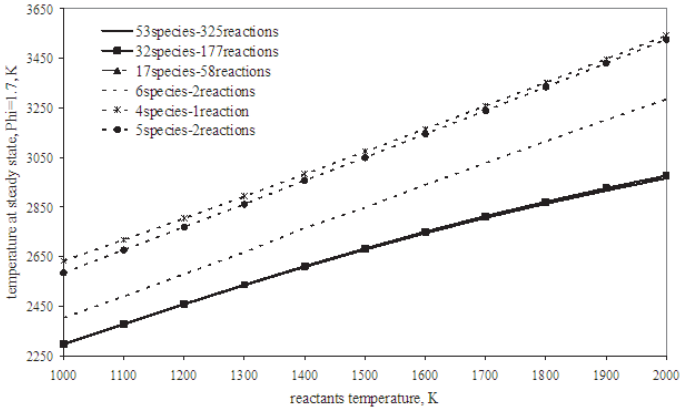


Fig. A9. Φ=1.7, P=3atm, temperature

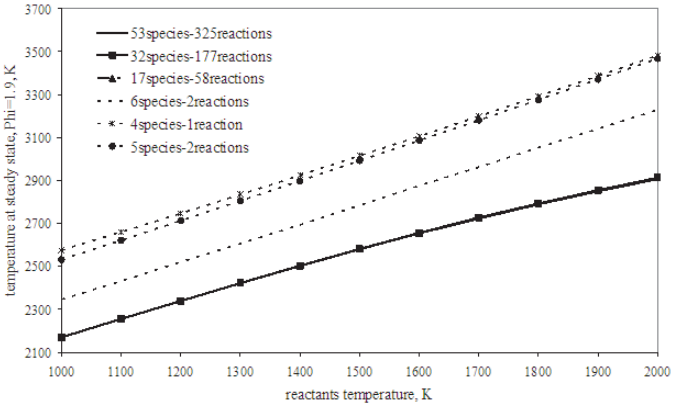


Fig. A10. Φ=1.9, P=3atm, temperature

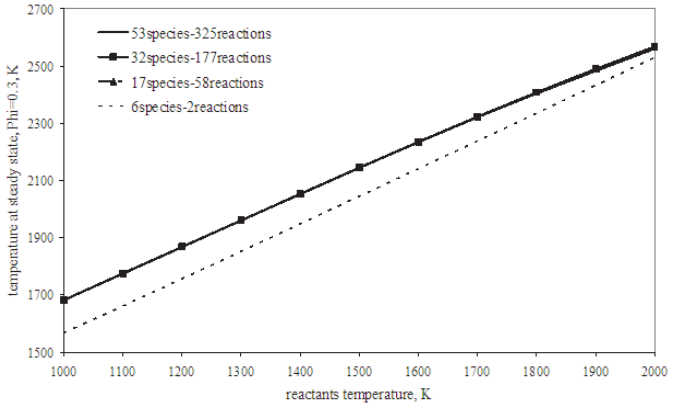


Fig. A11. Φ=0.3, P=5atm, temperature

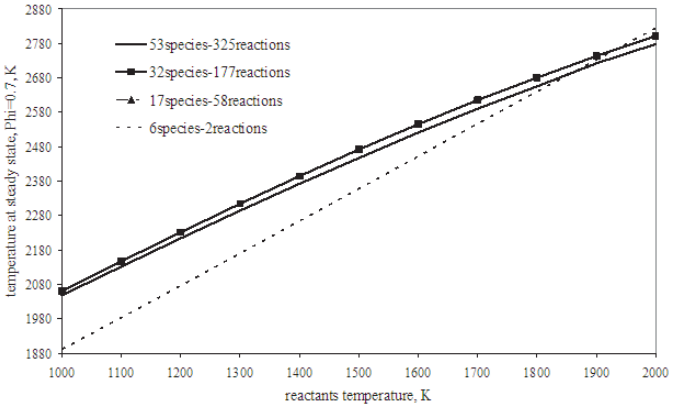


Fig. A12. Φ=0.5, P=5atm, temperature

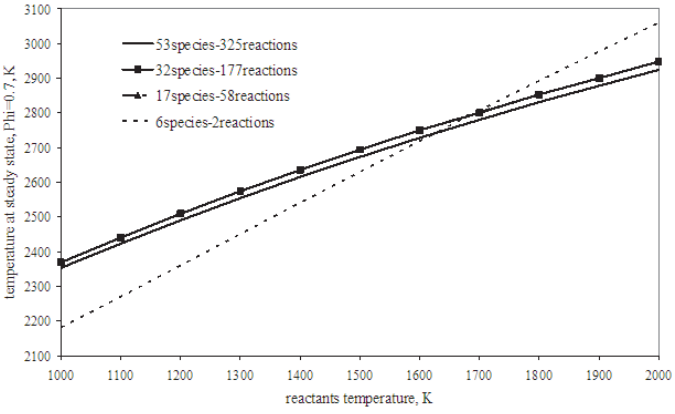


Fig. A13. Φ=0.7, P=5atm, temperature

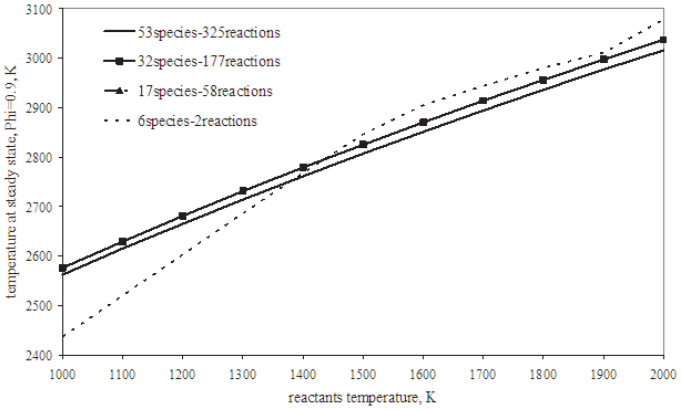


Fig. A14. Φ=0.9, P=5atm, temperature

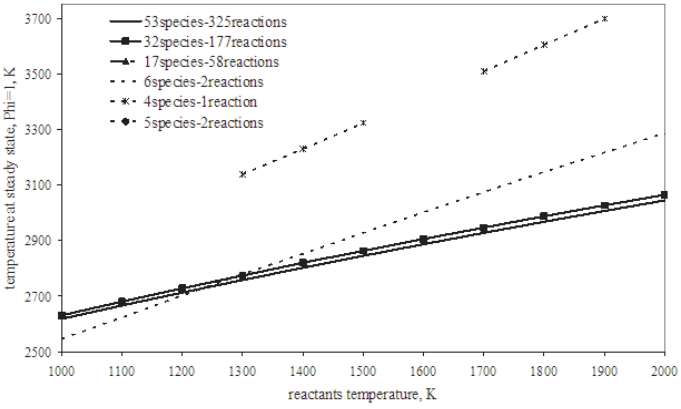


Fig. A15. Φ=1, P=5atm, temperature

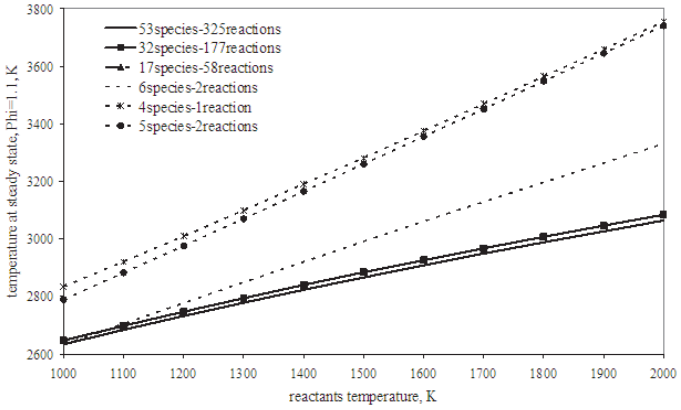


Fig. A16. Φ=1.1, P=5atm, temperature

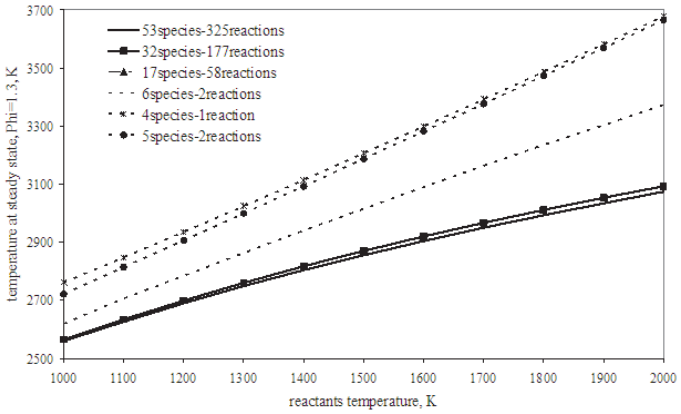


Fig. A17. Φ=1.3, P=5atm, temperature

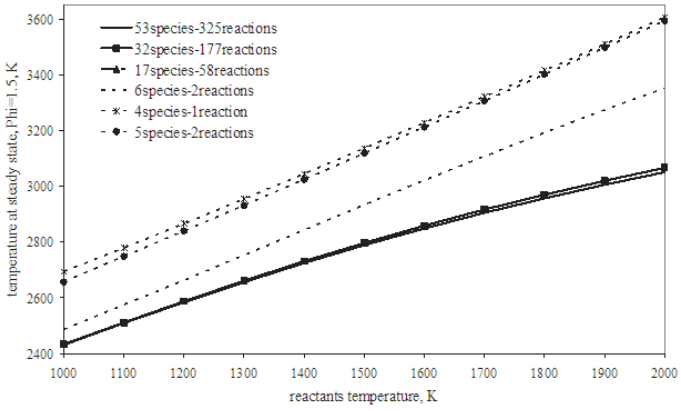


Fig. A18. Φ=1.5, P=5atm, temperature

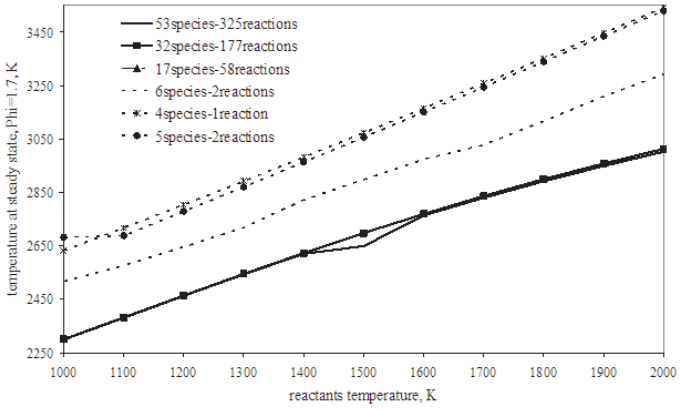


Fig. A19. Φ=1.7, P=5atm, temperature

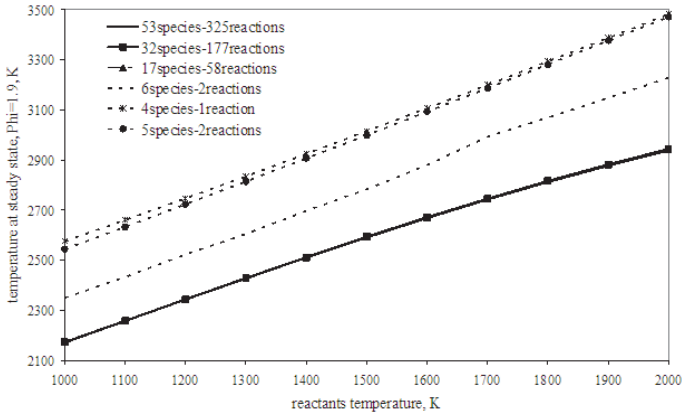


Fig. A20. Φ=1.9, P=5atm, temperature

8. References

Arione, L. (2010) "Development status of the LM10-MIRA engine for the LYRA Launch Vehicle", Proceedings of the "Space 2010" Conference, San Sebastian.

Bowman, C.T. (1986) Chemical Kinetics Models for Complex Reacting Flows, Ber. Bunsenges. Phys. Chem. 90, 934.

Bruno, C. (2001) "Chemical Microthrusters: Effect of Scaling on Combustion", AIAA Paper 2001-3711, AIAA JPC, Salt Lake City.

Bruno, C. (2003) Giacomazzi, E., Ingenito, A. Chemical Microrocket: Scaling and Performance Enhancement, EOARD Contract FA8655-02-M034, SPC 02-4034 Final report, London, UK.

Cozzi, F. (2007) Experimental Analysis of a Swirl Flow in a meso-scale combustor by LDV, XV A.I.VE.LA National Meeting, Milano, Italy.

Cozzi, F., Caratti, L. (2007) Temperature measurements by Spontaneous Raman Scattering in a meso-scale combustor, XV A.I.VE.LA National Meeting, Milano, Italy.

Cozzi, F., Coghe, A., Olivani, A., Rogora, M. (2007): Stability and Combustion Efficiency of a Meso-Scale Combustor Burning Different Hydrocarbon Fuels, 30th Meeting on Combustion, Ischia, Naple, Italy.

Dagaut, P., Boettner, J-C., Cathonnet, M. (1991) Methane Oxidation: Experimental and Kinetic Modeling Study, Combust. Sci. and Tech. 77, pp. 127-148.

DeGroot, W.A., and Oleson, S.R., (1996), "Chemical Microthruster Options", AIAA Paper 96-2863, presented at the AIAA Joint Propulsion Conference, Buena Vista, FL.

Gardiner, W.C. Jr. (1999) Gas-Phase Combustion Chemistry, Springer-Verlag, pp. 31-41.

GRI-Mech Version 1.2 released 11/16/94, http://www.me.berkley.edu/gri_mech/

GRI-Mech Version 3.0 released 07/30/99, http://www.me.berkley.edu/gri_mech/

Hurlbert, E., Angstadt, T., Villemarette, M., Collins, J., Allred, J., Mahoney, J., Peters, T., (2008) 870 lbf Reaction control system test using Lox/Ethanol and LOx/Methane at white sands test facility, AIAA 2008-5247, Hartford.

Heffington, W.M., Parks, G.W., Salzman, K.G.P., Penner, S.S. (1997) Studies of Methane Oxidation Kinetics, 16th Symp. (Int.) Combust. [Proc.], 997.

Hautman, D.J., Dryer, F.L., Shug, K.P., Glassman, I. (1981) A Multiple Step Overall Kinetics Mechanism for the Oxydation of Hydrocarbon, Combust. Sci. Technol. 25, 219.

Kawashima, H., Okita, K., Aoki, K., Azuma, N., Kumawaka, A., Onodera, T., Yoshida, S., Negeshi, H., Manako, H., Koganezawa, T. (2009) "Combustion and regenerative cooling characteristics of LOx/Methane engine", Transactions of the Japan Society for Aeronautical and Space Sciences, Space Technology Japan, Vol. 7, ists 26 (ISTS Special Issue: Selected papers from the 26th International Symposium on Space Technology and Science), pp Ta_7-Ta_11

Kee, R.J., Grcar, J.F., Smooke, M.D., Miller, J.A. (1985) A Fortran Program for Modeling Steady Laminar One-Dimensional Premixed Flames, Report No. SAND85-8240, Sandia National Laboratories.

Janson, S.W., (1994), "Chemical and Electric Micropropulsion Concepts for Nanosatellites", AIAA paper 94-2998, presented at the AIAA Joint Propulsion Conference, Indianapolis.

Minotti, A., Bruno, C., Cozzi, F. (2009) Numerical Simulation of a Micro Combustion Chamber, AIAA Aerospace Science Meeting and Exhibit, Orlando, FL, AIAA-2009-0447.

Mueller, J., (1997), "Thruster Options for Microspacecraft: A Review and Evaluation of Existing Hardware and Emerging Technologies", AIAA Paper 97-3058, presented at the AIAA Joint Propulsion Conference, Seattle, July 1997.

Paczko, G., Lefdal, P.M., Peters, N. (1988) Reduced Reaction Schemes for Methane, Methanol and Propane Flames, 21st Symp. (Int.) Combust. [Proc.], 739.

Stone, R., Tiliakos, N., Balepin, V., Tsai, C.-Y., Engers, R. (2008) "Altitude testing of LOx-Methane Rocket engine at ATK-GASL", AIAA 2008-3701, Seattle.

Trevino, C., Mendez, F. (1992) Reduced Kinetic Mechanism for Methane Ignition, 24TH Symp. (Int.) on Combust., The Combust. Inst., pp.121-127

Westbrook, K. C., Dryer, L. F. (1981) Simplified reaction mechanism for the oxidation of hydrocarbon fuels in flames, Combustion Science and Technology 27, pp. 31-43.

Part 3

Materials and Structures

14

Creep Behaviors and Influence Factors of FGH95 Nickel-Base Superalloy

Tian Sugui and Xie Jun
Shenyang University of Technology
China

1. Introduction

With the development of aerospace and ground transportation industry, the properties of aerospace engine are increasingly required to be improved. And especially, Turbine disk is one of the important component parts for advanced aero-engines, serving usually in the condition of 540～840°C at higher stress, so the turbine disks are required to be of high temperature tolerance and creep resistance (Flageolet B. et al., 2005; Liu D. M. et al., 2006). Thereinto, the continuous plastic flow of a material during creep can eventually result in large plastic deformations and significant modifications to the microstructure of the material, so that occurs the creep fracture of the turbine disk parts. Therefore, it is very important for the aero-engine materials to have a better property of the creep resistance.

Traditional wrought superalloys can hardly meet the requirements of the turbine disks in the advanced aerospace for their poor temperature tolerance and loading capacity resulted from their serious composition segregation in ingots and poor hot processability (Park N. K. & Kim I. S., 2001; Lherbier L. W. & Kent W. B., 1990; S. Terzi et al., 2008), especially the weaker cohesive force of grain boundaries (Paul L., 1988; Wang P. et al., 2008; Jia CH. CH. et al., 2006). While nickel based powder superalloys are an excellent material used for preparing the high temperature rotating section of the advanced aero-engine because its advantages are no macro-segregation in the ingot, chemical composition uniformity and high yield strength (Lu Z. Z. et al., 2005; Zhou J. B. et al., 2002).

FGH95 alloy is a nickel-base powder superalloy in which microstructure consists of γ matrix, γ' and carbide phases. The characteristics of FGH95 superalloy include the high extent of alloying and high volume fraction of γ'-phase (about 50%), besides the alloy possesses excellent integrating mechanical properties at 650°C (Domingue J. A., et al. 1980; Hu B. F. et al., 2006). Moreover, various size, morphology and distribution of γ' phase in the alloy can be obtained by different heat treatment regimes (Zainul H. D., 2007). The preparation technologies of FGH95 superalloy includes the powder pretreatment, hot isostatic pressing (HIP) and heat treatment. The heat treatment regimes of the alloy include the high temperature solution and twice aging treatment. After solution treated at high temperature, the alloy may adopt the different cooling methods, such as cooled in molten salt or in oil bath, and the microstructure and creep properties of the alloy are related to the chosen heat treatment regimes (Klepser C. A., 1995).

The deformation mechanism of the polycrystalline Ni-base superalloys during creep includes twinning, dislocations by-passing or shearing into the γ′ phase (Raujol S. et al., 2004 ; Kovarik L. et al., 2009). Actually, the mechanical properties and creep behaviors of the alloy are related to the quantities, morphology and distribution of γ′-phase, and especially, the configuration of the boundary and carbides have an important effect on the creep resistance of the alloy (Viswanathan G. B. et al., 2005; Hu B. F. et al., 2003). For example, after the alloy is solution treated for cooling in molten salt, the total number and size of secondary γ′ phase increase, which can effectively improve the plasticity of the alloy at high temperature (Zhang Y. W. et al., 2002). Because the various microstructures in the alloy may be obtained by different heat treatment regimes, it is very important to understand the influence of heat treatment regimes on the microstructure and creep resistance of the alloy.

In this chapter, the different HIP alloys are solution treated at different temperatures, and cooled in the molten salt or oil bath, respectively, then through twice aging treatment. Besides, some full heat treated alloys are aged for different time at high temperatures, and then the parameters of γ′, γ phases in the alloy are measured for evaluating the effect of the long term aging time on the misfits. The creep properties of the alloy treated by different heat treated regimes are measured under the conditions of the applied different temperatures and stresses, and the microstructures of the alloy are observed by using SEM and TEM for investigating the influences of the heat treatment regimes on the microstructure and creep properties. Additionally, the deformation mechanism and fracture feature of the alloy during creep are briefly discussed.

2. Experimental procedure

FGH95 powder particles of the nickel-base superalloy with the size of about 150 meshes were put into a stainless steel can for pre-treating at 1050 °C for 4 h. The can containing FGH95 powder alloy was hot isostatic pressed (HIP) for 4 h under the applied stress of 120 MPa at 1120 °C, 1150 °C and 1180 °C, respectively. The heat treatment and long term aged treatment regimes of the alloy are listed, respectively, in the Table 2.1 and Table 2.2. The cooled rates of the specimen in the oil bath and molten salt are measured to be about 205 °C/min and 110 °C/min, respectively. The error ranges of the used heating furnace are about ± 2 °C. The chemical composition of FGH95 superalloy is shown in Table 2.3.

HIP Temp., (°C)	Solution Temp., (°C)	Quenching Temp., (°C)	Double aging treatment
1120	1140 for 1 h 1150 for 1 h 1160 for 1 h 1165 for 1 h	cooled for 15 min in molten salt at 583 °C	870 °C×1h + 650 °C×24h
1180	1150 for 1 h 1160 for 1 h	cooled for 15 min in oil bath at 120 °C	
	1150 for 1 h 1160 for 1 h	cooled for 15 min in molten salt at 583 °C	

Table 2.1. Heat treatment regime of FGH95 superalloy

HIP Temp., (°C)	Heat treatment	Aging regime
1120	1155 for 1 h + cooled for 15 min in molten salt at 550 °C + 870 °C×1h +650 °C×24 h	450 °C×500 h 450 °C×1000 h 550 °C×500 h 550 °C×1000 h
1150 1180		----

Table 2.2. Long time aging treated regime of FGH95 superalloy

C	B	Cr	Co	Al	Ti	W	Mo	Nb	Ni
0.060	0.012	12.98	8.00	3.48	2.55	3.40	3.40	3.50	Bal

Table 2.3. Composition of FGH95 superalloy（mass fraction,%）

By means of the anode selective dissolving method, the volume fraction of γ′-phase in FGH95 alloy was measured to be about 47%. Thereinto, the electrolytic extraction of γ′ phase in the alloy was conducted for separating from the γ matrix under the condition of the temperature at 0°C and current density about 50mA/cm^2. The choosing electrolyte solution consisted of $(NH4)_2SO_4$ and citric acid, the experimental device of the electrolytic extraction was shown in Fig. 2.1. After the electrolytic extraction was conducted, the granularity distribution, phases constituting and the misfit of γ′, γ phases in the alloy were measured by means of the XRD analysis and SEM/EDS observation.

The ingot of FGH95 superalloy was cut into the specimens with the cross-section of 4.5 mm × 2.5 mm and the gauge length of 20 mm, and the size of the sample was shown in Fig. 2.2. Uniaxial constant load tensile testing was performed, in a GWT504 model creep testing machine, for measuring creep curves under the experimental conditions of 984 MPa ~ 1050 MPa and 630 °C ~ 670 °C. The yield strength of FGH95 alloy was measured to be 1110 MPa at 650 °C. The strain data of the alloy at different conditions were measured with an extensometer to portray the creep curves, twice of the each creep testing were conducted for ensuring the statistical confidence. The specimens of FGH95 alloy at different states were grinded and polished for observing the microstructure by using SEM and TEM, so that the influence of the heat treatment technics on the microstructure, the creep feature and fracture mechanism of the alloy was investigated.

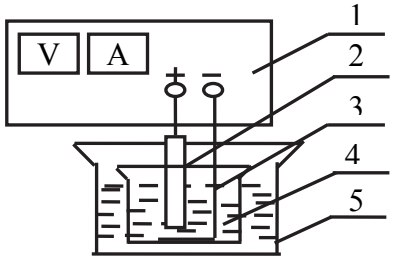


Fig. 2.1. Experimental equipment
1--Power supply, 2--sample, 3-- cathode, 4--solution, 5--container

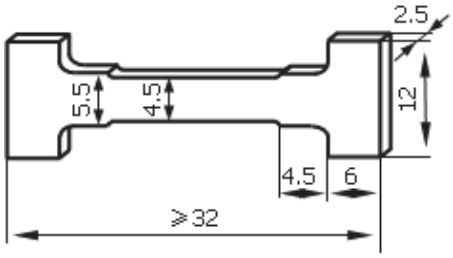


Fig. 2.2. Schematic diagram of the tensile creep sample

3. Influence of HIP temperatures on microstructure

3.1 Influence of HIP temperatures on microstructure of FGH95 alloys

After the alloy was hot isostatic pressed at different temperatures, the microstructure of the HIP alloys was shown in Fig. 3.1. The regions which were encompassed by coarse γ′ phase were defined as previous powder particles. After 1120 °C hot isostatic pressing molded, the alloy consisted of γ′ and γ phases, thereinto, the coarser γ′ phase distributed around the powder particles was defined as the previous particle boundaries (PPB). Therefore, the configuration of sphere-like previous particle was clearly appeared, and the power particle size was about 15~25 μm, as shown in Fig. 3.1(a).

With the HIP temperature increased to 1150 °C, the size of the powder particles was similar to the former, but the sphere-like configuration was not clearly. The PPBs consisted mainly of the coarse γ′ phase, and the size and amount of the coarser γ′ phase decreased slightly as shown in Fig. 3.1(b). As the HIP temperature increased to 1180 °C, the size of the grain grew up obviously, being about 20~40 μm. Besides, the grain boundaries appeared the straight-like feature, and the amount and size of the coarse γ′ phase decreased obviously as shown in Fig. 3.1(c). The dark regions around the powder particles were defined as the previous particle boundaries (PPB) in which the secondary γ′ phase was precipitated along the different orientations as marked by letters A and B.

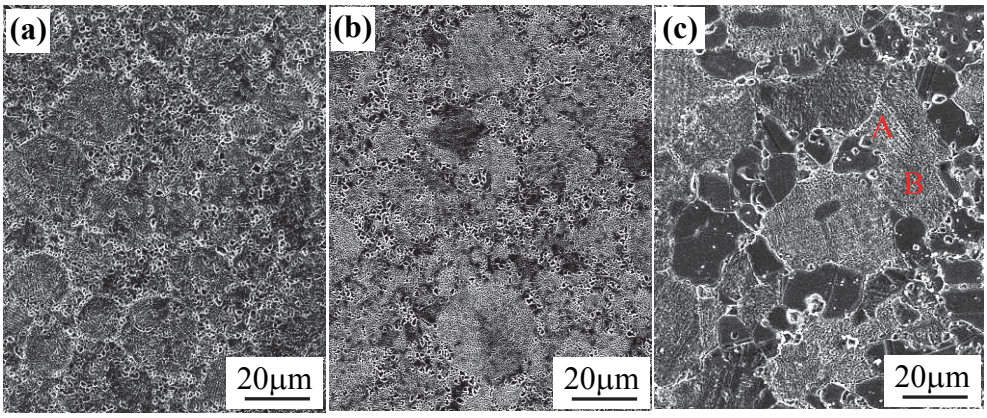


Fig. 3.1. Microstructure of the alloy after HIP treated at different temperatures.
(a) 1120 °C,(b) 1150 °C,(c) 1180 °C

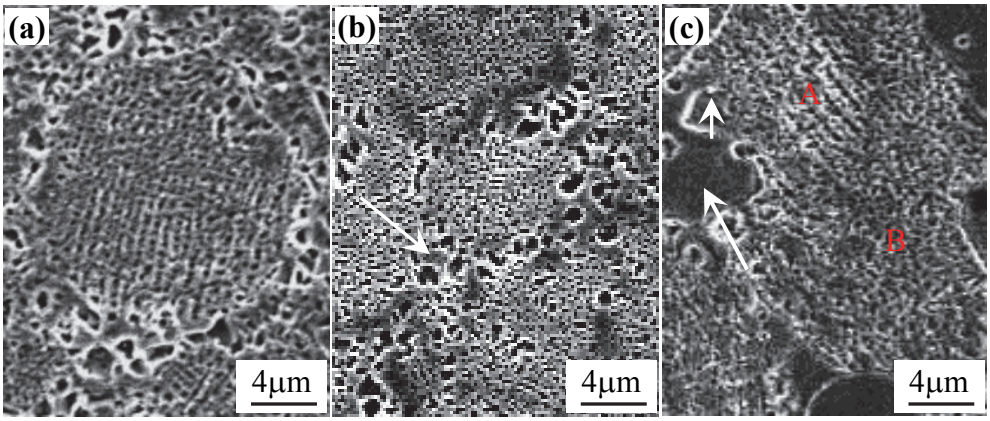


Fig. 3.2. Magnified morphology of the alloy HIP treated at different temperatures. (a) finer γ′-phase precipitated within the grain of the alloy treated by HIP at 1120 °C, (b) after HIP treated at 1150 °C, no γ′- phase particles precipitated in the regions near the coarser γ′-phase as marked by arrow, (c) after the alloy treated by HIP at 1180 °C, particle-like carbides precipitated along the boundary as marked by short arrow.

The magnified morphology of the alloys which were hot isostatic pressed at different temperatures was shown in Fig. 3.2. Thereinto, the γ′ phase displayed the black color due to the dissolved during chemical corrosion, and the γ matrix which is not dissolved displays the gray color. After HIP treated at 1120 °C, the coarser γ′ phase which distributed around the PPB was about 1~2 μm in size and was defined as the primary γ′ phase. Besides, the fine secondary γ′ phase was regularly distributed along the same orientation within the previous powder particles as shown in Fig. 3.2(b). As the HIP temperature increased to 1150 °C, the secondary γ′ phase about 0.1~0.3 μm in size was dispersedly precipitated within the grain, and the coarser γ′ phase still existed in the PPB regions. Moreover, the depleted zone of the fine γ′-phase appeared in the regions near the coarser γ′ phase, as marked by the arrow in Fig. 3.2(b).

The magnified morphology of the powder particle in the 1180 °C HIP alloy was shown in Fig. 3.2(c), indicating that the fine secondary γ′ particles with different orientations were precipitated within the same grain as marked by A and B in Fig. 3.2(c). No fine γ′ particles were precipitated in the PPB regions near the coarser γ′ phase, so the region was defined as the depleted zone of the fine γ′-phase as marked by the long arrow in Fig. 3.2(c). Moreover, some white carbide particles were precipitated in the PPB region as marked with the white short arrow in Fig. 3.2(c).

The microstructure of the different temperature HIP alloys which were solution treated at 1155 °C, cooled in the molten salt at 520 °C and twice aging treated was shown in Fig. 3.3. The microstructure of the 1120 °C HIP alloy after full heat treated was shown in Fig. 3.3(a), illustrating that a few of coarser γ′ phase was distributed along the grain boundaries. And the size of the coarse γ′ phase was about 1 ~ 2 μm, as marked with the white short arrow in Fig. 3.3(a). Besides, the fine γ′ phase was dispersedly precipitated within the grain as marked with the long arrow in Fig. 3.3(a). Comparing to Fig. 3.3(a), the microstructure of the 1150 °C HIP after full heat treatment had no obvious distinction to the former, and the grain size

was about 10~25μm. The secondary γ′ phase was dispersedly distributed in the alloy as shown in Fig. 3.3(b). With the HIP temperature increased to 1180 °C and after full heat treatment, the PPB trace was kept in the alloy as marked with the long arrow in Fig. 3.3(c), indicating that the grain grew up obviously after HIP treated at 1180 °C, and the coarse γ′ phase dissolved completely, only kept a few of primary γ′ phase which size was about 1μm in the grain boundaries as marked with the short arrow in Fig. 3.3(c). Moreover, some fine carbide particles were dispersedly precipitated in the alloy as shown in Fig. 3.3(c).

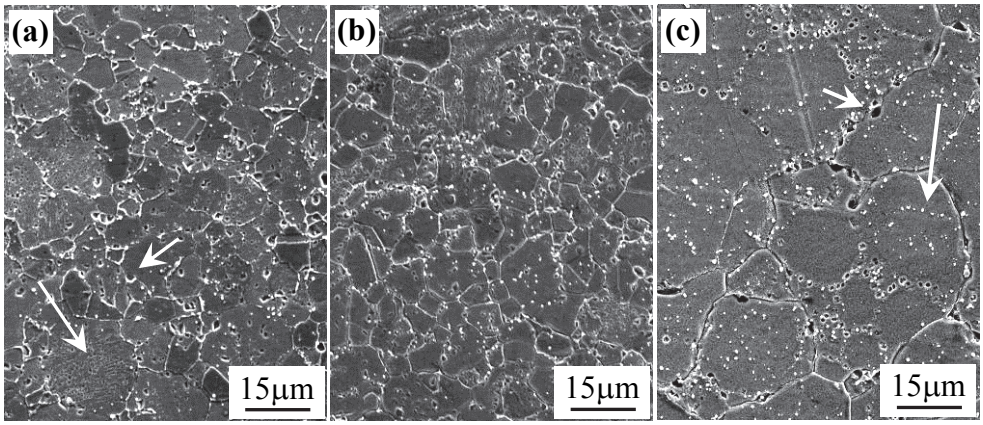


Fig. 3.3. Microstructure of the different temperatures HIP alloy after fully heat treated. (a) 1120 °C, (b) 1150 °C, (c) 1180 °C

The magnified morphology of the different temperature HIP alloy after fully heat treated was shown Fig. 3.4. After the 1120 °C and 1150 °C HIP alloys were fully heat treated, a few of coarse γ′ phase was precipitated along the boundary regions as marked with short arrow in Fig. 3.4(a) and (b). And the coarse γ′ phase appeared in the boundary as shown in Fig. 3.4(a).

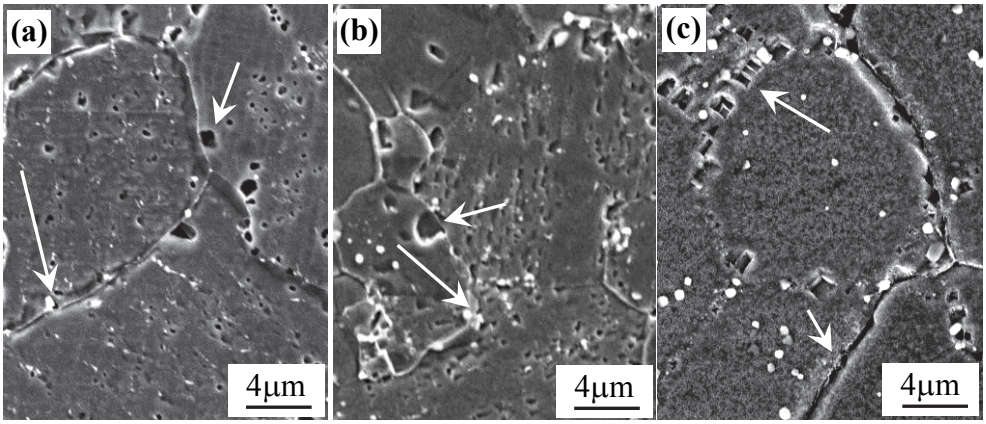


Fig. 3.4. Magnified morphology of the different temperature HIP alloy after fully heat treated (a) 1120 °C, (b) 1150 °C, (c) 1180 °C

Moreover, fine white particles were discontinuously precipitated along the boundaries as marked with the long arrow in Fig. 3.4(a) and (b), and the fine secondary γ′ phase and white particles were dispersedly distributed in the alloy as shown in Fig. 3.4(a) and (b). With the HIP temperature increased to 1180 °C and after fully heat treated, the sphere-like PPB kept in the alloy as marked with long arrow in Fig. 3.4(c), and the boundaries displayed an obvious straight feature as marked with short arrow in Fig. 3.4(c). Moreover, fine γ′ phase was dispersedly distributed within the grain, and the white particles of about 0.2 μm in size were dispersedly precipitated in the grain as shown in Fig. 3.4(c).

3.2 Analysis on the influence of HIP temperatures on microstructure

When the FGH95 alloy was HIP treated at 1120 °C and 1150 °C which is lower than the dissolving temperature of γ′ phase (T_m= 1160 °C), the coarse γ′ phase distributed around the PPB were not completely dissolved. Therefore, the γ′ phase which was reserved in the PPB regions was going to grow up during the slowly cooling stage of HIP, which resulted in forming the morphology of the coarser γ′ particles distributed along the PPB. The size of previous powder particles were kept in the alloy because no enough deformed energy which come from the HIP deformation promoted the powder particles growth up, as shown in Fig. 3.1(a) and (b).

When the alloy was hot isostatic pressed under the conditions of 120 MPa at 1180 °C, the coarser γ′ phase in the PPB regions was completely dissolved, which may enhance the diffusion rate of the elements and reduce the transferring resistance of the boundaries to promote the growth of the grain, therefore, the alloy displayed a larger grain size as shown in Fig. 3.1(c). At the same time, the nucleation and growth of the carbide occurred on the fine oxides which existed in the powder particle surface, and the carbide nucleation process only needs lower energy (Domingue J. A. et al., 1980). Consequently, some white carbide and coarser γ′ particles were discontinuously precipitated in the PPB regions as shown in Fig. 3.2(c), and the depleted zones of the fine γ′ phase appeared in the region near the coarse γ′ particles, due to the coarsening of γ′ particles to consume the much more elements of γ′ former, as shown in Fig. 3.2(c).

4. Influence of solution temperatures on microstructure

4.1 Influence of solution temperature on the microstructure

After the 1120 °C HIP alloy was solution treated at 1140 °C and twice aged, the grain size of the alloy was homogeneous, and fine γ′ phase was dispersedly precipitated within the grain, as shown in Fig. 4.1(a). Besides, much more coarse γ′ phase precipitated along the grain boundaries as marked with the arrow in Fig. 4.1(a). The magnified morphology was shown in Fig. 4.1(b), significant amount of fine γ′ phase were dispersedly distributed in the alloy, and the depleted zone of γ′-phase appeared around the coarse γ′ phase which distributed along the boundary regions, as shown in the Fig. 4.1(b).

With the solution temperature increased to 1150 °C, the microstructure of the alloy consisted of the γ′ and γ phases, the average grain size of the alloy was about 10~20 μm, as shown in Fig. 4.2(a), which indicated that some coarser γ′-particles are precipitated in the wider boundary regions, and the average size of the coarser γ′ phase was about 1~2.5 μm. The magnified morphology of the alloy was shown in Fig. 4.2(b), indicating that significant amount of the fine γ′ particles were dispersedly distributed within the grains, the size of the

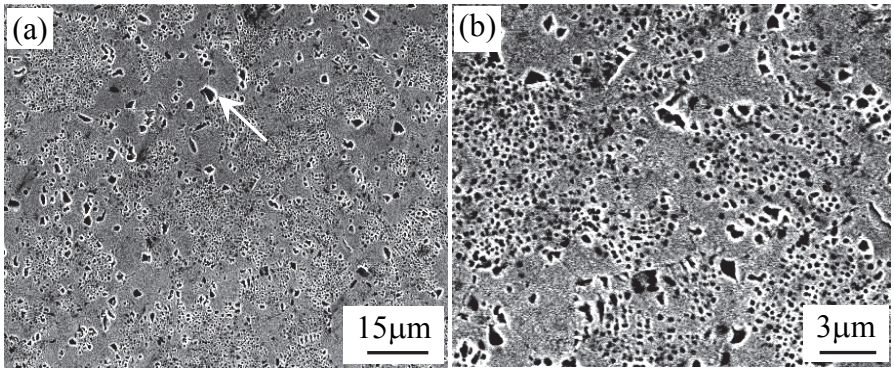


Fig. 4.1. Microstructure of the alloy after solution treated at 1140 °C. (a) after solution treated at 1140 °C, coarse γ′ phase distributed in the alloy, (b) no finer γ′ phase precipitated in the regions near the coarser γ′ phase

ones was about 0.1~0.3 μm. And no fine γ′-particles were precipitated in the regions near the coarser γ′-phase, the regions were defined as the depleted zone of the fine γ′-phase as marked by the arrow in Fig. 4.2(b), the magnified morphology of the depleted zone of γ′-phase was marked by the arrow in Fig. 4.2(c).

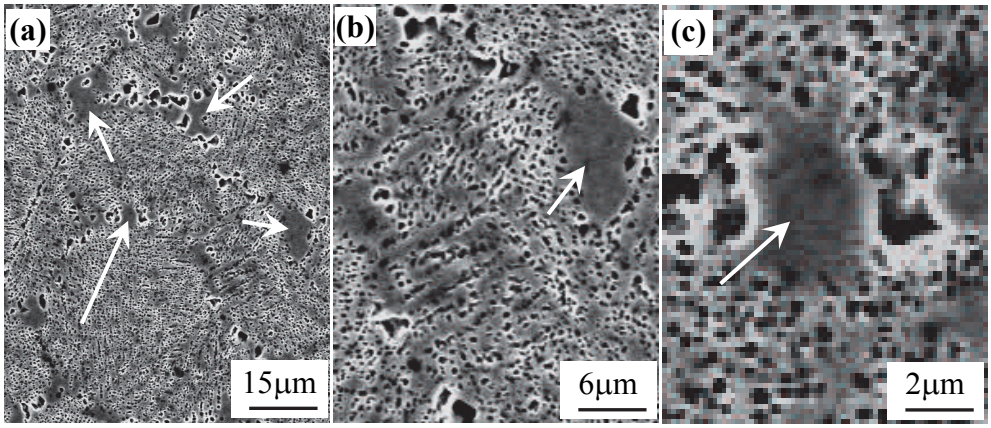


Fig. 4.2. Microstructure of the alloy after solution treated at 1150 °C. (a) After solution treated at 1150 °C, some coarser γ′ phase precipitated in the wider boundary regions, (b) fine γ′ phase distributed dispersedly within the grains, and the depleted zones of the fine γ′ phase marked by arrow, (c) magnified morphology of the depleted zone of γ′-phase

As the solution temperature increased to 1160 °C, the grain boundaries appeared obviously, and the average size of the grains in the alloy was about 15 ~ 25 μm as shown in Fig. 4.3(a). Compared to Fig. 4.2(a), the coarser γ′-precipitates in the boundary regions disappeared, and the white particles with size about 0.2 μm were precipitated within the grains and along the boundaries as marked by the arrows in Fig. 4.3(a). The magnified morphology of the alloy was shown in Fig. 4.3(b), indicating that the coarser γ′ particles and depleted zone of the fine

γ′ phase had disappeared, and the secondary γ′ phase was dispersedly distributed within the grains. The grain boundary in the alloy was marked by longer arrow and some particles were homogeneously precipitated along the boundaries and within the grains as marked by the short arrows in Fig. 4.3(b).

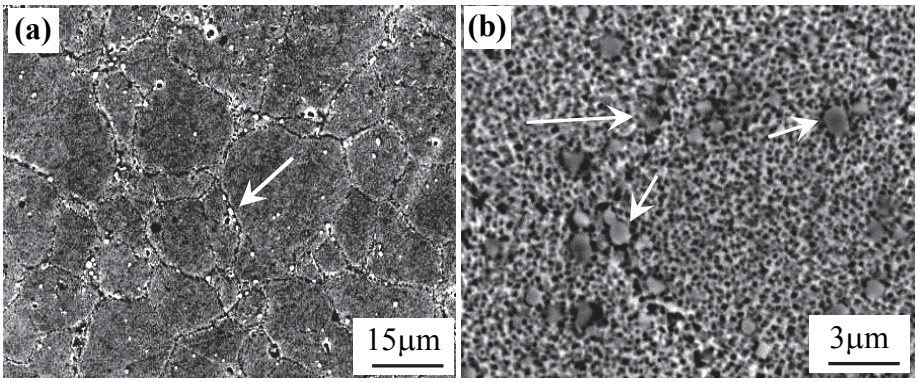


Fig. 4.3. Microstructure of the alloy after solution treated at 1160 °C. (a) after solution treated at 1160 °C, the coarser γ′ phase disappearing and the grain boundaries appearing obviously in the alloy, (b) the secondary γ′ phase distributed dispersedly within the grain, and some particles precipitated in the alloy as marked by the arrows

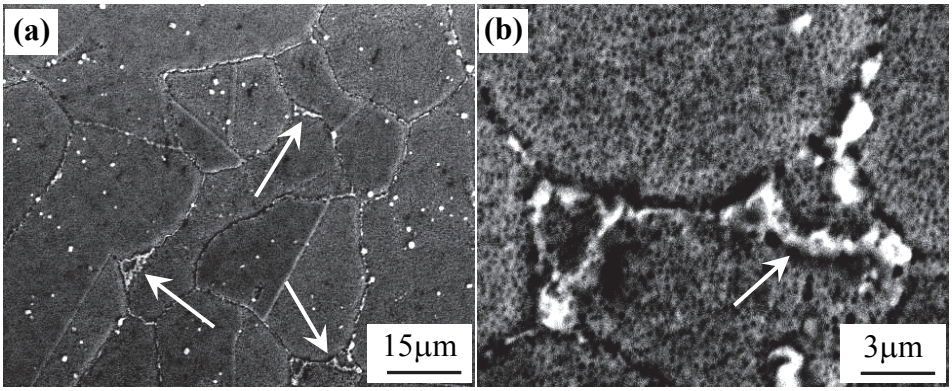


Fig. 4.4. Microstructure of the alloy after solution treated at 1165 °C. (a) After solution treated at 1165 °C, the average size of the grains in the alloy increasing obviously, and films of white phase precipitated along the boundaries as marked by arrows, (b) the secondary γ′ phase distributed dispersedly within the grains, and films of the white phase distributed along the grain boundaries

As the solution temperature increased to 1165 °C, the grain sizes in the alloy increased, the linear-like boundaries appeared in the alloy, and the films of the white phase were continuously precipitated along the boundaries as marked by the arrows in Fig. 4.4(a). The magnified morphology of the white phase was shown in Fig. 4.4(b), indicating that white particles were continuous precipitated to form the films along the boundaries as marked by

the arrow in Fig. 4.4(b), and significant amount of the fine secondary γ′ phase were precipitated within the grains, no depleted zones of the fine γ′ phase were detected in the alloy. The grain sizes after the alloy was solution treated at various temperatures were measured as listed in Table 4.1. Moreover, by means of composition analysis under SEM/EDS, it is indicated that the elements Nb, Ti and C are richer in the white particles which were located within the grain and boundary regions as shown in Fig. 4.3 and Fig. 4.4, respectively.

After the alloy was solution treated at 1160 °C, the fine γ′-particles with the size of about 0.1 μm were dispersedly precipitated within the grains as shown in Fig. 4.5(a), the particles can effectively hinder the dislocation movement to enhance the creep resistance of the alloy. The smaller space between the fine γ′-particle was measured to be about 0.03 μm, the bigger space between the fine γ′-particles is measured to be about 0.12 μm as marked by letters L_1 and L_2 in Fig. 4.5(a), respectively. It was indicated by TEM/EDS analysis that the elements Nb, Ti and C richer in the carbide particle which was located in the boundary as shown in Fig. 4.5(b), and the particle was identified as (Nb, Ti)C phase by means of the diffraction spots analysis as marked in Fig. 4.5(c).

Solution temp. (°C)	1140	1150	1160	1165
Average grain sizes (μm)	10~20	10~20	15~25	20~40

Table 4.1. Grain sizes of the alloy solution treated at different temperatures

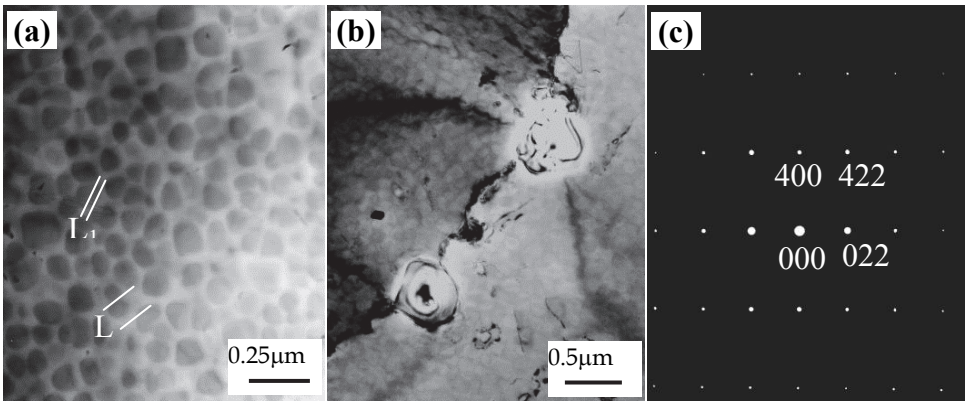


Fig. 4.5. Morphologies of γ′ and carbide phases. (a) Fine γ′ phase precipitated dispersedly within the grain, (b) carbide particles precipitated along the boundaries, (c) SAD patterns

4.2 Influence of quenching method on the microstructure

After the 1180 °C HIP alloy was solution treated at different temperatures and cooled in oil bath at 120 °C and twice aging treated, the alloy displayed the various sizes of grains in which the fine γ′ phase was dispersedly and regularly precipitated, as marked by white long arrow in Fig. 4.6(a). The bigger grains was about 40~60 μm in size, the smaller grains was about 20 μm in size, the bunch-like coarser γ′ particles were distributed in the PPB regions, or some coarser γ′ particles were congregated in the local regions as marked by black arrow in Fig. 4.6(a). The depleted zone of the fine γ′-phase appeared in the regions near the coarser γ′ phase as marked with the white short arrow in Fig. 4.6(a). As the solution temperature

increased to 1160 °C, the coarser γ′ phase in the alloy was completely dissolved due to the higher solution temperature. Moreover, the fine γ′ phase and the white particles were dispersedly distributed within the grains, respectively, and some white particles were precipitated along the boundaries as marked by white arrow in Fig. 4.6(b). Furthermore, the twinning appeared within the grain as marked by black arrow in Fig. 4.6(b).

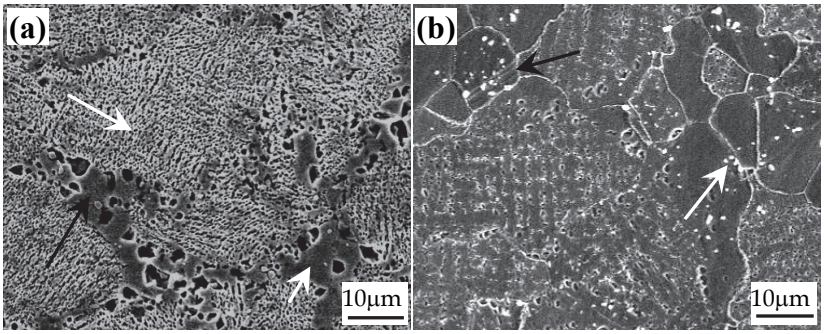


Fig. 4.6. Morphology of the alloy after solution treated at different temperatures and cooled in oil bath. (a) After solution treated at 1150 °C, the fine γ′ phase distributed dispersedly within the grains, and the depleted zone of γ′-phase as marked by shorter arrow. (b) after solution treated at 1160 °C, the fine γ′ phase and carbide particles distributed dispersedly within the grains, respectively

After solution treated at 1150 °C, then cooled in molten salt and twice aged, the microstructure of the alloy consisted of the particle-like γ′phase with different size, the coarser γ′-phase was distributed in the PPB regions as marked by white arrow, and the fine γ′ phase was dispersedly distributed within the grain as marked by black arrow in Fig. 4.7(a). Furthermore, the depleted zone of the fine γ′ phase appeared in the region near the coarser γ′-phase as marked by white short arrow in Fig. 4.7(a), and the size of the fine γ′ phase was about 0.15 μm.

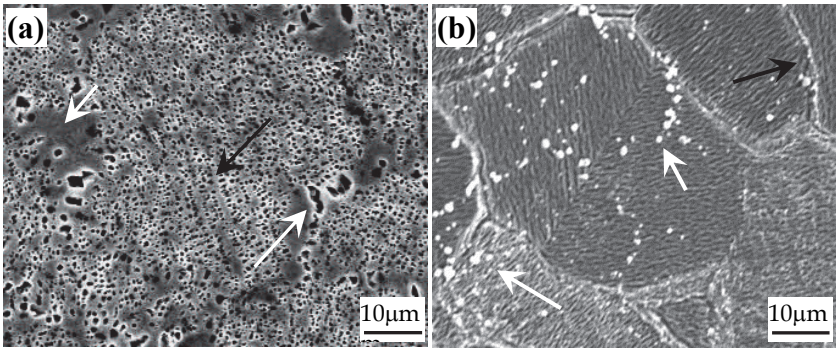


Fig. 4.7. Morphology of the alloy after solution treated at different temperatures, and cooled in molten salt. (a) After solution treated at 1150 °C, the fine γ′ phase distributed dispersedly in the grains, and the coarse γ′ particles precipitated along the boundaries, (b) after solution treated at 1160 °C, straight-like boundaries marked by black arrow, and the fine particles distributed dispersedly within the grains.

5. Influence of aging time on microstructure and misfits

5.1 Influence of aging time on microstructure

After the 1120 °C HIP alloy was solution treated at 1155 °C, cooled in molten salt at 520 °C and twice aging treated, the microstructure of the full treated alloys which were kept at 450 °C and 550 °C for 500 h respectively was shown in Fig. 5.1(a) and (b). Compared to the alloy which was not long time aged, no obvious distinction on the microstructure was detected in the alloy which was long time aged for 500 h at 450 °C, as shown in Fig. 5.1(a). Thereinto, no obvious change on the grain size was found, and a few coarse γ′ particles about 1~2 μm in size was precipitated along the boundaries as shown in the Fig. 5.1(a). With the long time aging temperature increased to 550 °C, the grain size about 15~20 μm was similar to the former, and some coarse γ′ phase were distributed along the boundaries as shown in Fig. 5.1(b).

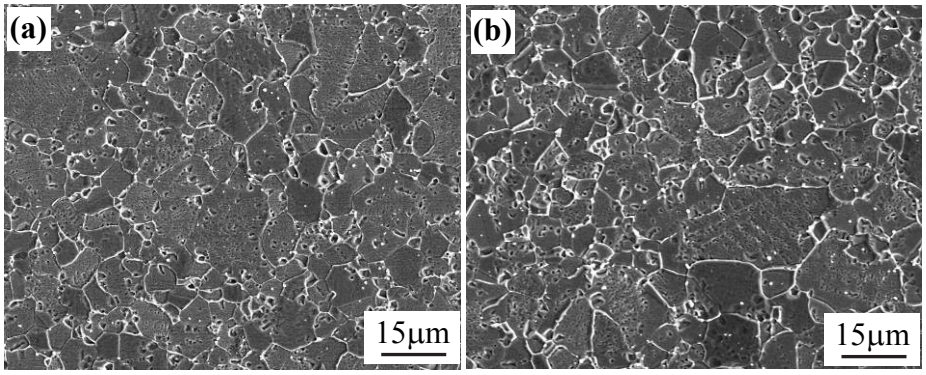


Fig. 5.1. Microstructure of the alloy aged for 500 h at different temperatures. (a) 450 °C, (b) 550 °C

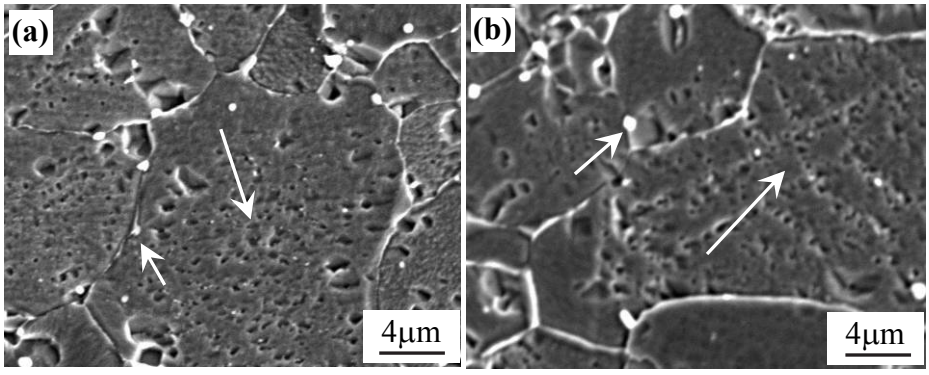


Fig. 5.2. Magnified morphology of the alloy aged for 500 h at different temperatures. (a) 450 °C, (b) 550 °C

The magnified morphology of the alloys aged for 500 h at 450 °C and 550 °C was shown in Fig. 5.2(a) and (b), respectively. After the alloy aged for 500 h at 450 °C, fine white carbide

particles were precipitated along the boundaries as marked with short arrow in Fig. 5.2(a). Moreover, the secondary γ′ phase about 0.1~0.2 μm in size was dispersedly distributed in the grain as marked with the long arrow in Fig. 5.2(a). With the temperature increased to 550 °C, the microstructure of the alloy was similar to the former, and the fine carbide particles about 0.1~0.25 μm in size were distributed along the boundaries as marked with the short arrow in Fig. 5.2 (b). Besides, the fine γ′ precipitates were dispersedly distributed in the alloy as marked with the long arrow in Fig. 5.2(b).

The microstructure of the full heat treated alloys aged for 1000 h at 450 °C and 550 °C was shown in Fig. 5.3(a) and (b), respectively. Fig. 5.3(a) showed that the microstructure was similar to the one of the aged-free alloy, and the grain size was about 15~20 μm. Moreover, a few coarse γ′ phase was precipitated along the grain boundaries, and the size of γ′ precipitates was about 1~2 μm as shown in Fig. 5.3(a). As the aging temperature increased to 550 °C and kept for 1000 h, no obvious distinction on the size of the grain and coarse γ′ phase was detected, as shown in Fig. 5.3(b).

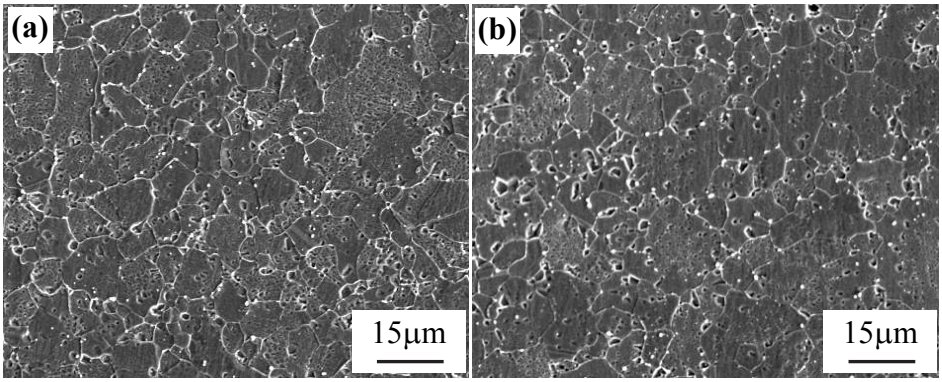


Fig. 5.3. Microstructure of the alloy aged for 1000 h at different temperatures. (a) 450 °C, (b) 550 °C

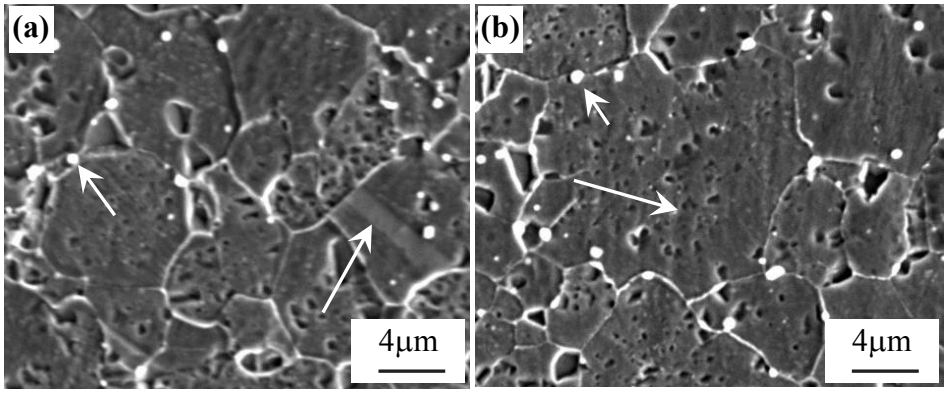


Fig. 5.4. Magnified morphology of the alloy aged for 1000 h at different temperatures. (a) 450 °C, (b) 550 °C

After aged for 1000 h at 450 °C and 550 °C, the magnified morphology of the alloy was shown in Fig. 5.4. Figure 5.4(a) shown that, after aged for 1000 h at 450 °C, some fine white carbide particles were precipitated along the boundaries as marked by the short arrow in Fig. 5.4 (a). Besides, twinning appeared obviously in the alloy as marked with the long arrow in Fig. 5.4(a). With the aging temperature increased to 550 °C, the microstructure of the alloy was similar to the former, and some fine white carbides distributed along the boundaries as marked with the short arrow in Fig. 5.4(b). Moreover, the secondary γ′ phase about 0.1~0.2 μm in size was dispersedly precipitated in the grain as marked with the long arrow in Fig. 5.4(b).

5.2 Influence of aging time on misfits

After FGH95 alloy was full heat treated, the microstructure of the alloy consisted of γ-matrix, γ′ phase and carbide particles. Thereinto, the ordered γ′ phase was coherently embedded in the γ matrix. Besides, both γ and γ′ phases have the FCC structure and close lattice parameters. Moreover, the coherent interfaces was kept in between the γ′/γ phases, and the certain misfits occurred in the interfaces of γ′/γ phases. The equations of calculating the parameters and misfit may be expressed as follow:

$$2d\sin\theta = \lambda \tag{5.1}$$

$$a = d\sqrt{\mathrm{h}^2 + \mathrm{k}^2 + \mathrm{l}^2} \tag{5.2}$$

$$\delta(\%) = \frac{2(a_{\gamma'} - a_{\gamma})}{(a_{\gamma'} + a_{\gamma})} \times 100\% \tag{5.3}$$

Where, d is crystal plane distance, θ is diffraction angle, λ = 0.154 nm, α is lattice parameter, (h, k, l) is crystal plane indexes and δ is the lattice misfit.

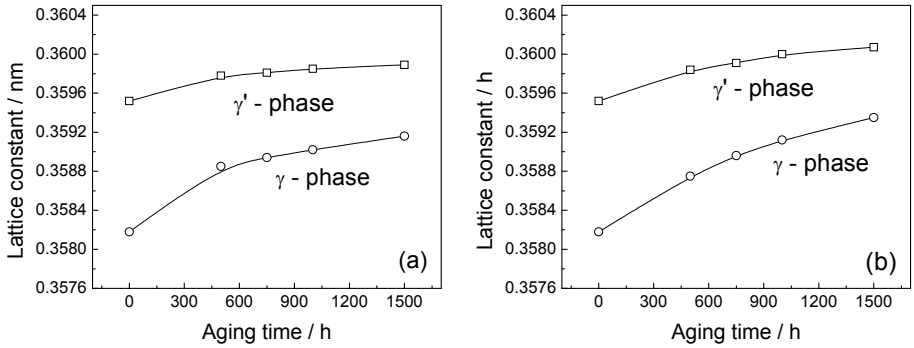


Fig. 5.5. Dependence of aging time and lattice parameters of γ/γ′ phases at different temperatures. (a) 450 °C, (b) 550 °C

After the fully heat treated alloys aged for 500 h, 750 h, 1000 h and 1500 h at 450 °C and 550 °C, the XRD curves of the alloys were measured, in which the located angles were

chosen to separate the diffraction peaks of γ and γ′ phases for measuring the crystal plane distance of γ and γ′ phase in the alloy. In the further, the lattice parameters and misfits of the alloys were calculated according to the equations (5.2) and (5.3), and the dependence of the lattice parameters and aged time at different temperatures was shown in Fig. 5.5. Figure 5.5(a) shown that the parameters of γ and γ′ phases increased gradually with the aging time at 450 °C, and the parameter increment of γ phase was more than the one of γ′ phase. With the aging temperature increased to 550 °C, the trend of the parameter increment was similar to the former, but the parameter increment of γ and γ′ phases was larger than the former, as shown in Fig. 5.5(b).

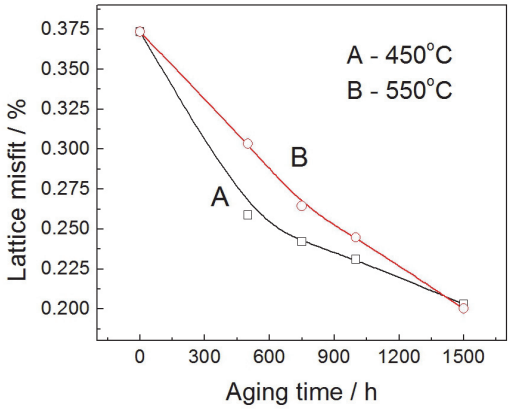


Fig. 5.6. Relationship between the lattice misfit and aging time at different temperatures

The relationship between the lattice misfit and aging time at different temperatures was shown in Fig. 5.6. This indicated that the lattice misfits decreased from 0.3734% to 0.2030% as the aging time prolonging at 450 °C as marked with letter A in Fig. 3.6. With the aging temperature increased to 550 °C, the lattice misfit decreased from 0.3734% to 0.2002% as marked with letter B in Fig. 3.6. Thereinto, the bigger difference between the lattice misfits of the alloys aged at 450 °C and 550 °C appeared in the periods of aging for 500 h and 1000 h.

6. Creep behaviors and relative parameters

6.1 Effect factors of solution temperature on creep properties

6.1.1 Influence of solution temperature on creep properties

After the 1120 °C HIP alloy solution treated at 1160 °C and cooled in molten salt at 583 °C and twice aging treated, creep curves of the alloy under different conditions are shown in Fig. 6.1. Figure 6.1(a) was the creep curves measured under different applied stresses at 650 °C. Under the applied stress of 1020 MPa, the alloy displays the shorter initial creep stage and longer steady creep stage, besides, the alloy displays the lower strain rate and longer creep lifetime about 156 h. With the applied stress enhanced to 1034 MPa, the strain rate of the alloy during steady state creep increased, and the lifetime of the alloy decreased to 104 h. As the applied stress enhanced to 1050 MPa, the creep lifetime of the alloy decreased sharply to 51 h. This indicates that, in the ranges of the applied stresses, the alloy is not sensitive to the applied stress at 650 °C.

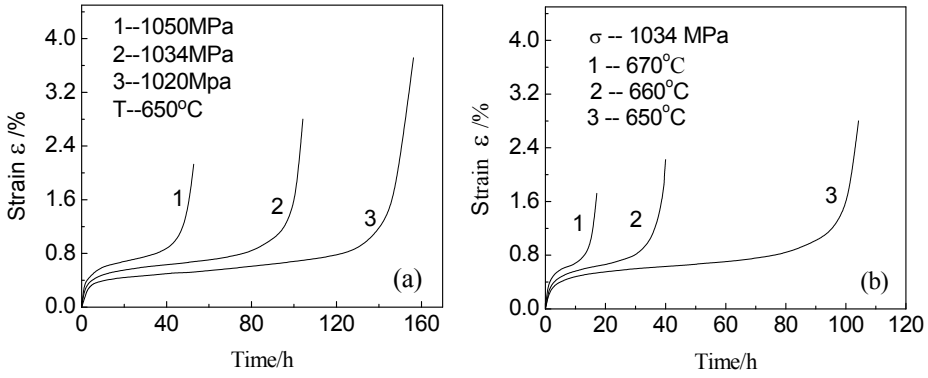


Fig. 6.1. After fully heat treated, creep curves of 1120 °C HIP alloy at different conditions (a) Applied various stresses at 650 °C, (b) 1034 MPa at various temperatures

The creep curves of the alloy under the conditions of 1034 MPa and different temperatures are shown in Fig. 6.1(b), illustrating that the alloy possesses the lower strain rate and longer creep lifetime under the applied stress of 1034 MPa at 650 °C, and the strain rate of the alloy during steady state creep was measured to be about 0.00367%/h, the creep lifetime of the alloy was about 104 h. With the creep temperature increased to 660 °C, the strain rate of the alloy during steady state creep increased to 0.00825%/h, and the creep lifetime decreased sharply to 40 h. As the temperature increased to 670 °C, the strain rate during steady state creep further enhanced to 0.01792%/h, and the lifetime of the alloy decreased to 17 h. This indicates that the alloy has an obvious sensitive to the applied temperature.

Under the applied stress of 1034 MPa at 650 °C, creep curves of the alloy solution treated at different temperatures were shown in Fig. 6.2. The creep curve of the alloy solution treated at 1150 °C was marked by letter A in Fig. 6.2, through which the strain rate of the alloy during steady state creep was measured to be 0.0102%/h, the lasting time was about 40 h, and the creep lifetime of the alloy was measured to be 67 h. The creep curve of the alloy solution treated at 1160 °C was marked by letter B in Fig. 6.2, indicating that the alloy displays a lower strain rate during steady-state creep, and the creep lifetime of the alloy was measured to be 104 h. When the solution temperature increased to 1165 °C, the creep lifetime of the alloy was measured to be about 9 h, as marked by letter C in Fig. 6.2.

After solution treated at different temperatures, the creep lifetimes of the alloy under the applied various stresses and temperatures are measured as listed in Table 6.1. It may be understood from Table 6.1 that, under the applied stress of 1034 MPa at 650 °C, the lifetime of the alloy solution treated at 1150 °C was measured to be 67 h. As the solution temperature enhanced to 1160 °C, the lifetime of the alloy increased to 104 h. When the solution temperature enhanced further to 1165 °C, the lifetime of the alloy decreased rapidly to 9 h, as shown in Table 6.1. This indicates that the solution temperature has an obvious influence on the creep lifetimes of the alloy.

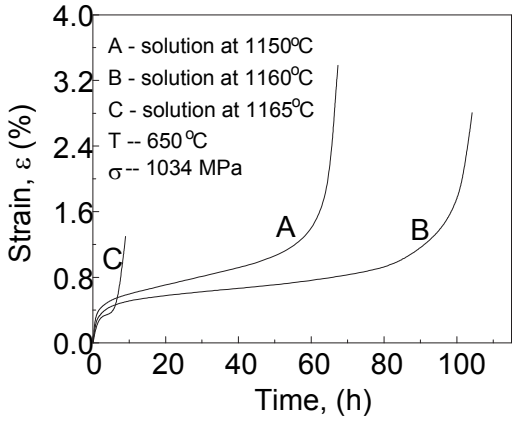


Fig. 6.2. Under the applied stress of 1034 MPa at 650 °C, creep curves of the alloy solution treated at different temperatures

After solution treated at 1150 °C, the lifetime of the alloy under the applied stress of 1020 MPa at 650 °C was about 95 h, and the strain rate during steady state creep was measured to be about 0.00827%/h. The lifetime of the alloy decreased to 67 h as the applied stress increased to 1034 MPa at 650 °C, and the creep lifetime of the alloy decreased to 37 h as the applied stress increased to 1050 MPa at 650 °C. And the alloy solution treated at 1160 °C displays also the similar regularity under the experimental conditions. This indicates that the alloy has an obvious sensitive to the applied stresses and temperatures.

Solution Temp. (°C)	650 °C									1034 MPa					
	1020 MPa			1034 MPa			1050 MPa			660 °C			670 °C		
	t_f (h)	δ (%)	$\dot{\varepsilon}$ (%/h)	t_f (h)	δ (%)	$\dot{\varepsilon}$ (%/h)	t_f (h)	δ (%)	$\dot{\varepsilon}$ (%/h)	t_f (h)	δ (%)	$\dot{\varepsilon}$ (%/h)	t_f (h)	δ (%)	$\dot{\varepsilon}$ (%/h)
1150	95	4.2	0.00827	67	3.4	0.0102	37	3.2	0.0129	26	3.6	0.0208	12	2.3	0.0418
1160	156	3.7	0.00307	104	2.8	0.00367	51	2.1	0.00456	40	2.2	0.00825	17	1.7	0.0179
1165	–	–	–	9	1.3	0.00304	–	–	–	–	–	–	–	–	–

Table 6.1. Effect of solution temperature on stress rupture properties of FGH95 alloy

6.1.2 Influence of quenching methods on creep properties

After the alloy was solution treated at 1150 °C and cooled in oil bath at 120 °C, the creep curves of the alloy at different conditions were measured as shown in Fig. 6.3. Under the applied stress of 1034 MPa at different temperatures, the creep curves of the alloy were shown in Fig. 6.3(a), it indicated that the alloy possessed the lower strain rate and longer creep lifetime of about 380 h at 630 °C.

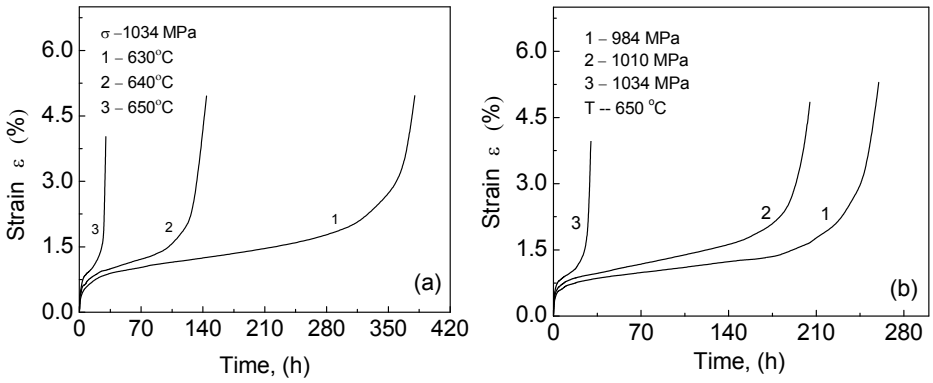


Fig. 6.3. Creep curves of the alloy cooled in oil bath at different conditions. (a) Creep curves of the alloy under the applied stress of 1034 MPa at various temperatures, (b) creep curves of the alloy under the applied different stresses at 650 °C

As the creep temperature increased to 640 °C, the lifetime of the alloy decreased to 144 h. When the creep temperature increased to 650 °C, the lifetime of the alloy decreased rapidly to 32 h. This indicates that the alloy had an obvious sensitive to the applied temperatures. While the creep curves of the alloy under the applied different stresses at 650 °C were measured as shown in Fig. 6.3(b). Under the applied stress of 984 MPa, the alloy displayed a shorter initial creep stage and a longer steady state stage of creep, and the creep lifetime of the alloy was measured to be about 260 h. The lifetime of the alloy decreased to 205 h as the applied stress enhanced to 1010 MPa, but the creep lifetime of the alloy at 650°C/1034 MPa deceased rapidly to 32 h. This indicates that the alloy possessed an obvious sensitive to the applied stress when the applied stress was over 1010 MPa.

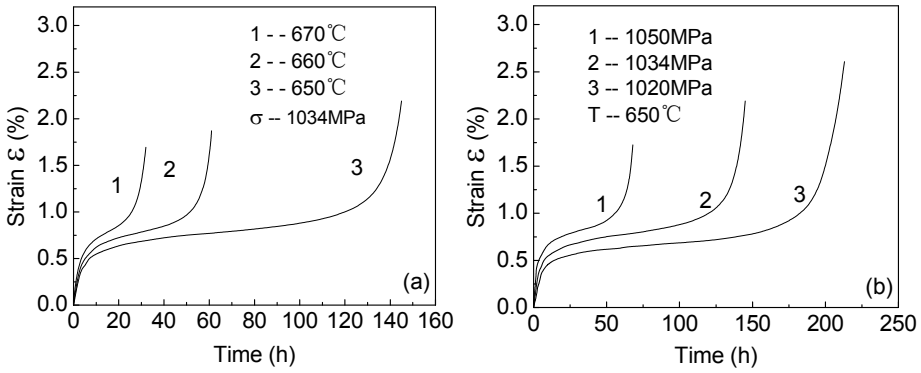


Fig. 6.4. Creep curves of the alloy cooled in molten salt at 583 °C at different conditions. (a) Creep curves of alloy under applied stress of 1034 MPa at various temperatures, (b) creep curves of alloy under the applied different stresses at 650 °C

After the solution treated at 1160 °C and cooled in molten salt at 583 °C, the creep curves of the alloy at different conditions were shown in Fig. 6.4. Under the applied stress of 1034 MPa at different temperatures, the creep curves of the alloy were measured as shown in the Fig. 6.4(a), indicating that the alloy displayed a lower strain rate and creep lifetime of about 145 h at 650 °C. As the temperature enhanced to 660 °C and 670 °C, the strain rates of the alloy during the steady state creep increased, and the creep lifetimes decreased to 60 h and 32 h, respectively. This indicates that the alloy has an obvious sensitive to the applied temperatures.

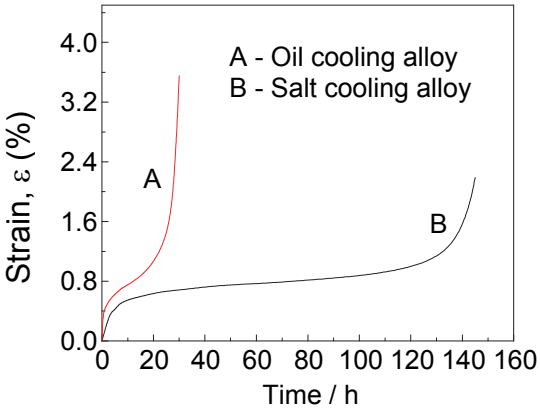


Fig. 6.5. Under the applied stress of 1034 MPa at 650 °C, creep curves of the alloy treated by different regimes

The creep curves of the alloy under the applied different stresses at 650 °C were shown in Fig. 6.4(b). The creep feature of the alloy under the applied stress of 1020 MPa displayed a shorter initial stage and longer steady state stage, and the creep lifetime of the alloy was measured to be 213 h. As the applied stress enhanced to 1034 MPa and 1050 MPa, the creep lifetimes of the alloy were measured to be 145 h and 68 h, respectively. This displays only a smaller decreasing extent of the creep lifetimes, therefore, it may be concluded that no obvious sensitivity is displayed in the alloy in the ranges of the applied stresses.

Cooling medium	Solution Temp. (°C)	Creep lifetime, t_f (h)	Elongation (%)
Oil bath	1150	32	3.9
	1160	121	2.6
Molten salt	1150	67	3.0
	1160	145	2.5

Table 6.2. Creep properties of FGH95 alloy treated by different regimes at 1034 MPa /650 °C

Under the applied stress of 1034 MPa at 650 °C, the creep curves of the alloy cooled in different mediums were shown in Fig. 6.5, indicating that, compared to the oil cooled alloy,

the alloy cooled in molten salt displayed a lower stain rate and longer creep lifetime. Moreover, the creep properties of the alloy treated by different regimes under different conditions are measured as listed in Table. 6.2, this indicates that the alloy solution treated at high temperature possesses the better creep properties.

6.2 Creep equation and relative parameters

Transient strain of the alloy occurs when the loading is applied at high temperature, and the strain rate of the alloy decreases as the creep goes on. The strain rate keeps constant once the creep enters the steady state stage, therefore, the strain rate of the alloy during steady state creep may be expressed by Dorn's law as follows:

$$\dot{\varepsilon}_{ss} = A\sigma_A^n \exp(-\frac{Q_a}{RT}) \tag{6.1}$$

where $\dot{\varepsilon}_{ss}$ is the strain rate during steady state creep, A is a constant related to the microstructure, σ_A is the applied stress, n is the apparent stress exponent, R is the gas constant, T is the absolute temperature, and Q_a is the apparent creep activation energy.

6.2.1 Influence of solution temperatures on creep relative parameters

After the 1120 °C HIP alloy solution treated at various temperatures, the creep curves of the alloy were measured in the ranges of 650 °C ~ 670 °C and 1020 MPa ~ 1050 MPa. The dependence of the strain rates of the alloy during steady-state creep on the applied temperatures and stresses was shown in Fig. 6.6. Figure 6.6(a) showed the relationship between the strain rates and the temperatures under the applied stress of 1034 MPa, and the dependence of the strain rates on the applied stresses at 650 °C was shown in Fig. 6.6(b). According to the data during the steady state creep, the creep activation energies and stress exponents of the alloys, which were solution treated at 1150 °C and 1160 °C, were calculated to be Q_A= 510.1 ± 20 kJ/mol, Q_B = 580.3 ± 20 kJ/mol and n_A=15.4, n_B=14.1, respectively.

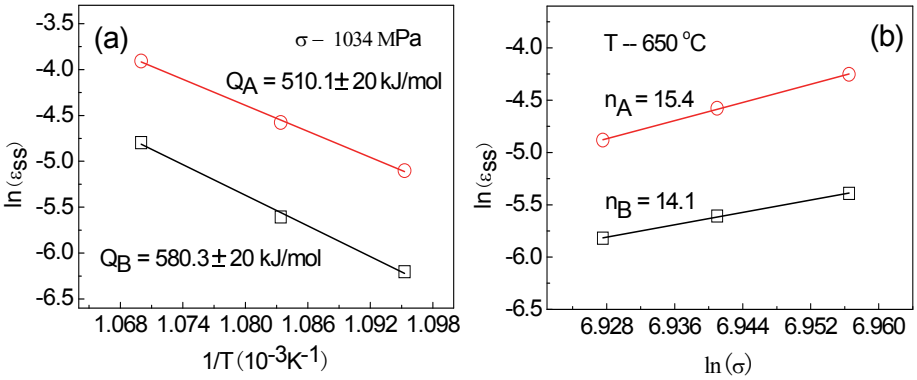


Fig. 6.6. Dependence of the strain rates during steady state creep on the applied temperatures and stresses for the alloy solution treated by different temperatures. (a) Strain rate vs. temperatures at 1034 MPa, (b) strain rate vs. the applied stress at 650 °C

6.2.2 Influence of quenching methods on creep relative parameters

After the 1180 °C HIP alloy was solution treated and cooled in different mediums, the dependences of the strain rates during steady state creep on the applied temperatures and stresses in the ranges of 650 °C ~ 670 °C and 1020 MPa ~ 1050 MPa are shown in Fig. 6.7. The relationship between the strain rates and the temperatures under applied stress of 1034 MPa was shown in Fig. 6.7(a), according to the data, the creep activation energies of the alloy treated by different regimes were calculated to be Q_A= 381.1 ± 20 kJ/mol and Q_B = 590.3 ± 20 kJ/mol, respectively. The dependence of the strain rates on the applied stresses at 650 °C was shown in Fig. 6.7(b), according to the data, the stress exponents of the alloy treated by different regimes were calculated to be n_A = 17.9 and n_B = 13.8, respectively. The values of the creep activation energies and stress exponents were measured as listed in Table 6.3. It may be concluded according to the data in Table 6.3 that the alloy cooled in molten salt displayed a better creep resistance.

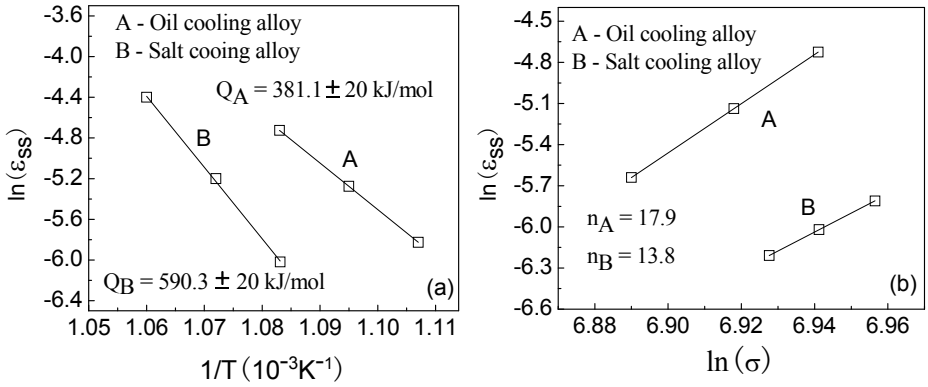


Fig. 6.7. Relationship between the strain rates and the applied temperatures, stresses for the alloy treated by different regimes. (a) Strain rates & temperatures, (b) strain rates & the applied stresses

Cooling medium	Activation energies (kJ/mol)	Stress exponents (n)
Oil bath	381.1 ± 20	17.9
Molten salt	590.3 ± 20	13.8

Table 6.3. Activation energies and stress exponents during the steady-state creep of the alloy treated by different regimes

7. Deformation mechanisms of the alloy during creep

7.1 Deformation features of alloy during creep

After the 1120 °C HIP alloy solution was treated at 1160 °C, the morphology of the alloy crept for 104 h up to rupture under the applied stress of 1034 MPa at 650 °C was shown in Fig. 7.1. Figure 7.1(a) displayed the feature of the stacking fault as marked by white arrow, and two (1/3)<112> super-Shockleys partials were located on the two sides of the stacking fault. Some dislocation loops with various sizes appeared clearly in the γ matrix of the alloy,

the smaller dislocation loop was marked with black arrow in Fig. 7.1(a), and larger dislocation loop was located in the region above the stacking fault. It could be deduced by analysis that the <110> super-dislocation shearing into the γ or γ′-phase may be decomposed to form the configuration of (1/6)<112> shockleys dislocations or (1/3)<112> super-Shockleys partials plus the stacking fault.

Some blocky carbide particles were precipitated in the local area of the alloy as marked by black arrow in Fig. 7.1(b). The grain boundary in the alloy was marked by white arrow in Fig. 7.1(c), and dislocations tangles were piled up in the region B near the boundary, which suggested that the boundaries may effectively hinder the dislocation movement during creep. Besides, the morphology of <110> super-dislocation shearing into the γ′-phase was marked by the black arrow, and the stacking fault formed from the dislocation decomposition was marked by the letter A in Fig. 7.1(c). The blocky carbides were precipitated along the boundary as marked by the black arrow in Fig. 7.1(d), and some dislocations were piled up in the regions near the carbides, which indicated that the carbide particles can effectively hinder the dislocation movement. But, the facts that significant amount of dislocations were piled up in the regions near the boundaries can cause the stress concentration to promote the initiation and propagation of the micro-crack along the boundaries as the creep goes on.

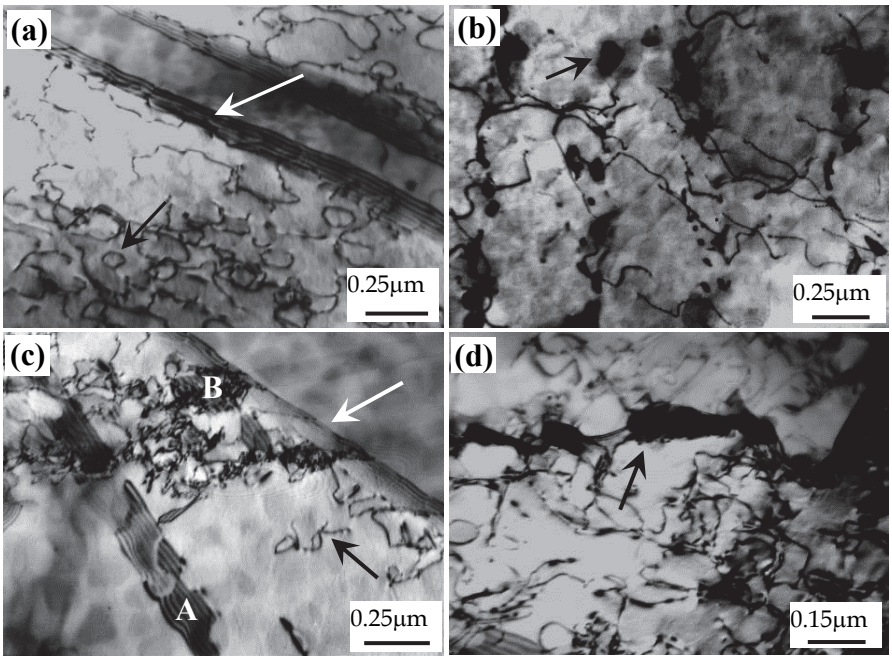


Fig. 7.1. Microstructure of the alloy after (solution treated at 1160 °C) crept for 104 h up to fracture at 650 °C/1034 MPa. (a) dislocation loops in the alloy, (b) carbide particles precipitated within the grain, (c) morphology of the super-dislocation shearing into γ′ phase as marked by black arrow, the stacking fault in the region A and dislocation tangles piled up in the B region near the boundary, (d) carbide particles precipitated along the boundary, and dislocation tangles piled up in the regions near the boundaries

After the 1180 °C HIP alloy solution treated at 1150 °C and cooled in the oil bath, the microstructures of the alloy crept for 260 h up to fracture under the applied stress of 984 MPa at 650 °C was shown in Fig. 7.2. The fine sphere-like γ′ particles about 0.1~0.2 μm in size were dispersedly distributed in the alloy as shown in Fig. 7.2(a), which may hinder the dislocation movement for improving the creep resistance of the alloy. During creep, the double orientation slipping of dislocations were activated in the local regions of γ matrix as shown in Fig. 7.2(b), the deformed dislocations in the form of the tangles were arranged along the level direction, the bundle-like dislocations were arranged along the upright-like direction as marked by arrows, respectively, in Fig. 7.2(b). During creep, the slipping dislocations moving to the boundary were stopped as marked by white arrow in Fig. 7.2(c), indicating that the boundary had an obvious effect on hindering the dislocations movement. Therefore, it may be concluded that the deformation feature of the alloy during creep is the double orientations slipping of dislocations activated in the γ matrix phase.

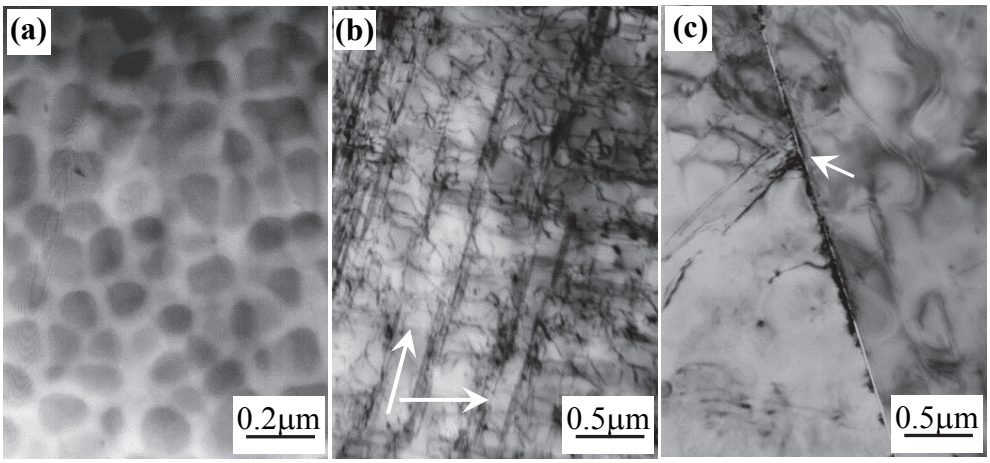


Fig. 7.2. After solution treated at 1150 °C and cooled in oil bath, microstructure of the alloy crept for 260 h up to fracture under the applied stress of 980 MPa at 650 °C. (a) Fine γ′ phase precipitated dispersedly within the grains, (b) double orientations slipping of dislocation as marked by arrows, (c) dislocations slipping stopped on the boundary

After the 1180 °C HIP alloy was solution treated at 1160 °C and cooled in the molten salt, the microstructures of the alloy crept for 72 h up to fracture under the applied stress of 1034 MPa at 650 °C was shown in Fig. 7.3. The carbide particles were precipitated along the boundaries as marked by arrow in Fig. 7.3(a), which resulted in the boundary with the uneven feature. The dislocation-free region appeared in the left side of the boundary, significant amount of dislocations were congregated in the another side of the boundaries, and the carbide particle was precipitated along the boundary as marked by the arrow in Fig. 7.3(b). It may be analyzed according to Fig. 7.3(b) that the boundaries and the carbides precipitated along the boundaries may hinder the dislocations movement for improving the creep resistance of the alloy, so that significant amount of dislocations are piled up in the right-side region of the boundary as shown in Fig. 7.3(b).

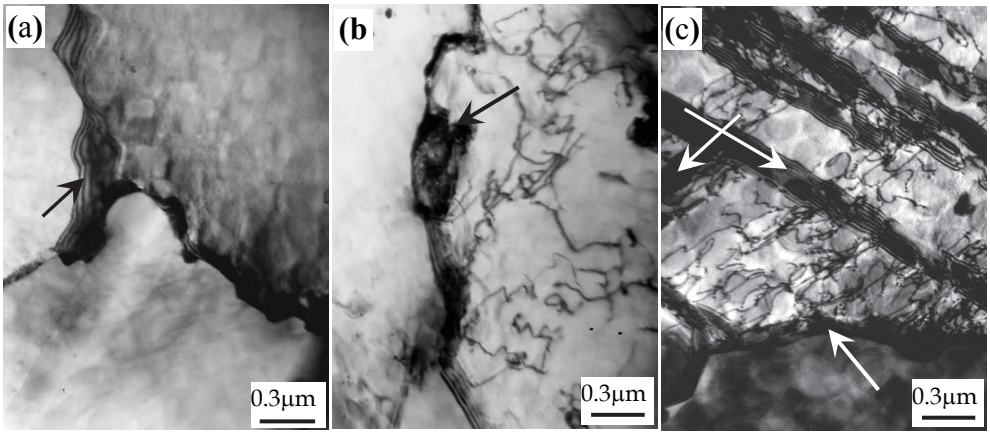


Fig. 7.3. After solution treated at 1160 °C and cooled in molten salt, microstructure of the alloy crept up to fracture at 1034 MPa / 650 °C. (a) Carbide particles precipitated along the boundary, (b) dislocation tangles piled up in the regions near the boundaries, (c) stacking fault and dislocations appeared within the grain

In the local region of the alloy, the double orientation slipping of dislocations were activated within the grain, the boundary was marked by arrow in the bottom of Fig. 7.3(c), the directions of the double orientation slipping of dislocations were marked with arrows in Fig. 7.3(c). The arrow which points to left side corresponds to the orientation of the dislocations tangle, the arrow which points to right-side corresponds to the direction of the stacking fault stripes. It may be deduced by analysis that the stacking fault is formed in between two (1/3)<112> super-Shockleys partials which originates from the decomposition of <110> super-dislocation. The configuration of the partial dislocation + stacking fault can restrain the cross-slipping of the dislocations to improve the creep resistance of the alloy.

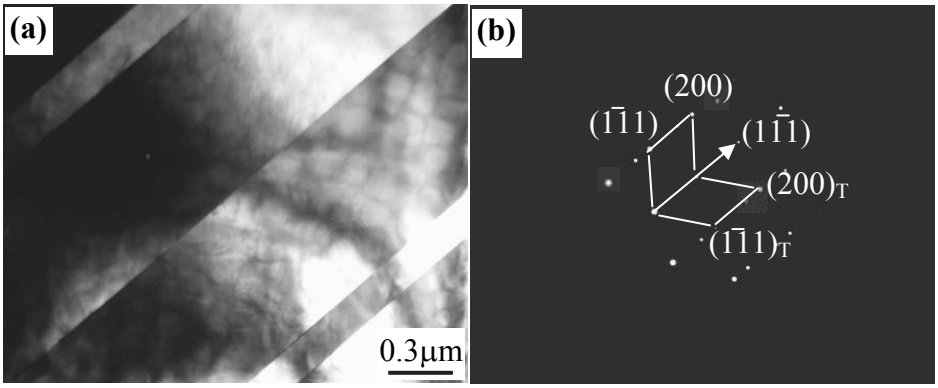


Fig. 7.4. Twinning in the alloy during creep at 650 °C/1034 MPa. (a) twinning, (b) SAD patterns

After the alloy was crept up to fracture under the applied stress of 1034 MPa at 650 °C, the feature of the twinning deformation was shown in Fig. 7.4(a), and the twinning plane was

identified as (111) plane by means of the diffraction pattern analysis as shown in Fig. 7.4(b). The blocky carbide was precipitated along the grain boundary as marked by shorter arrow in Fig. 7.5(a), and some micro-twinning was detected in the alloy as marked by black arrow. In another local area, the morphology and diffraction spots of the micro-twinning were shown in Fig. 7.5(b) and (c), this is well agreed with the result in literature (Unocic R. R. et al., 2008). The one end of the micro-twinning was stopped at the grain boundary, which suggests that the grain boundaries had an obvious effect on hindering the twinning deformation.

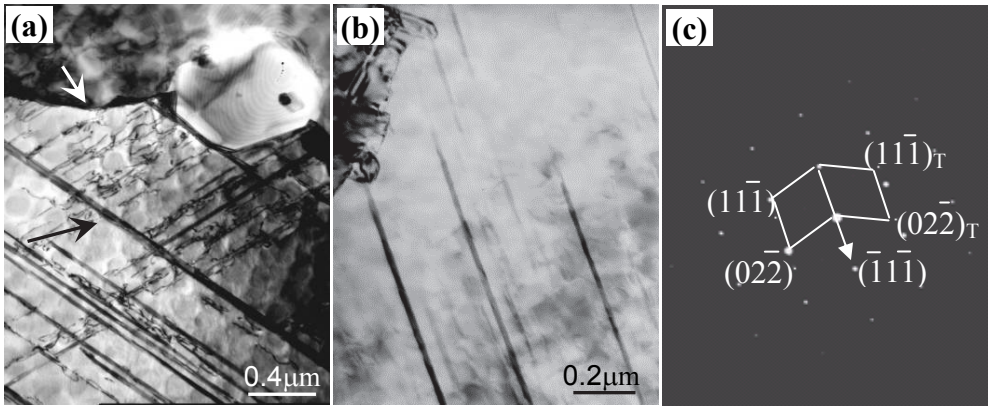


Fig. 7.5. Morphologies of the carbide precipitated along the boundary and micro-twinning formed in the alloy during creep under the applied stress of 1034 MPa at 650 °C. (a) Carbide precipitated along the boundary, (b) micro-twinning, (c) SAD patterns

7.2 Analysis on deformation features of the alloy during creep

7.2.1 Dislocation model

Significant amount of the fine γ′ particles are precipitated within the grains, which may effectively hinder the dislocation movement. When the deformed dislocations move over the γ′ phase during creep, the dislocation loops are kept around the γ′ particles as shown in Fig. 7.1(a), which suggests that the deformation feature of the alloy during creep is the dislocations moving over the γ′-phase by Orowan bypassing mechanism. It is reasonable consideration that the various spaces between the γ′ particles appear in different regions, the dislocations may bow out along the wider channels between the two γ′ particles during creep, and the bowing dislocations move over the γ′-phase by Orowan mechanism to encounter for forming the dislocation loops, as marked with black arrow in Fig. 7.1(a). When the dislocations bow out along the channels during creep, the applied stress which is enough to overcome the Orowan resistance can be expressed as follows:

$$\tau_{or} = \frac{\mu \cdot b}{L} \tag{7.1}$$

Where, μ is the shear modulus, b is the Burgers vector, and L is the space between two γ′ particles. This indicates that the resistance of the dislocation movement increases with the diminishing of the space between the γ′-particles, and the resistance of alloy enhances with the volume fraction of γ′-particles.

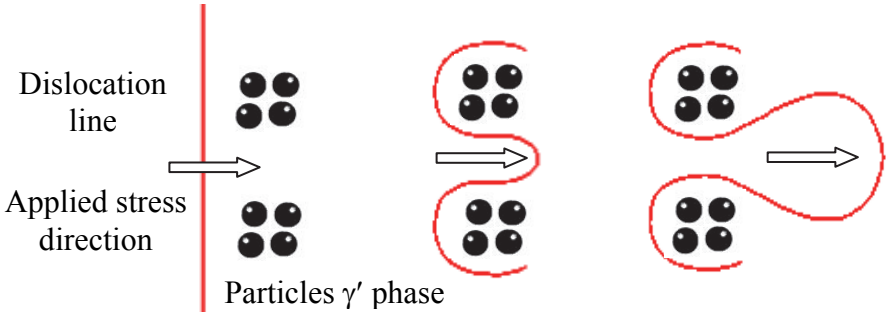


Fig. 7.6. Schematic diagram of dislocation bowing out in the congregated regions of the γ′ particles

The space between the γ′ particles diminishes when some γ′ particles are congregated together, the smaller space is measured to be L_1 = 0.03 μm as shown in Fig. 4.5(a), but the bigger space between the γ′-particles is measured to be L_2 = 0.12 μm. Compared to the bigger space, the dislocations bowing out in the regions with smaller space needs fourfold Orowan resistance according to the formula (7.1), therefore, it is difficult for the dislocations to bow out along the channels with smaller space, but they can bow out along the channels with larger space to form the larger loops as shown in Fig. 7.1(a). When the dislocations bow out along the channels with bigger space (L), the formation process of the bigger dislocation loops is schematically shown in Fig. 7.6.

In the later stage of creep, significant amount of dislocations move to the region near the γ′ particles to generate stress concentration, and when the stress value originated from the stress concentration exceeds the yield strength of γ′ phase, the <110> super-dislocation may shear into the γ′-phase, as marked by black arrow in Fig. 7.1(c). Furthermore, the <110> super-dislocations may be decomposed to form the configuration of two (1/3)<112> super-Shockleys partials and the stacking fault, as marked by letter A in Fig. 7.1(c). The critical stress of dislocation shearing into γ′ phase increases with the yield strength. And the critical stress (τ_{cs}) can be expressed as follows (Zhang J. SH., 2007) :

$$\tau_{cs} = \frac{\eta_{APB}}{b}\left(\frac{0.3 \cdot \eta_{APB} \cdot f \cdot r}{T}\right)^{1/2} \tag{7.2}$$

Where T is the dislocation line tension, r is the radius of the γ′-particle, $\boldsymbol{b}$ is the Burgers vector, f is the volume fraction of γ′ phase, and η_{APB} is the antiphase boundary energy originated from the dislocation shearing γ′ phase. It may be understood from equation (7.2) that the critical stress of dislocation shearing into γ′ phase increases with the size (r), volume fraction (f) of γ′-phase and antiphase boundary energy (η_{APB}) to improve the creep resistance of the alloy.

7.2.2 Micro-twinning model

The observed shearing process spawned the model (Unocic R. R. et al., 2008) is illustrated in Fig. 7.7. The process of shearing the larger γ′ particles is presumed to occur by the cooperative movement of the coupled (a/2)[112] dislocation (3Dα), each of which is dissociated into three (a/3)[112] partials (Dα).

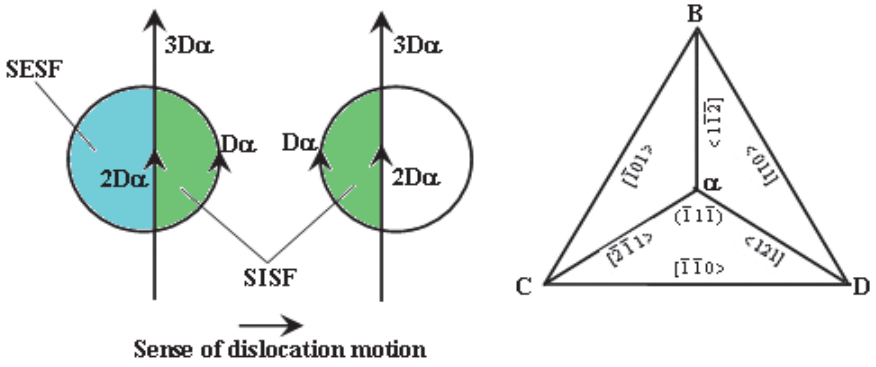


Fig. 7.7. Model of γ′ precipitate shearing by coupled Shockley partials for creating SISF/SESF pairs. After Hirth and Lothe (Unocic R. R. et al., 2008, as cited in Hirth J. P. & Lothe J., 1968)

The a<112> dislocations are hypothesized to originate from the interaction of two different a<101> super- dislocations originating from different slip systems. For example:

$$a[011] + a[101] = a[112] \tag{7.3}$$

Clearly, this model then requires a high symmetry orientation such that two slip systems experience a relatively large shear stress.

In situ deformation at higher temperature gives rise to a distinctly different mode of shearing in which the extended faults propagate continuously and viscously through both particles and matrix. These extended faults are associated with partials that move in a correlated manner as pairs. Koble (Koble M., 2001) induced that these partials may be a/6<112> partials of the same Burgers vector, and that they may be traveling in parallel {111} planes, as illustrated in Fig. 7.8. Without detailed confirmation of this hypothesis, Kolbe further deduced that these were in fact micro-twins, and that the temperature dependence of the process may be associated with recording that would ensure in the wake of twinning a/6<112> partials as they traverse the γ′ particles. The shear strain rate can be expressed as follow:

$$\dot{\gamma} = \rho_{tp} b_{tp} v = \rho_{tp} b_{tp} \frac{D_{ord} (b_{tp} / x^2)}{\ln[f_2 \Delta\Gamma / (2\tau_{eff} b_{tp} - f_2 \Gamma_{tt})}} \tag{7.4}$$

Where, Γ_{pt} is the energy of two layered pseudo-twin, and b_{pt} is the magnitude of the Burgers vector of the twinning partials, Γ_{tt} is the energy of two layered true twin, ρ_{pt} is the density of mobile twinning partials, D_{ord} is the diffusion coefficient for ordering, x is the short range diffusion length (assumed to be several nearest neighbor distances, or ~2b), f_2 is the volume fraction of the secondary γ′ precipitates, f_3 is the volume fraction of the tertiary γ′ precipitates. And the effective stress (τ_{eff}) , in the presence of tertiary γ′ precipitates, is given by:

$$\tau_{eff} = \tau - \frac{f_3 \Gamma_{pt}}{2 b_{tp}} \tag{7.5}$$

The experimental values of parameters such as dislocation density ρ_{pt}, volume fraction of the secondary γ′ precipitate that are critical to the prediction can be determined directly from TEM observations. Disk alloys in this temperature regime typically exhibit the creep curves having a minimum rate, with a prolonged increase of creep rate with time. As the fine γ′ phase volume fraction decreases during thermal exposure, it is possible that the operation of 1/2[110] matrix dislocations becomes increasingly important. The coarse microstructure (small value of f_3) resulting from a slow cooling rate, the deformation is dominated by 1/2<110> dislocation activated in the matrix, and SESF shearing in the secondary γ′ precipitates.

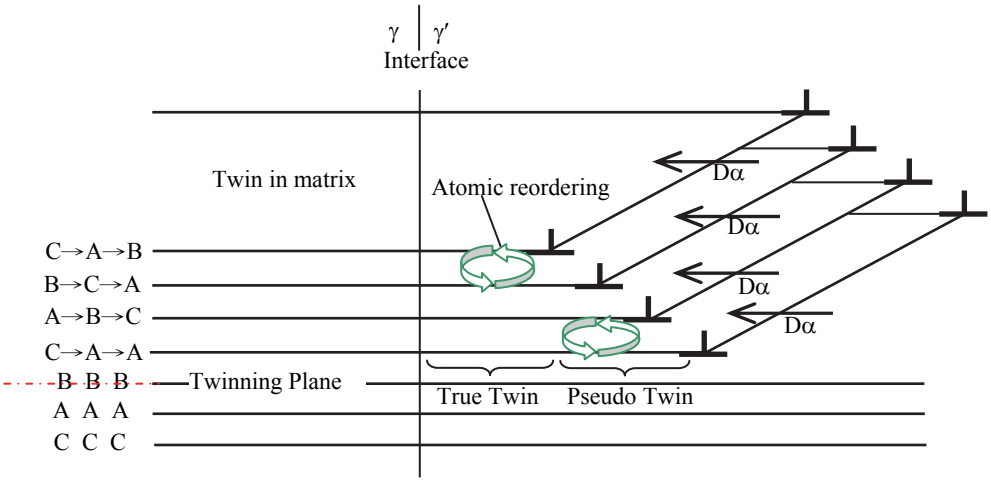


Fig. 7.8. Schematic representation of micro-twinning mechanism from shear by identical Shockley partials (Dα) transcending both the γ matrix and γ′ precipitate in adjacent {111} planes which then require atomic reordering in γ′ to convert stacks of CSF into a true twinned structure. After Kolbe (Koble M., 2001)

8. Fracture features of the alloy during creep

8.1 Influence of solution temperature on fracture feature of alloy during creep

After the 1120 °C HIP alloy was solution treated at 1150 °C and isothermal quenched in molten salt at 583 °C, the morphology of the alloy crept for different time under the applied stress of 1034 MPa at 650 °C was shown in Fig. 8.1. The applied stress direction was marked with the arrow in Fig. 8.1(a), after the alloy was crept for 40 h, some slipping traces appeared on the surface of the sample, and some parallel slipping traces were displayed within the same grain. Moreover, the various orientations of the slipping trace appeared within the different grains. Besides, the kinking of the slipping traces appeared in the region of the boundaries as marked by arrow in Fig. 8.1(a). After crept for 67 h up to rupture, the surface morphology of the alloy was shown in Fig. 8.1(b), indicating that the amount of the slipping trace increased as the creep went on, and the slipping traces were deepened to form the slipping steps on the surface of the specimen. Moreover, the bended slipping traces appeared in the boundary regions, as marked by longer arrow in Fig. 8.1(b), which was

attributed to the effect of the flow metal in the γ′-free phase zone where is lower in strength. Besides, the cracks were initiated in the distortion regions of slipping traces as marked by shorter arrow in Fig. 8.1(b).

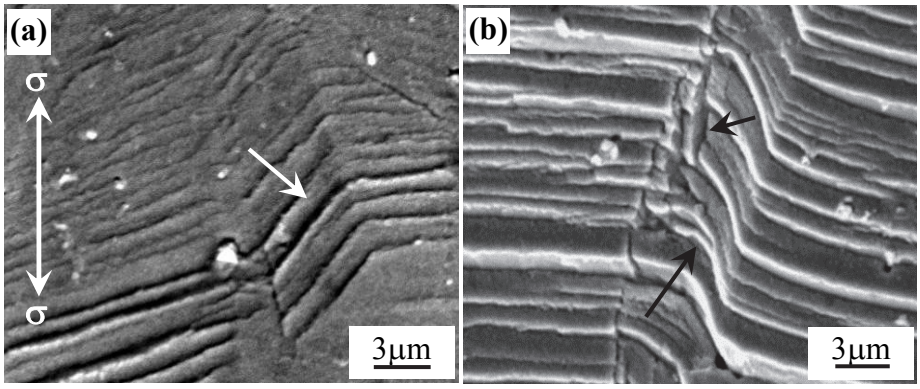


Fig. 8.1. Surface morphology of the alloy crept for different time up to fracture. (a) After crept for 40 h, a few slipping traces appeared within the different grains, (b) after crept up to fracture, significant amount of the slipping traces appeared on the sample surface, and cracks appeared in the region near the boundary as marked by arrow

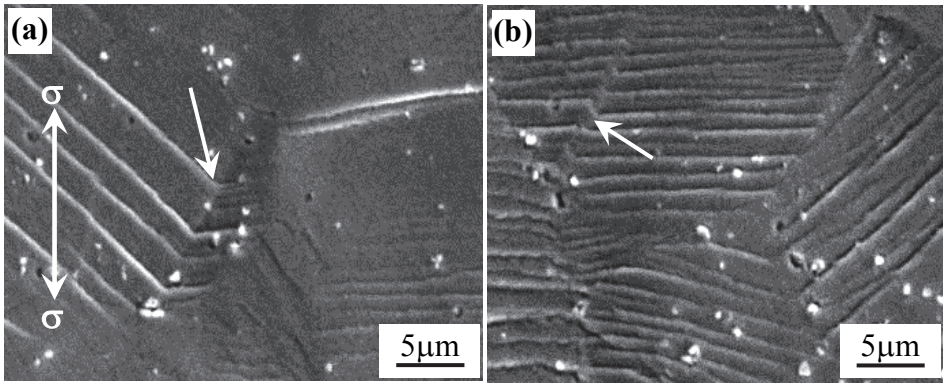


Fig. 8.2. After solution treated at 1160 °C, surface morphology of the alloy crept for different time. (a) After crept for 60 h, a few slipping traces appeared within the different grains, (b) after crept for 80 h, significant amount of the slipping traces appeared in the surface of the sample

After 1120 °C HIP alloy was solution treated at 1160 °C and twice aged, the morphology of the alloy crept for different time under the applied stress of 1034 MPa at 650 °C was shown in Fig. 8.2. The direction of the applied stress was marked by arrows, after the alloy was crept for 60 h, the morphology of the slipping traces on the sample surface was shown in Fig. 8.2(a), which displayed the feature of the single orientation slipping appearing within the different grains. And the intersected of the slipping traces appeared in the boundary region as marked by arrow in Fig. 8.2(a), which indicated that the boundary may hinder the

slipping of the traces to change their direction. When crept for 80 h, the quantities of the slipping traces on the sample surface increased obviously, as shown in Fig. 8.2(b), and some white blocky carbide particles were precipitated within the grains.

After solution treated at 1160 °C and twice aged, the surface morphology of the alloy crept up to rupture under the applied stress of 1034 MPa at 650 °C was shown in Fig. 8.3. As the creep went on, the quantities of the slipping traces increases gradually (the direction of the applied stress shown in Fig. 8.3(a), which may bring out the stress concentration to promote the initiation of the micro-cracks along the boundary which was vertical to the stress axis as marked by the letter A and B in Fig. 8.3(a). In the other located region, the morphology of the crack initiation was marked by letter C in Fig. 8.3(b), the micro-cracks displayed the non-smooth surface as marked by arrow, and the white carbide particle was located in the crack, it indicated that the carbide particles precipitated along the boundary may restrain the cracks propagating along the boundaries to enhance the creep resistance of the alloy.

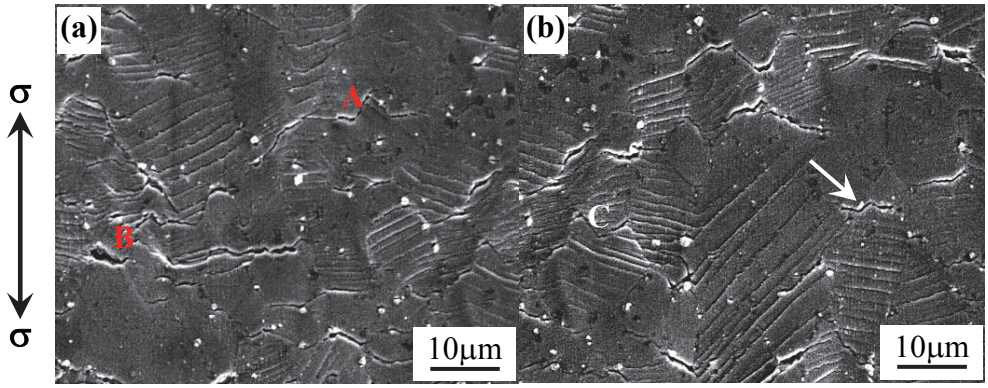


Fig. 8.3. Cracks initiated and propagated along the boundary. (a) Crack initiated along the boundaries vertical to the stress axis, (b) crack propagated along the boundaries as marked by arrow

After the alloy crept up to fracture, the morphology of the sample polished and eroded was shown in Fig. 8.4. Some carbide particles were located in the boundaries as shown in Fig. 8.4(a), which may hinder the slipping of the dislocation for enhancing the creep resistance of the alloy. Moreover, the unsmooth surface of the cracks appeared in the fracture regions as marked by white arrow in Fig. 8.4(a). However, when no carbide particles were precipitated along the boundaries, the crack after the alloy crept up to fracture displayed the smooth surface as marked with the letter D and E in Fig. 8.4(b).

It may be thought by analysis that, although the carbide particles may hinder the dislocations movement for improving the creep resistance of alloy, the carbides located in the regions near the boundaries may bring about the stress concentration to promote the initiation and propagation of the cracks along the boundary as marked with the arrow in Fig. 8.4(a). Therefore, the fracture displayed the non-smooth surface due to the pinning effect of the carbide particles precipitated along the boundaries to restrain the boundaries slipping during creep. Though the carbide particles precipitated along the boundaries can improve the cohesive strength of the boundaries, the micro-cracks are still initiated and propagated along the boundaries, which suggests that the boundaries are still the weaker regions for causing fracture of the alloy during creep.

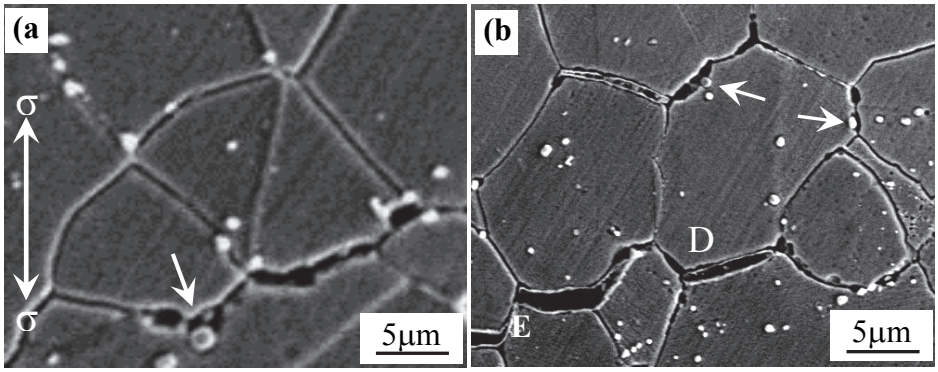


Fig. 8.4. After solution treated at 1160 °C, surface morphology of the alloy crept up to fracture. (a) Carbide particles near the crack along the boundary marked by arrow, (b) morphology of cracks propagated along the boundary marked by arrow

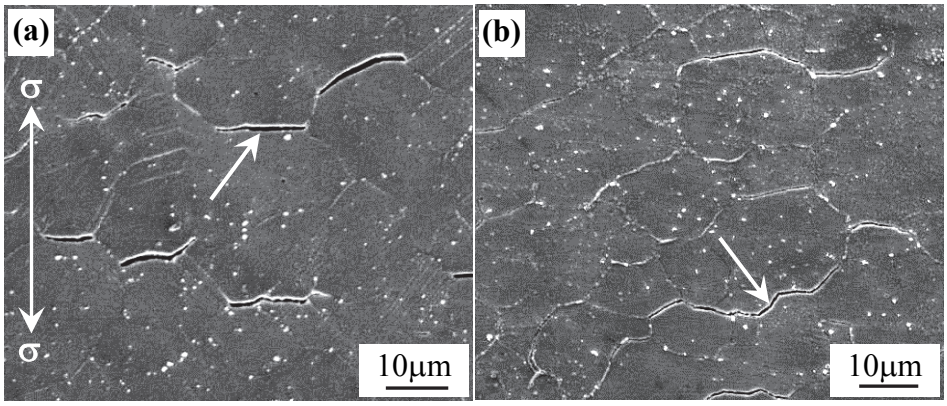


Fig. 8.5. After solution treated at 1165 °C, surface morphology of the alloy crept for 9 h up to fracture. (a) Crack initiated along the boundary as marked by arrow, (b) cracks propagated along the boundary as marked by arrow

After solution treated at 1165 °C and aged, the surface morphology of the alloy crept for 9 h up to rupture under the applied stress of 1034 MPa at 650 °C was shown in Fig. 8.5. A few slipping trace appeared only on the surface of the alloy, and some micro-cracks were initiated along the boundaries vertical to the applied stress axis, as marked by arrow in the Fig. 8.5(a). As the creep went on, the morphology of the micro-crack propagated along the boundary was shown in Fig. 8.5(b), in which the fracture of the alloy displayed the smooth surface. It may be deduced according to the feature of the smooth fracture that the carbide films precipitated along the boundaries has an important effect on decreasing the stress fracture properties of the alloy. The carbide films were formed along the boundaries during heat treated, which reduced the cohesive strength between the grains. Therefore, the micro-crack was firstly initiated along the boundaries with the carbide films, and propagated along the interface between the carbide films and grains, which resulted in the formation of the smooth surface on the fracture, and decreased to a great extent the creep properties of the alloy.

After the alloy was crept for 9 h up to rupture under the applied stress of 1034 MPa at 650 °C, the surface morphology after the sample was polished and eroded was shown in Fig. 8.6. The carbide films were continuously formed along the boundaries as marked with the long arrow in Fig. 8.6(a), the direction of the applied stress was marked by arrow, the micro-crack was initiated along the carbide film, as marked by shorter arrow in Fig. 8.6(a). As the creep went on, the morphology of the crack propagated along the boundaries was shown in Fig. 8.6(b), the fracture after the crack was propagated displayed the smooth surface, and the white carbide film was reserved between the tearing grains marked by arrow in Fig. 8.6(b), which displayed an obvious feature of the intergranular fracture of the alloy during creep. It can be thought by analysis that the carbide films precipitated along the boundaries, during heat treated, possessed the hard and brittle features and weakened the cohesive strength between the grains. Therefore, the micro-crack was firstly initiated along the carbide films and propagated along the interface between the grains and carbide films, which resulted in the formation of the smooth surface on the fracture, so the alloy had the lower toughness and shorter creep lifetime. Moreover, it was identified by means of composition analysis under SEM/EDS that the elements Nb, Ti, C and O were rich in the white particles on the surface of the samples, as shown in Fig. 8.2, Fig. 8.3 and Fig. 8.5, respectively, therefore, it is thought that the white particles on the surface of the samples are the oxides of the elements Nb, Ti and C.

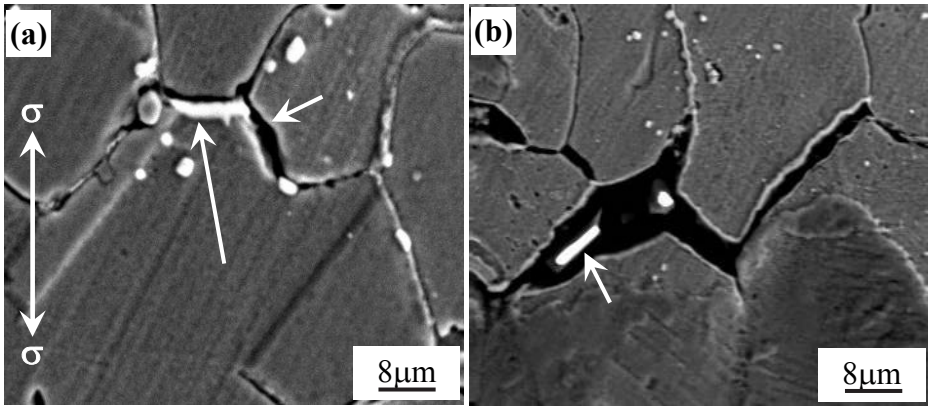


Fig. 8.6. After solution treated at 1165 °C, surface morphology of the alloy crept for 9 h up to fracture. (a) Crack initialed along the boundary marked by arrow, (b) morphology of cracks propagated along the boundary marked by arrow.

8.2 Influence of quenching temperatures on fracture feature of alloy during creep

After the 1180°C HIP alloy was solution treated at 1150 °C and cooled in oil bath at 120 °C, the morphologies of the alloy crept for 260 h up to rupture under the applied stress of 984 MPa at 650 °C were shown in Fig. 8.7. If the PPB region between the powder particles was regard as the grain boundaries as shown in Fig. 8.7(a), the grain boundaries after the alloy was crept up to rupture were still wider, and the ones were twisted into the irregular piece-like shape as marked by arrow in Fig. 8.7(a).

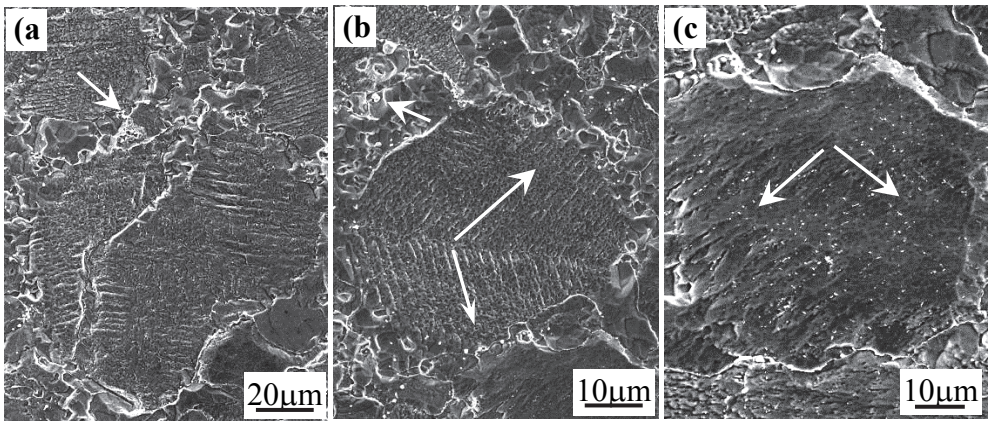


Fig. 8.7. Microstructure of alloy after crept up to fracture under the applied stress of 984 MPa at 650 °C. (a) Wider grain boundaries broken into the irregular shape as marked by arrow, (b) traces with double orientations slipping feature appeared within the grain as marked by arrows, (c) finer particles precipitated along the slipping traces

Some irregular finer grains were formed in the boundary regions, and displaying a bigger difference in the grain sizes. Some coarser γ′ precipitates were precipitated in the boundaries region in which the creep resistance is lower due to the spareness of the finer γ′ phase. The severed deformation of the alloy occurred firstly in the boundary regions during high stress creep, which resulted in the boundaries broken into the irregular piece-like shape. At the same time of the severed deformation, the traces with double orientations slipping feature appeared within the grains as marked by arrows in Fig. 8.7(b), and some particles were precipitated in the boundaries region as marked by short arrow in Fig. 8.7(b). Moreover, the finer white particles were precipitated in the regions of the double orientations slipping traces as marked by arrows in Fig. 8.7(c), and the white particles were distinguished as the carbides containing the elements Nb, Ti and C by means of SEM/EDS composition analysis.

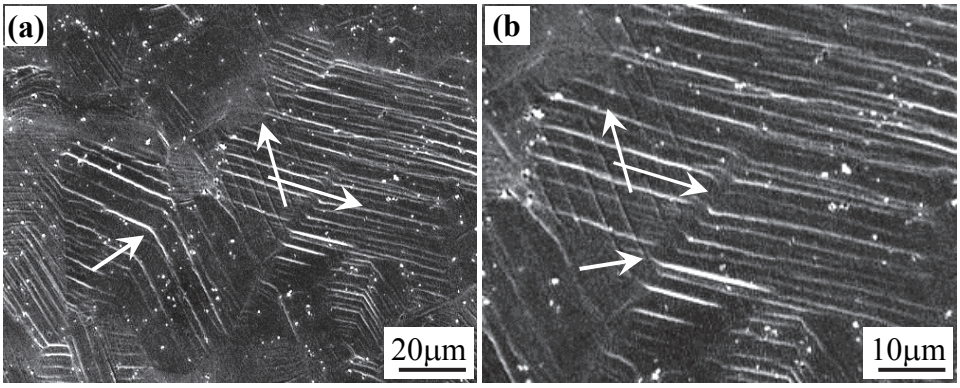


Fig. 8.8. Microstructure after the molten salt cooled alloy crept up to fracture under the applied stress of 1034 MPa at 650 °C. (a) Traces of the double orientations slipping appeared within the grains, (b) magnified morphology of the slipping traces

After solution treated at 1150 °C, and cooled in molten salt at 583 °C, the morphology of the alloy crept for 67 h up to rupture under the applied stress of 1034 MPa at 650 °C was shown in Fig. 8.8. This indicated that the traces with the double orientations slipping feature appeared within the grain, and the various orientations of the slipping traces appeared in the different grains, thereinto, the directions of the thicker and fine traces were marked by the arrows, respectively, in Fig. 8.8(a). Moreover, the traces with the cross-slipping feature were marked by shorter arrow in Fig. 8.8(a).

8.3 Analysis on fracture features during creep

After solution treated at various temperatures, the alloy had different creep properties due to the difference of microstructure as shown in Table 6.2. When solution treated at 1150 °C, the alloy possessed a uniform grain size and wider PPB regions between the grains. Moreover, some coarser γ′ precipitates were distributed along the PPB regions in which no fine γ′-phase was precipitated in the regions near the coarser γ′-phase, as shown in Fig. 4.2(a), the regions possessed a lower creep strength due to the cause of the γ′-free phase zone. After the alloy was solution treated at 1160 °C and twice aged, the coarser γ′ precipitates along the boundary regions disappeared, the boundaries appeared obviously in between the grains. And the cohesive strength between the grains was obviously improved due to the pinning effect of the fine carbide particles, as shown in Fig. 4.3(b), therefore, the alloy displayed a better creep resistance and longer the lifetime.

After the 1120 °C HIP alloy was solution treated at 1160 °C and twice aged, the alloy was crept for 104 h up to fracture under the applied stress of 1034 MPa at 650 °C, the fracture after the alloy was crept up to rupture displayed the initiating and propagating feature of the cuneiform crack as marked by letters A and B in Fig. 8.3. The schematic diagram of the crack initiated along the triangle boundary is shown in Fig. 8.9, where σ_n is the normal stress applied on the boundary, L is the boundary length, h is the displacement of the cuneiform crack opening, α is the crack length, θ is the inclined angle of the adjacent boundaries.

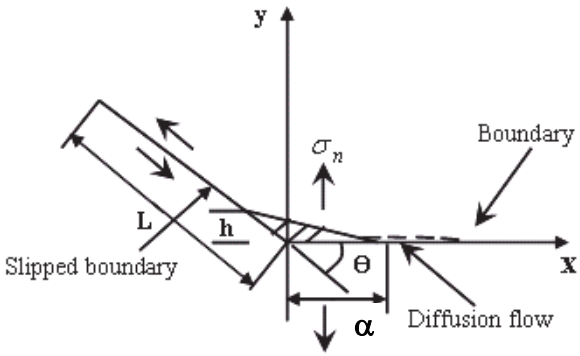


Fig. 8.9. Schematic diagram of the crack initiated along the triangle boundary

Under the action of the applied stress, significant amount of the activated dislocations are piled up the regions near the boundary to bring the stress concentration, which results in the initiation of the crack in the region near the triangle boundary, and the crack is

propagated along the boundary as the creep goes on. Thereinto, the critical length (α_C) of the instable crack propagated along the boundary can be expressed as follows (Yoo M. H., 1983).

$$a_c = \frac{Gh^2}{2\pi(1-\nu)\cdot\gamma_f} \tag{8.1}$$

Where, G is shearing modulus, ν is Poisson ratio, γ_f is the crack propagating work, h is the displacement of the cuneiform crack opening. This indicates that critical length (α_c) of the instable crack propagated along the boundary increases with the displacement of the crack opening, and is inversely to the crack opening work. Thereinto, the displacement of the crack opening increases with the creep time, which can be express as follows:

$$h(t) = \frac{\tau}{\beta}\left[1-\exp(-\frac{4\cdot t\cdot\alpha\cdot\delta_B\cdot\sin\theta}{\pi\eta_B})\right] \tag{8.2}$$

Where $h_w=(\tau/\beta)$ is the max displacement of the crack opening, τ is the resolving shear stress component applied along the boundary, t is the time of the crack propagation, δ_B is the boundary thickness, η_B is the sticking coefficient of the boundary slipping, β is the material constant.

The Eq. (8.2) indicates that the displacement of the crack opening (h) increases with the time and length of crack propagation. When two cuneiform-like cracks on the same boundary are joined each other due to their propagation, the intergranular rupture of the alloy occurs to form the smooth surface on the fracture. The schematic diagram of two cuneiform-like cracks initiated and propagated along the boundary for promoting the occurrence of the intergranular fracture is shown in Fig. 8.10. If the carbide particles are dispersedly precipitated along the boundaries, the ones may restrain the boundaries slipping for improving the creep resistance of the alloy to form the non-smooth surface on the fracture, as marked by arrow in Fig. 8.3(b).

After solution treated at 1165 °C and twice aged, the grain size of the alloy increased obviously, and the carbide films were formed along the boundaries as shown in Fig. 4.4, which weakened the cohesive strength between the grains. Therefore, the cracks were easily initiated and propagated along the boundaries adjoined to the carbide films, which may sharply reduce the lifetime and plasticity of the alloy during creep.

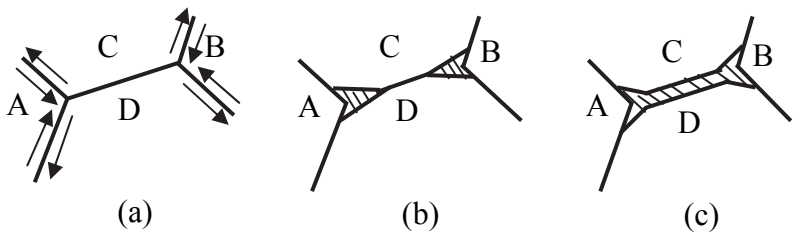


Fig. 8.10. Schematic diagram of the cuneiform-like cracks initiated and propagated along the boundary. (a) Triangle boundary, (b) initiation of the cuneiform-like crack, (c) propagation of the crack along the boundary

Because the boundaries and the carbide particles can effectively hinder the dislocation movement, and especially, the carbide particles can improve the cohesive strength between the grains and restrain the boundaries slipping during creep, therefore, it may be concluded that the carbide particles precipitated along the boundaries have an important effect on improving the creep resistance of the alloy. Although the carbide particles precipitated along the boundaries can improve the strength of the boundaries, the micro-cracks are still initiated and propagated along the boundaries, which suggests that the boundaries are still the weaker regions for causing fracture of the alloy during creep. And once, the carbide is continuously precipitated to form the film along the boundary, which may weaken the cohesive strength between the grains to damage the creep lifetimes of the alloy. The analysis is in agreement with the experimental results stated above.

When the alloy was solution treated at 1150 °C and cooled in oil bath at 120 °C, the carbon atoms were supernaturally dissolved in the matrix of the alloy due to quenching at lower temperature. The concentration supersaturation in the alloy promoted the carbon atoms for precipitating in the form of the fine carbide particles during creep under the applied higher tensile stress at 650 °C, in especially, the slipping trace regions support a bigger extruding stress for inducing the carbon atoms to precipitate in the form of the fine carbide particles along the slipping traces as shown in Fig. 8.7(c). This is thought to be a main reason of the fine carbides precipitated along the slipping traces.

On the other hand, when the alloy was solution treated at 1150 °C and cooled in molten salt at 583 °C, although the slipping traces appeared still in the matrix of the alloy during creep, no fine carbide particles were precipitated along the slipping traces, as shown in Fig. 8.8, due to the concentration supersaturation of the carbon atoms in the matrix is lower than the one of the alloy cooled in oil bath at 120 °C.

9. Conclusion

By means of hot isostatic pressing and heat treated at different temperatures, creep curves measurement and microstructure observation, an investigation had been made into the influence of hot isostatic pressing and heat treatment on the microstructure and creep behaviors of FGH95 nickel-base superalloy. Moreover, the deformation and fracture mechanisms of the alloy were discussed. The conclusions were mainly listed as follows:

1. When the alloy was hot isostatic pressed below the dissolving temperature of γ′ phase, as the HIP temperature increased, the size and amount of primary coarse γ′ phase decreased gradually in the PPB regions, and the size of the grains was equal to the one in the previous powder particles. With the HIP temperature increased to 1180°C, the coarse γ′ phase in the PPB was completely dissolved, and the grain of the alloy grew up obviously.
2. When the solution temperature was lower than the dissolving temperature of γ′ phase, after solution treated at 1140 °C, finer γ′ phase was dispersedly precipitated within the grains, and some coarser γ′ precipitates were distributed in the wider boundary regions where appeared the depleted zone of the fine γ′-phase. With the solution temperature increased, the amounts of the coarser γ′ phase and the zone of γ′-free phase decreased gradually.
3. After solution temperature increased to 1160 °C, the coarser γ′ phase in the alloy was fully dissolved, the fine secondary γ′ phase with high volume fraction was dispersedly

distributed within the grains, and the particles of (Nb, Ti)C carbide were precipitated along the boundaries. When the alloy was solution treated at 1165 °C, the size of the grains was obviously grown up, and the carbides were continuously precipitated to form the films along the boundaries.

4. During long term aging in the ranges of 450 °C and 550 °C, no obvious change in the grain size was detected in the alloy as the aging time prolonged, but the γ′ phase grew up slightly. With the aging time prolonging, the lattice parameters of the γ and γ′ phases increases slightly, but the misfit of γ γ′ phases decreased slightly.
5. Under the applied stress of 1034 MPa at 650 °C, the solution treated alloy cooled in molten salt displayed a better creep resistance. In the ranges of the applied temperatures and stresses, the creep activation energy of the alloy was measured to be Q = 590.3±20 kJ/mol.
6. The deformation mechanisms of the alloy during creep were the twinning, dislocations by-passing or shearing into the γ′ phase. The <110> super-dislocations shearing into the γ′ phase may be decomposed to form the configuration of (1/3)<112> super-Shockleys partial plus stacking fault.
7. During creep, the deformed features of the solution treated alloy cooled in oil bath was that the double orientation slipping of dislocations were activated, and the fine carbide particles were precipitated along the regions of the slipping traces. And the depleted zone of the fine γ′ phase was broken into the irregular piece-like shape due to the severe plastic deformation.
8. The deformed features of the alloy treated in molten salt were that the twinning and dislocation tangles were activated in the matrix of the alloy. Thereinto, the fact that the particles-like carbides were dispersedly precipitated within the grains and along the boundary might effectively restrain the dislocation slipping and hinder the dislocations movement, which is one important factor of the alloy possessing the better creep resistance and the longer creep lifetime.
9. In the later stage of creep, the slipping traces with the single or double orientations features appeared on the surface of the alloy. As the creep went on, the amount of the slipping traces increased to bring about the stress concentration, which might promote the initiation and propagation of the micro-cracks along the boundaries, this was thought to be the main fracture mechanism of the alloy during creep.

10. References

Domingue J. A., Boesch W. J., Radavich J. F.. (1980). S*uperalloys1980*. pp.335 – 344.

Flageolet B.; Jouiad M.; Villechaise P., et al. (2005). *Materials Science and Engineering A*, Vol. 399, pp. 199 – 205, ISSN: 0921 – 5093.

Hirth J. P. & Lothe J. (1968). Theory of Dislocations, 2nd ed., Wiley, New York, p.319

Hu B. F., Chen H. M., Li H. Y., et al. (2003). *Journal of Materials Engineering*, No.1, pp. 6 – 9, ISSN: 1001 – 4381.

Hu B. F., Yi F. Zh., et al. (2006). *Journal of University of Science and Technology Beijin*g, Vol.28, No.12, pp. 1121 – 1125, ISSN: 1001 – 053X.

Jia CH. CH., Yin F. ZH., Hu B. F., et al. (2006). *Materials Science and Engineering of Powder Metallurgy*, Vol.11, No.3, pp.176 – 179, ISSN: 1673 – 0224

Klepser C. A.. (1995). Scripta Metallurgical, Vol.33, No.4, pp. 589 – 596, ISSN: 1359 – 6462.

Kovarik L. , Unocic R. R. , Li J. , et al. (2009). *Journal of the Minerals*, Vol. 61, No.2, pp. 42 – 48, ISSN: 1047 – 4838.

Koble M.. (2001). *Materials Science and Engineering A*, Vol. 319-321, PP. 383 – 387, ISSN: 0921 – 5093.

Lherbier L.W. & Kent W. B.. (1990). *The International Journal of Powder Metallurgy*, Vol.26, No.2, pp. 131 – 137, ISBN: 0361 – 3488.

Liu D. M., Zhang Y., P. Liu Y., et al. (2006). *Powder Metallurgy Industry*, Vol.16, No.3 pp. 1-5, ISSN: 1006 – 6543.

Lu Z. Z., Liu C. L., Yue Z. F.. (2005). *Materials Science and Engineering A*, Vol. 395, pp. 153 – 159, ISSN: 0921 – 5093.

Park N. K. & Kim I. S.. (2001). *Journal of Materials Processing Technology*, Vol. 111, No.2, pp. 98 – 102, ISSN: 0924 – 0136.

Paul L.. (1988). *Powder Metallurgy Superalloys*, pp. 27 – 36.

Raujol S., Pettinari F., Locq D., et al. (2004). *Materials Science and Engineering A*, Vol. 387 – 389, pp.678 – 82, ISSN: 0921 – 5093.

Terzi S., Couturier R., Guetal L., et al. (2008). *Materials Science and Engineering A*, Vol. 483-484, pp. 598 – 601, ISSN: 0921 – 5093.

Viswanathan G. B., Sarosi P. M., Henry M. F., et al. (2005). *Acta Materialia*, Vol.53, pp. 3041~3057, ISSN: 1359 – 6454.

Unocic R. R., Viswanathan G. B., Sarosi P. M., et al. (2008). *Materials Science and Engineering A*, Vol. 483 – 484, pp. 25 – 32, ISSN: 0921 – 5093..

Wang P., Dong J. X., Yang L., et al. (2008). Material Review, Vol.22, No.6, pp. 61 – 64, ISSN: 1005 – 023X.

Yoo M. H. & Trinkaus H.. (1983). Metall. Trans., Vol. 14, No.4, pp. 547 – 561, ISSN: 1073-5623.

Zainul H. D.. (2007). *Materials and Design*, Vol.28, pp.1664 – 1667, ISSN: 0261 – 3069.

Zhang J. SH.. (2007). *High Temperature Deformation and Fracture of Material*. Beijing: Science Press, pp. 102 – 105, ISBN: 978 – 7 – 03 – 017774 – 2.

Zhang Y. W., Zhang Y., Zhang F. G., et al. (2002). *Transactions of Materials and Heat Treatment*, Vol.23, No.3, pp. 72 – 75, ISSN: 1009 – 6264.

Zhou J. B., Dong J. X., Xu Zh. Ch., et al. (2002). *Heat Treatment of Metals*, vol.27, No.6, pp.30 – 32. ISSN: 0254 – 6051.

15

Multi-Dimensional Calibration of Impact Models

Lucas G. Horta, Mercedes C. Reaves,
Martin S. Annett and Karen E. Jackson
NASA Langley Research Center, Hampton, VA
USA

1. Introduction

As computational capabilities continue to improve and the costs associated with test programs continue to increase, certification of future rotorcraft will rely more on computational tools along with strategic testing of critical components. Today, military standards (MIL-STD 1290A (AV), 1988) encourage designers of rotary wing vehicles to demonstrate compliance with the certification requirements for impact velocity and volume loss by analysis. Reliance on computational tools, however, will only come after rigorous demonstration of the predictive capabilities of existing computational tools. NASA, under the Subsonic Rotary Wing Program, is sponsoring the development and validation of such tools. Jackso*n* (2006) discussed detailed requirements and challenges associated with certification by analysis. Fundamental to the certification effort is the demonstration of verification, validation, calibration, and algorithms for this class of problems. Work in this chapter deals with model calibration of systems undergoing impact loads.

The process of model calibration, which follows the verification and validation phases, involves reconciling differences between test and analysis. Most calibration efforts combine both heuristics and quantitative methods to assess model deficiencies, to consider uncertainty, to evaluate parameter importance, and to compute required model changes. Calibration of rotorcraft structural models presents particular challenges because the computational time, often measured in hours, limits the number of solutions obtainable in a timely manner. Oftentimes, efforts are focused on predicting responses at critical locations as opposed to assessing the overall adequacy of the model. For example (Kamat, 1976) conducted a survey, which at the time, studied the most popular finite element analysis codes and validation efforts by comparing impact responses from a UH-1H helicopter drop test. Similarly, (Wittlin and Gamon, 1975) used the KRASH analysis program for data correlation of the UH-1H helicopter. Another excellent example of a rotary wing calibration effort is that of (Cronkhite and Mazza, 1988) comparing results from a U.S. Army composite helicopter with simulation data from the KRASH analysis program. Recently, (Tabiei, Lawrence, and Fasanella, 2009) reported on a validation effort using anthropomorphic test dummy data from crash tests to validate an LS-DYNA (Hallquist, 2006) finite element model. Common to all these calibration efforts is the use of scalar deterministic metrics.

One complication with calibration efforts of nonlinear models is the lack of universally accepted metrics to judge model adequacy. Work by (Oberkampf et al., 2006) and later (Schwer et al., 2007) are two noteworthy efforts that provide users with metrics to evaluate nonlinear time histories. Unfortunately, seldom does one see them used to assess model

adequacy. In addition, the metrics as stated in (Oberkampf et al., 2006) and (Schwer et al., 2007) do not consider the multi-dimensional aspect of the problem explicitly. A more suitable metric for multi-dimensional calibration exploits the concept of impact shapes as proposed by (Anderson et al., 1998) and demonstrated by (Horta et al., 2003). Aside from the metrics themselves, the verification, validation, and calibration elements, as described by (Roache, 1998; Oberkampf, 2003; Thacker, 2005; and Atamturktur, 2010), must be adapted to rotorcraft problems. Because most applications in this area use commercially available codes, it is assumed that code verification and validation have been addressed elsewhere. Thus, this work concentrates on calibration elements only. In particular, this work concentrates on deterministic input parameter calibration of nonlinear finite element models. For non-deterministic input parameter calibration approaches, the reader is referred to (Kennedy and O'Hagan, 2001; McFarland et al., 2008).

Fundamental to the success of the model calibration effort is a clear understanding of the ability of a particular model to predict the observed behavior in the presence of modeling uncertainty. The approach proposed in this chapter is focused primarily on model calibration using parameter uncertainty propagation and quantification, as opposed to a search for a reconciling solution. The process set forth follows a three-step approach. First, Analysis of Variance (ANOVA) as described in work by (Sobol et al., 2007; Mullershon and Liebsher, 2008; Homma and Saltelli, 1996; and Sudret, 2008) is used for parameter selection and sensitivity. To reduce the computational burden associated with variance based sensitivity estimates, response surface models are created and used to estimate time histories. In our application, the Extended Radial Basis Functions (ERBF) response surface method, as described by (Mullur, 2005, 2006) has been implemented and used. Second, after ANOVA estimates are completed, uncertainty propagation is conducted to evaluate uncertainty bounds and to gage the ability of the model to explain the observed behavior by comparing the statistics of the 2-norm of the response vector between analysis and test. If the model is reconcilable according to the metric, the third step seeks to find a parameter set to reconcile test with analysis by minimizing the prediction error using the optimization scheme proposed (Regis and Shoemaker 2005). To concentrate on the methodology development, simulated experimental data has been generated by perturbing an existing model. Data from the perturbed model is used as the target set for model calibration. To keep from biasing this study, changes to the perturbed model were not revealed until the study was completed.

In this chapter, a description of basic model calibration elements is described first followed by an example using a helicopter model. These elements include time and spatial multi-dimensional metrics, parameter selection, sensitivity using analysis of variance, and optimization strategy for model reconciliation. Other supporting topics discussed are sensor placement to assure proper evaluation of multi-dimensional orthogonality metrics, prediction of unmeasured responses from measured data, and the use of surrogates for computational efficiency. Finally, results for the helicopter calibrated model are presented and, at the end, the actual perturbations made to the original model are revealed for a quick assessment.

2. Problem formulation

Calibration of models is a process that requires analysts to integrate different methodologies in order to achieve the desired end goal which is to reconcile prediction

with observations. Although in the literature the word "model" is used to mean many different forms of mathematical representations of physical phenomena, for our purposes, the word model is used to refer to a finite element representation of the system. Starting with an analytical model that incorporates the physical attributes of the test article, this model is initially judged based on some pre-established calibration metrics. Although there are no universally accepted metrics, the work in this paper uses two metrics; one that addresses the predictive capability of time responses and a second metric that addresses multi-dimensional spatial correlation of sensors for both test and analysis data. After calibration metrics are established, the next step in the calibration process involves parameter selection and uncertainty estimates using engineering judgment and available data. With parameters selected and uncertainty models prescribed, the effect of parameter variations on the response of interest must be established. If parameter variations are found to significantly affect the response of interest, then calibration of the model can proceed to determine a parameter set to reconcile the model. These steps are described in more detail, as follows.

2.1 Time domain calibration metrics

Calibration metrics provide a mathematical construct to assess fitness of a model in a quantitative manner. Work by (Oberkampf, 2006) and (Schwer, 2007) set forth scalar statistical metrics ideally suited for use with time history data. Metrics in terms of mean, variance, and confidence intervals facilitate assessment of experimental data, particularly when probability statements are sought. For our problem, instead of using response predictions at a particular point, a vector 2-norm (magnitude of vector) of the system response is used as a function of time. An important benefit of using this metric is that it provides for a direct measure of multi-dimensional closeness of two models. In addition, when tracked as a function of time, closeness is quantified at each time step.

Because parameter values are uncertain, statistical measures of the metric need to be used to conduct assessments. With limited information about parameter uncertainty, a uniform distribution function, which is the least informative distribution function, is the most appropriate representation to model parameter uncertainty. This uncertainty model is used to create a family of N equally probable parameter vectors, where N is arbitrarily selected. From the perspective of a user, it is important to know the probability of being able to reconcile measured data with predictions, given a particular model for the structure and parameter uncertainty. To this end, let $Q(t,p) \triangleq \left\| v(t,p) \right\|_2$ be a scalar time varying function, in which the response vector v is used to compute the 2-norm of the response at time t, using parameter vector p. Furthermore, let $\underline{\sigma}(t) = \min_{\forall p} Q(t,p)$ be the minimum value over all parameter variations, and let $\bar{\sigma}(t) = \max_{\forall p} Q(t,p)$ be the maximum value. Using these definitions and N LS-DYNA solutions corresponding to equally probable parameter vectors, a calibration metric can be established to bound the probability of test values falling outside the analysis bounds as;

$$\mathrm{M}_1 = \mathrm{Prob}(\underline{\sigma}(t) < Q_e(t) \cap Q_e(t) > \bar{\sigma}(t)) << 1/N \tag{1}$$

where $Q_e(t)$ is the 2-norm of responses from the experiment. Note that N controls tightness of the bounds and also the number of LS-DYNA solutions required.

The use of norms, although convenient, tends to hide the spatial relationships that exist between responses at different locations in the model. In order to study this spatial multi-dimensional dependency explicitly, a different metric must be established.

2.2 Spatial multi-dimensional calibration metric

Spatial multi-dimensional dependency of models has been studied in classical linear dynamic problems in terms of mode shapes or eigenvectors resulting from a solution to an eigenvalue problem. Unfortunately, the nonlinear nature of impact problems precludes use of any simple eigenvalue solution scheme. Alternatively, an efficient and compact way to study the spatial relationship is by using a set of orthogonal impact shape basis vectors. Impact shapes, proposed by (Anderson, 1998 and later by Horta, 2003), are computed by decomposing the time histories using orthogonal decomposition. For example, time histories from analysis or experiments can be decomposed using singular value decomposition as

$$y(x,t) = \sum_{i=1}^{n} \sigma_i \phi_i(x) g_i(t) \tag{2}$$

In this form, the impact shape vector ϕ_i sized m x 1, contains the spatial distribution information for m sensors, g(t) contains the time modulation information, σ contains scalar values with shape participation factors, and n is the number of impact shapes to be included in the decomposition, often truncated based on allowable reconstruction error. Although Eq. (2) is written in continuous time form, for most applications, time is sampled at fixed intervals such that $t = k\Delta T$ where the integer k=0,…,L and ΔT is the sample time. From Eq. (2), the fractional contribution of the i^{th} impact shape to the total response is proportional to δ_i, defined as;

$$\delta_i = \sigma_i \Big/ \sum_{l=1}^{n} \sigma_l \tag{3}$$

Mimicking the approach used in classical dynamic problems, impact shapes can now be used to compare models using orthogonality. Orthogonality, computed as the dot product operation of vectors (or matrices), quantifies the projection of one vector onto another. If the projection is zero, vectors are orthogonal, i.e., uncorrelated. This same idea applies when comparing test and analysis impact shapes. Numerically, the orthogonality metric is computed as;

$$M_2 = \breve{\Phi}^T \breve{\Psi} \tag{4}$$

where $\breve{\Phi}$ is sized m x l with l measured impact shapes at m locations and $\breve{\Psi}$, sized m x l, are shapes computed using simulation data. Note that both $\breve{\Phi}$ and $\breve{\Psi}$ are normalized matrices such that $\breve{\Phi}^T \breve{\Phi} = I$ and $\breve{\Psi}^T \breve{\Psi} = I$. Because individual impact shape vectors are stacked column-wise, metric M_2 is a matrix sized l x l with diagonal values corresponding to the vector projection numerical value. If vectors are identical then their projection equals 1. Consequently, when evaluating models, multi-dimensional closeness with experiment is judged based on similarity of impact shapes and shape contributions. Two direct benefits of using impact shapes are discussed in the next two sections.

2.2.1 Algorithm for response interpolation

Adopting impact shapes as a means to compare models has two advantages. First, it allows for interpolation of unmeasured response points, and second, it provides a metric to conduct optimal sensor placement. During most test programs, the number of sensors used is often limited by the availability of transducers and the data acquisition system. Although photogrammetry and videogrammetry measurements provide significantly more data, even these techniques are limited to only those regions in the field of view of the cameras. At times, the inability to view responses over the full structure can mislead analysts as to their proper behavior. For this purpose, a hybrid approach has been developed to combine measured data with physics-based models to provide more insight into the full system response. Although the idea is perhaps new in the impact dynamics area, this approach is used routinely in modal tests where a limited number of measurements is augmented with predictions using the analytical stiffness matrix. This approach takes advantage of the inherent stiffness that relates the motion at different locations on the structure. Because in impact dynamic problems, the stiffness matrix is likely to be time varying, implementation of a similar approach is difficult. An alternative is to use impact shapes as a means to combine information from physics-based models with experimental data. Specifically, responses at unmeasured locations are related to measured locations through impact shapes. To justify the approach, Eq. (2) is re-written as;

$$\begin{pmatrix} y(t) \\ y_e(t) \end{pmatrix} = \sum_{i=1}^{n} \begin{bmatrix} \phi_i \\ \kappa_i \end{bmatrix} \sigma_i g_i(t) = \begin{bmatrix} \Phi \\ \mathrm{K} \end{bmatrix} SG(t) \tag{5}$$

where the matrix partitions are defined as

$$\begin{bmatrix} \Phi \\ \mathrm{K} \end{bmatrix} \triangleq \begin{bmatrix} \phi_1 & \cdots & \phi_n \\ \kappa_1 & \cdots & \kappa_n \end{bmatrix}; S \triangleq \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \end{bmatrix}; G(t) \triangleq \begin{bmatrix} g_1(t) \\ \\ g_n(t) \end{bmatrix} \tag{6}$$

In contrast to Eq. (2), Eq. (5) shows explicitly responses at an augmented set of locations named $y_e(t)$, constructed using impact shapes κ_i at q unmeasured locations.
Using Eq. (5) with experimental data, the time dependency of the response can be computed as

$$SG(t) = \left(\Phi^T \Phi\right)^{-1} \Phi^T y(t) \tag{7}$$

Consequently, predictions for unmeasured locations can now be computed as

$$y_e(t) = \mathrm{K}\left(\Phi^T \Phi\right)^{-1} \Phi^T y(t) \tag{8}$$

Although Eq.(7) requires a matrix inversion, the rank of this matrix is controlled by sensor placement. Hence, judicious pretest sensor placement must be an integral part of this process. Fortunately, because the impact shapes are computed using singular value decomposition, they form an orthonormal set of basis vectors, i.e. $(\Phi^T \Phi)^{-1} \cong I$. It is important to note that measured data are used to compute the impact shapes (at sensor locations) and the time dependent part of the response, whereas data from the analytical model are used to compute impact shapes at all unmeasured locations.

2.2.2 Optimum sensor placement for impact problems

Optimal sensor placement must be driven by the ultimate goals of the test. If model calibration is the goal, sensor placement must focus on providing information to properly evaluate the established metrics. In multi-dimensional calibration efforts using the orthogonality metric, sensor placement is critical because if sensors are not strategically placed, it is impossible to distinguish between impact shapes. Fortunately, the use of impact shapes enables the application of well established sensor placement algorithms routinely used in modal tests. Placement for our example used the approach developed by (Kammer, 1991). Using this approach sensors are placed to ensure proper numerical conditioning of the orthogonality matrix.

2.3 Parameter selection

The parameter selection (parameters being in this case material properties, structural dimensions, etc.) process relies heavily on the analyst's knowledge and familiarity with the model and assumptions. Formal approaches like Phenomena Identification and Ranking Table (PIRT), discussed by (Wilson and Boyack, 1998), provide users with a systematic method for ranking parameters for a wide class of problems. Elements of this approach are used for the initial parameter selection. After an initial parameter selection is made, parameter uncertainty must be quantified empirically if data are available or oftentimes engineering judgment is ultimately used. With an initial parameter set and an uncertainty model at hand, parameter importance is assessed using uncertainty propagation. That is, the LS-DYNA model is exercised with parameter values created using the (Halton, 1960) deterministic sampling technique. Time history results are processed to compute the metrics and to assess variability. A by-product of this step produces variance-based sensitivity results which are used to rank the parameters. In the end, adequacy of the parameter set is judged based on the probability of one being able to reconcile test with analysis. If the probability is zero, as will be shown later in the example, the parameter selection must be revisited.

2.4 Optimization strategy

With an adequate set of parameters selected, the next step is to use an optimization procedure to determine values that reconcile test with the analysis. A difficulty with using classical optimization tools in this step is in the computational time it takes to obtain LS-DYNA solutions. Although in the helicopter example the execution time was optimized to be less than seven minutes, the full model execution time is measured in days. For this reason, ideally optimization tools for this step must take advantage of all LS-DYNA solutions at hand. To address this issue, optimization tools that use surrogate models in addition to new LS-DYNA solutions are ideal. For the present application the Constrained Optimization using Response Surface (CORS) algorithm, developed by (Regis and Shoemaker, 2005), has been implemented in MATLAB for reconciliation. Specifically, the algorithm starts by looking for parameter values away from the initial set of LS-DYNA solutions, then slowly steps closer to known solutions by solving a series of local constrained optimization problems. This optimization process will produce a global optimum if enough steps are taken. Of course, the user controls the number of steps and therefore the accuracy and computational expense in conducting the optimization. In cases where the predictive capability of the surrogate model is poor, CORS adds solutions in needed areas. Because parameter uncertainty is not used explicitly in the optimization, this approach is considered to be deterministic. If a probabilistic approach was used instead (see

Kennedy and O'Hagan, 2001; McFarland et al., 2008).), in addition to a reconciling set, the user should also be able to determine the probability that the parameter set found is correct. Lack of credible parameter uncertainty data precludes the use of probabilistic optimization methods at this time, but future work could use the same computational framework.

2.5 Analysis of variance

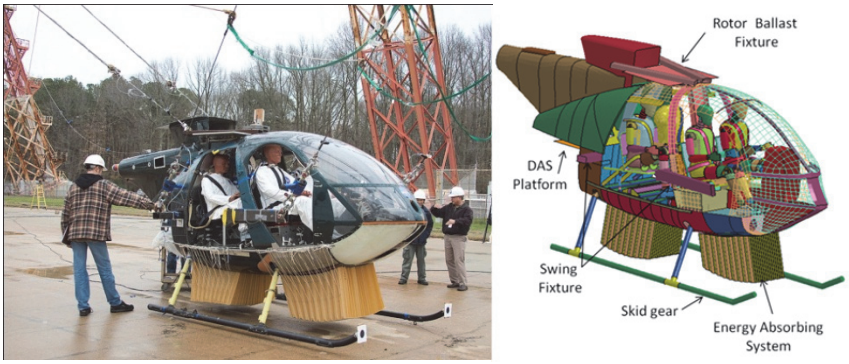
Parameter sensitivity in most engineering fields is often associated with derivative calculations at specific parameter values. However, for analysis of systems with uncertainties, sensitivity studies are often conducted using ANOVA. In classical ANOVA studies, data is collected from multiple experiments while varying all parameters (factors) and also while varying one parameter at a time. These results are then used to quantify the output response variance due to variations of a particular parameter, as compared to the total output variance when varying all the parameters simultaneously. The ratio of these two variance contributions is a direct measure of the parameter importance. Sobol et al. (2007) and others (Mullershon and Liebsher, 2008; Homma and Saltelli, 1996; and Sudret, 2008) have studied the problem as a means to obtain global sensitivity estimates using variance based methods. To compute sensitivity using these variance based methods, one must be able to compute many response predictions as parameters are varied. In our implementation, after a suitable set of LS-DYNA solutions are obtained, response surface surrogates are used to estimate additional solutions.

2.6 Response surface methodology

A response surface (RS) model is simply a mathematical representation that relates input variables (parameters that the user controls) and output variables (response quantities of interest), often used in place of computationally expensive solutions. Many papers have been published on response surface techniques, see for example (Myers, 2002). The one adopted here is the Extended Radial Basis Functions (ERBF) method as described by (Mullur, 2005, 2006). In this adaptive response surface approach, the total number of RS parameters computed equals $N(3n_p+1)$, where n_p is the number of parameters and N is the number of LS-DYNA solutions. The user must also prescribe two additional parameters: 1) the order of a local polynomial (set to 4 in the present case), and 2) a smoothness parameter (set to 0.15 here). Finally, the radial basis function is chosen to be an exponentially decaying function $e^{-(p-p_i)^2/2r_c^2}$ with characteristic radius r_c set to 0.15. A distinction with this RS implementation is that ERBF is used to predict full time histories, as opposed to just extreme values. In addition, ERBF is able to match the responses used to create the surrogate with prediction errors less than 10^{-10}.

3. Description of helicopter test article

A full-scale crash test of an MD-500 helicopter, as described by (Annett and Polanco, 2010), was conducted at the Landing and Impact Research (LandIR) Facility at NASA Langley Research Center (LaRC). Figure 1a shows a photograph of the test article while it was being prepared for test, including an experimental dynamic energy absorbing honeycomb structure underneath the fuselage designed by (Kellas, 2007). The airframe, provided by the US Army's Mission Enhanced Little Bird (MELB) program, has been used for civilian and military applications for more than 40 years. NASA Langley is spearheading efforts to develop analytical models capable of predicting the impact response of such systems.



a) during test preparations b) "as-tested" FEM

Fig. 1. MD-500 helicopter model.

4. LS-DYNA model description

To predict the behavior of the MD-500 helicopter during a crash test, an LS-DYNA (Hallquist, 2006) finite element model (FEM) of the fuselage, as shown in Figure 1b, was developed and reported in (Annett and Polanco, 2010). The element count for the fuselage was targeted to not exceed 500,000 elements, including seats and occupants; with 320,000 used to represent the energy absorbing honeycomb and skid gear. Shell elements were used to model the airframe skins, ribs and stiffeners. Similarly, the lifting and pullback fixtures, and the platform supporting the data acquisition system (mounted in the tail) were modeled using rigid shells. Ballast used in the helicopter to represent the rotor, tail section, and the fuel was modelled as concentrated masses. For materials, the fuselage section is modeled using Aluminum 2024-T3 with elastic-plastic properties, whereas the nose is fiberglass and the engine fairing is Kevlar fabric. Instead of using the complete "as-tested" FEM model, this study uses a simplified model created by removing the energy absorbing honeycomb, skid gears, anthropomorphic dummies, data acquisition system, and lifting/pull-back fixtures. After these changes, the resulting simplified model is shown in Figure 2. Even with all these components removed, the simplified model had 27,000 elements comprised primarily of shell elements to represent airframe skins, ribs and stiffeners. The analytical test case used for calibration, simulates a helicopter crash onto a hard surface with vertical and horizontal speeds of 26 ft/sec and 40 ft/sec, respectively. For illustration, Figure 3 shows four frames from an LS-DYNA simulation as the helicopter impacts the hard surface.

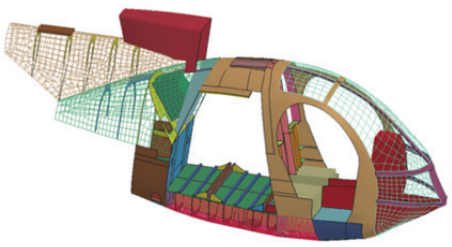


Fig. 2. Simplified finite element model.

5. Example results

Results described here are derived from the simplified LS-DYNA model, as shown in Figure 2. This simplified model reduced the computational time from days to less than seven minutes and allowed for timely debugging of the software and demonstration of the methodology, which is the main focus of the chapter. Nonetheless, the same approach can be applied to the complete "as-tested" FEM model without modifications.

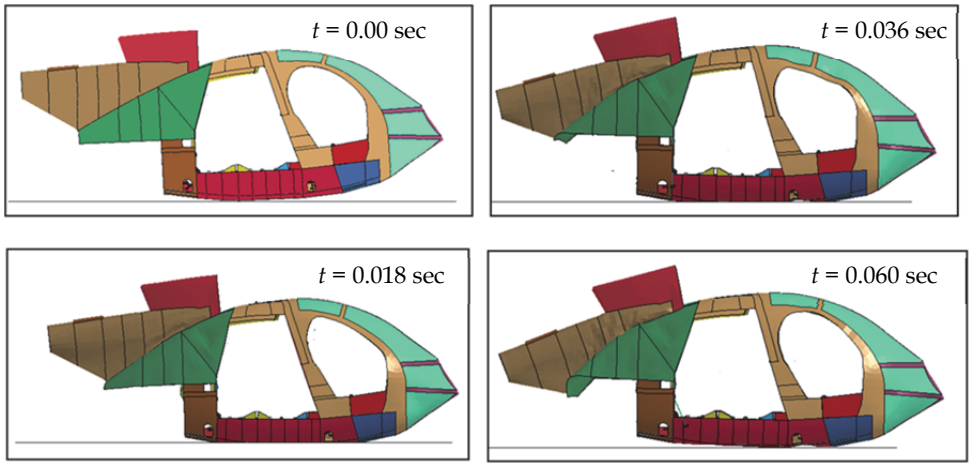
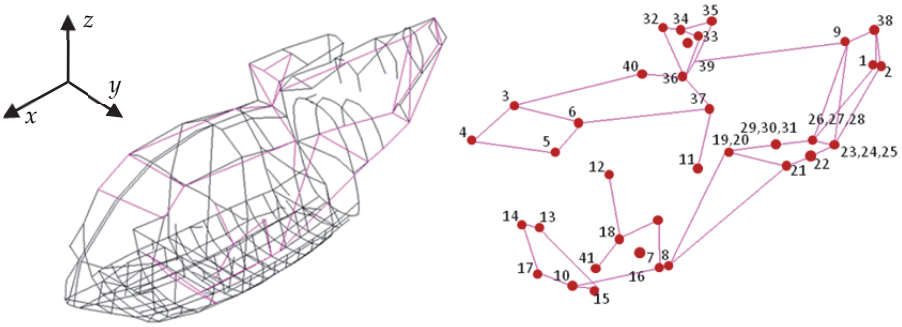


Fig. 3. Four frames of the LS-DYNA simulation as the helicopter impacts the hard surface.

For evaluation purposes, simulated data are used in lieu of experimental data. Because more often than not analytical model predictions do not agree with the measured data, the simplified model was arbitrarily perturbed. Knowledge of the perturbations and areas affected are not revealed until the entire calibration process is completed. Data from this model, referred to as the perturbed model, takes the place of experimental data. In this study, no test uncertainty is considered. Therefore, only 1 data set is used for test.



a) wireframe with 405 nodes b) sensor placement

Fig. 4. Helicopter wireframe for a) simplified model b) simulated test sensor placement and numbering.

Figure 4(a) depicts a wireframe of the simplified model showing only 405 nodes. Superimposed is a second wiring frame with connections to 34 nodes identified by an optimal sensor placement algorithm. At each node there can be up to 3 translational measurements, however, here the placement algorithm was instructed to place only 41 sensors. Figure 4(b) shows the location for the 41 sensors. Results from the optimal sensor placement located 8 sensors along the x direction, 10 sensors along the y direction, and 23 sensors along the z direction.

5.1 Initial parameter selection

Calibration efforts begin by selecting model parameters thought to be uncertain. Selecting these parameters is perhaps the most difficult step. Not knowing what had been changed in the perturbed model, the initial study considered displacements, stress contours, and plastic strain results at different locations on the structure before selecting the modulus of elasticity and tangent modulus at various locations. The parameters and uncertainty ranges selected are shown in Table 1. Without additional information about parameter uncertainty, the upper and lower bounds were selected using engineering judgment with the understanding that values anywhere between the bounds were equally likely.

No.	Parameter Description	Nominal	LowerBound	Upper Bound
1	E back panel (lbs/in^2)	10,000,000	8,000,000	12,000,000
2	E subfloor ribs (lbs/in^2)	10,000,000	8,000,000	12,000,000
3	E keel beam web (lbs/in^2)	9,880,000	7,904,000	11,856,000
4	E stinger upper tail (lbs/in^2)	10,000,000	8,000,000	12,000,000
5	E stinger lower tail (lbs/in^2)	10,000,000	8,000,000	12,000,000
6	E_t subfloor ribs (lbs/in^2)	134,200	107,360	161,040
7	E_t keel beam web (lbs/in^2)	134,200	107,360	161,040
8	E_t lower tail stinger (lbs/in^2)	134,200	107,360	161,040

Table 1. Initial parameter set description

With the parameter uncertainty definition in Table 1, LS-DYNA models can be created and executed to study the calibration metrics as described earlier. As an example, 150 LS-DYNA runs with the simplified model were completed while varying parameters over the ranges shown in Table 1. To construct the uncertainty bounds for each of the 150 runs, $Q(t,p)$ is computed from velocities at 41 sensors (see Figure 4) and plotted in Figure 5 as a function of time; analysis (dashed-blue) and the simulated test (solid-red). With this sample size, the probability of being able to reconcile test with analysis during times when test results are outside the analysis bounds is less than 1/150 (recall that the simulated test data is from the perturbed model). Figure 5 shows that during the time interval between 0.01 and 0.02 seconds, the analysis bounds are above the test. Therefore, it is unlikely that one would be able to find parameter values within the selected set to reconcile analysis with test. This finding prompted another look at parameter selection and uncertainty models to determine a more suitable set.

5.2 Revised parameter selection

A second search for a revised parameter set involved conversations with the model developer and additional runs while varying parameter bounds to see their effect on M_1.

The second set of parameters selected, after considering several intermediate sets, consisted of thicknesses at various locations in the structure. A concern with varying thickness is its effect on structural mass. However, because 80% of the helicopter model is comprised of non-structural masses, thickness changes had little impact on the total mass. Table 2 shows a revised parameter set and ranges selected for the second study.

6. Evaluation of calibration metrics with revised parameter set

Results for metric M_1 using the revised parameter set are shown in Figure 6; solid red is $Q_e(t)$ with the simulated test data and in dotted blue lines are analysis bounds using 50 LS-DYNA runs. With 50 runs, the probability that LS-DYNA would produce results outside these bounds is less than 1/50. Consequently, if the test results are outside these bounds, the probability of reconciling the model with test is also less than 1/50. Even though Figure 6 shows that, in certain areas, test results are very close to the analysis bounds; this new parameter set provides enough freedom to proceed with the calibration process.

No.	Parameter Description	Nominal	Lower Bound	Upper Bound	Calibrated Value
1	Keel beam stiffener thickness (in)	0.020	0.015	0.025	0.0161
2	Belly panel thickness (in)	0.090	0.08	0.135	0.1008
3	Keel beam thickness (in)	0.040	0.035	0.045	0.0358
4	Lower tail thickness (in)	0.040	0.035	0.045	0.0414
5	Back panel thickness (in)	0.020	0.015	0.025	0.0166
6	Upper tail thickness (in)	0.020	0.015	0.025	0.0168

Table 2. Revised parameter description

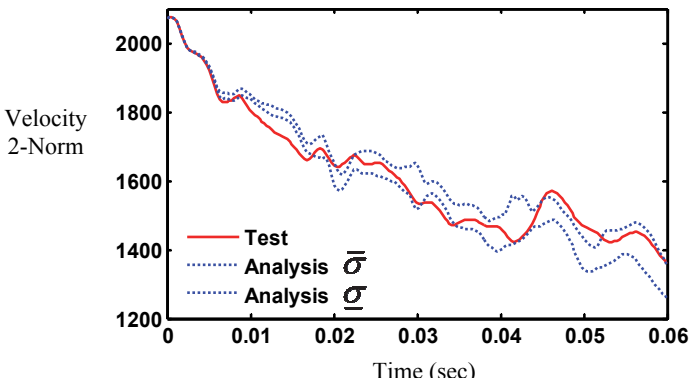


Fig. 5. Velocity vector 2-norm for analysis (with 150 LS-DYNA runs) and for simulated test.

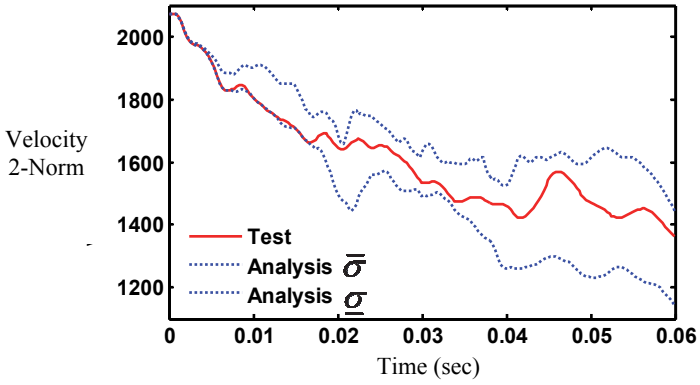


Fig. 6. Velocity vector 2-norm for analysis (with 50 LS-DYNA runs) and for simulated test.

Thus far, in this study, metric M_1 has been used exclusively to evaluate parameter adequacy and uncertainty bounds. What is missing from this evaluation is how well the model predicts the response at all locations. Considering that impact shapes provide a spatial multi-dimensional relationship among different locations, two models with similar impact shapes, all else being equal, should exhibit similar responses at all sensor locations. With this in mind, orthogonality results for the simplified model versus "test", i.e., the perturbed model, are shown in Figure 7. Essentially, the matrix M_2, as defined in Eq. (4), is plotted with analysis

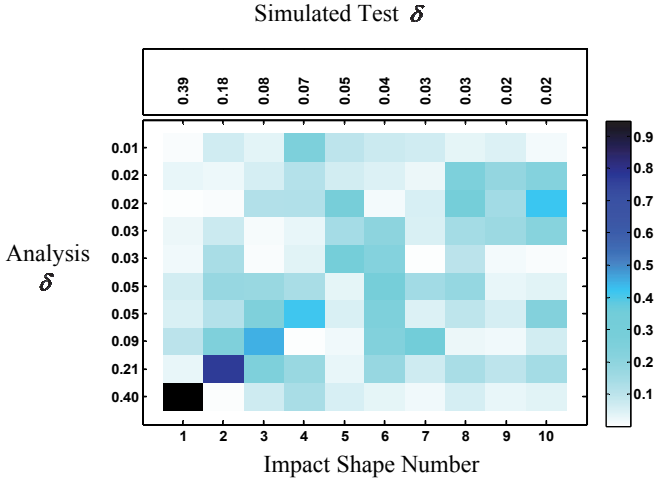


Fig. 7. Orthogonality results using impact shapes from the simulated test and baseline model.

along the ordinate and test along the abscissa. Colors represent the numerical value of the vector projections, e.g. a value of 1 (black) indicates perfect matching between test and analysis. Listed on the labels are the corresponding shape contribution to the response for both analysis (ordinate) and simulated test (top axis). For example, the first impact shape for analysis (bottom left) contributes 0.4 of the total response as compared to 0.39 for test. It is apparent that initial impact shape matching is poor at best with the exception of the first two shapes. An example of an impact shape is provided in Figure 8. Here, a sequence of 8 frames for the test impact shape number 2 (contribution $\delta_2 = 0.18$) expanded to 405 nodes, is shown. Motion of the tail and floor section of the helicopter dominates.

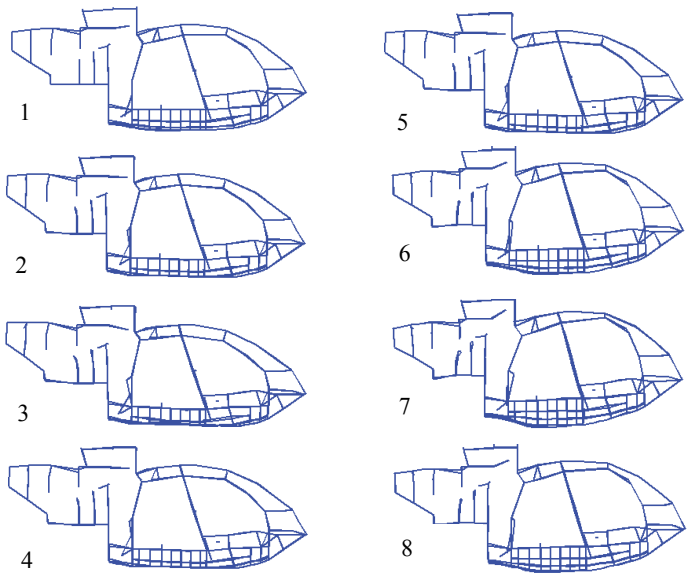


Fig. 8. Test impact shape number 2 ($\delta_2 = 0.18$) animation sequence.

6.1.1 Sensitivity with revised parameter set

Another important aspect of the calibration process is in understanding how parameter variations affect the norm metric $Q(t,p)$. This information is used as the basis to remove or retain parameters during the calibration process. As mentioned earlier, sensitivity results in this study look at the ratio of the single parameter variance to the total variance of $Q(t,p)$. This ratio is plotted in Figure 9 for each of the six parameters considered (as defined in Table 2). Along the abscissa is time in seconds and the ordinate shows contribution to variance. Colors are used to denote individual parameter contribution; total sum should approach 1 when no parameter interaction exists. In addition, $Q(t,p)$ is shown across the top, for reference. Because only 50 LS-DYNA runs are executed, an ERBF surrogate model is used to estimate responses with 1000 parameter sets for variance estimates. From results in Figure 9, note that parameter contributions vary significantly over time but for simulation times greater than 0.04 sec the upper tail thickness clearly dominates.

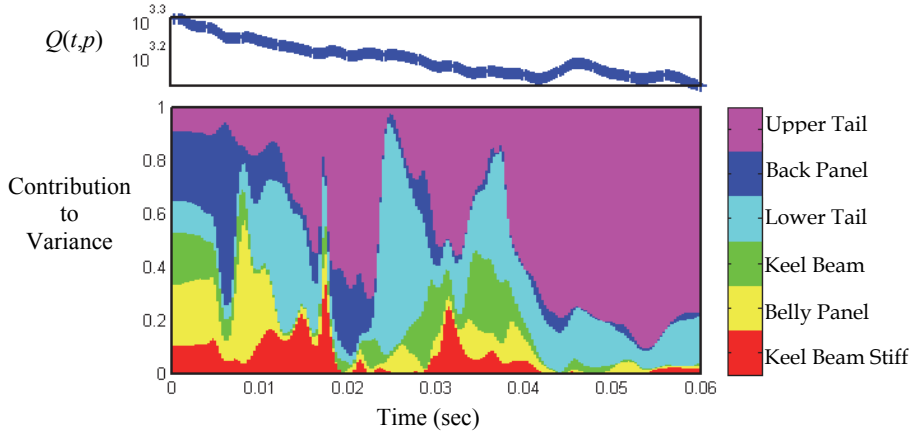


Fig. 9. Norm metric $Q(t,p)$ and sensitivity for thickness parameters.

6.2 Optimization for calibration parameters

The optimization problem is instructed to find parameter values to minimize the natural log of the prediction error using CORS. Starting with only 50 LS-DYNA solutions, CORS is allowed to compute 60 additional LS-DYNA runs. After completing the additional runs, the calibrated parameter set in Table 2 produced an overall prediction error reduction of 8%. Because the number of LS-DYNA runs is set by the user, the calibrated parameter set in Table 2 is not a converged set but rather an intermediate set. Nonetheless, it is referred to as the calibrated parameter set. The optimization process is purposely set up this way to control the computational expense in computing new LS-DYNA solutions. To show the impact of the parameter changes on the vector norm $Q(t,p)$, Figure 10 shows a comparison

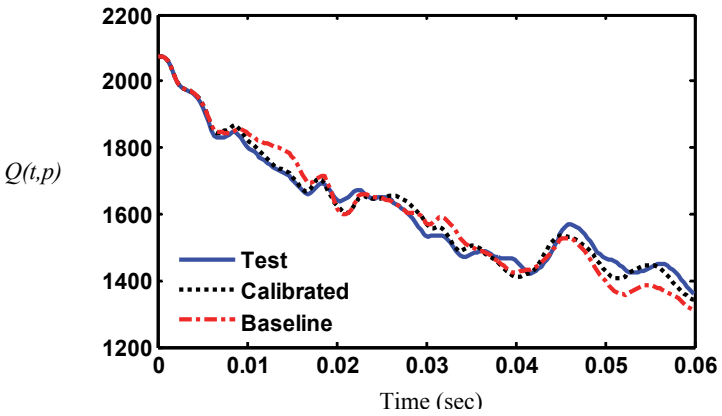


Fig. 10. Comparison of the velocity 2-norm using simulated test, baseline, and calibrated model data.

of results with data from the simulated test (solid blue), the baseline model using nominal values (dotted red), and the calibrated model (dashed-dot black). When comparing vector norms an overall assessment of the model predictive capability is quantified. However, users may prefer to evaluate improvements at specific sensor locations. For example, Figure 11 shows the velocity for the simulated test, baseline model, and the calibrated models at location 34 (see fig 1) in the z-direction. Improvements like those shown in Figure 11 are common for most locations but not all. Furthermore, after calibration, significant improvements are also apparent in the orthogonality results, as shown in Figure 12. As in Figure 7, results from the calibrated model are along the ordinate and the abscissa contains simulated test impact shapes. When comparing results against the baseline model (see Figure 7), it is clear that impact shape predictions have improved significantly, i.e., responses from the calibrated model and test are closer.

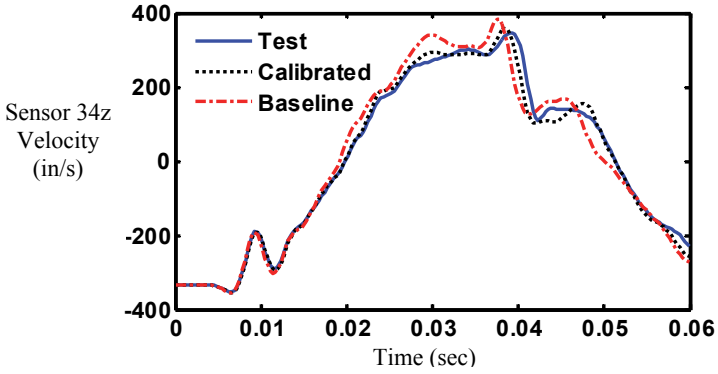


Fig. 11. Comparison of velocity response for the simulated test, baseline, and calibrated model data.

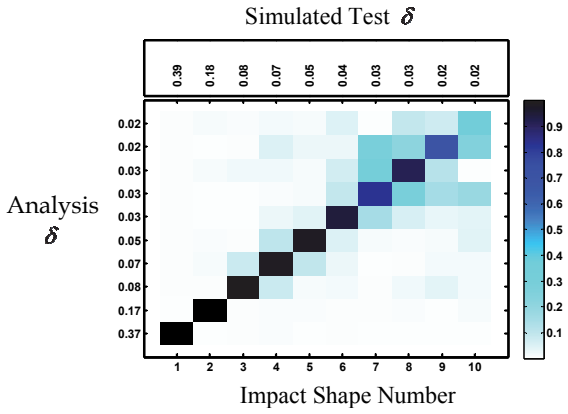
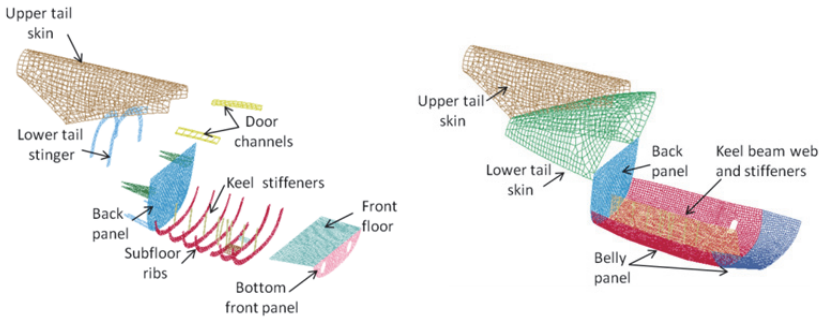


Fig. 12. Orthogonality results using impact shapes from the simulated test and calibrated model.

6.3 Revealing the correct answer

After completing the calibration process it is instructive to examine the actual changes made to the simplified model that resulted in the perturbed model. Because data from the perturbed model served as our target test data set, revealing the changes effectively gives the reader the true answer to the problem. Figure 13a depicts all the sections altered in the perturbed model. Alterations consisted of ± 20 % change in thickness at these locations. In contrast, Figure 13b shows the sections updated during the calibration process. It is interesting to see that although exact changes were not identified, regions where changes needed to be made were identified.



a) altered sections in perturbed model b) sections updated during calibration

Fig. 13. Sections altered to create the perturbed model and changes suggested during calibration.

7. Concluding remarks

An approach to conduct model calibration of nonlinear models has been developed and demonstrated using data from a simulated crash test of a helicopter. Fundamental to the approach is the definition and application of two calibration metrics: 1) metric M_1 compares the statistical bounds of the 2-norm of the velocity response from analysis versus test, and 2) metric M_2 evaluates the orthogonality between test and analysis impact shapes. The ability to reconcile analysis with test, assessed using metric M_1, is evaluated by computing the system response when parameters are varied to establish response bounds. Once the adequacy of the model is established, the process of reconciling model and test proceeds using constrained optimization to minimize the natural log of the prediction error. The optimization approach takes advantage of surrogate models to reduce the computational time associated with executing hundreds of LS-DYNA runs. Because the computational time is significant, users must trade the number of LS-DYNA solutions allowed against prediction error.

For the simulated helicopter example studied in this chapter, a flawed parameter set was evaluated and found to be inadequate using metric M_1. After determining a revised set of parameters, results from calibration show an overall prediction error reduction of 8%. Although this overall reduction may appear to be relatively small, improvements at specific sensor locations can be significantly better. Furthermore, improvements in orthogonality values after calibration resulted in matching 7 impact shapes with orthogonality values greater than 0.9 as compared to the initial model that only matched 1 impact shape.

Improvements in orthogonality produced multi-dimensional improvements in the overall model predictive capability. Finally, because the calibration process was demonstrated using simulated experiment data, it was shown that although exact changes were not identified, regions where changes needed to be made were identified.

8. Acknowledgement

The authors would like to thank Susan Gorton, Manager of the Subsonic Rotary Wing Project of NASA's Fundamental Aeronautics Program, for her support of this activity.

9. References

Anderson, M.C., Gan, W., & Haselman, T.K. (1998). Statistical Analysis of Modeling Uncertainty and Predictive Accuracy for Nonlinear Finite Element Models. Proceedings of the 69th Shock and Vibration Symposium, Minneapolis/St. Paul, MN, USA.

Anonymous, (1988). Military Standard MIL-STD-1290A (AV), Light Fixed and Rotary-Wing Aircraft Crash Resistance. Department of Defense, Washington, DC, USA, September 1988.

Annett, M.S, & Polanco, M.A. (2010). System-Integrated Finite Element Analysis of a Full-Scale Helicopter Crash Test with Deployable Energy Absorbers. *Proceedings of the 66th American Helicopter Annual Forum*, Phoenix, AZ, USA, May 2010.

Atamturktur, S, Hemez, F, & Unal, C. (2010). Calibration Under Uncertainty for Finite Element Models of Masonry Monuments. *Los Alamos National Laboratory (LANL) Technical Report,* LA-14414, February 2010.

Cronkhite, J.D., & Mazza, L.T. (1988). KRASH Analysis Correlation with the Bell ACAP Full-Scale Aircraft Crash Test. *Proceedings of the National Technical Specialists' Meeting on Advanced Rotorcraft Structures: Requirements Vs. Opportunities*, a Look at the Future. VA, USA, October 1988.

Hallquist, J.O. (2006). *LS-DYNA Keyword User's Manual. Version 971*, Livermore Software Technology Company, Livermore, CA, USA.

Halton, J.H. (1960). On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals. *Numerische Mathematik*, Vol. 2, pp. 84-90.

Homma, T, & Saltelli, A. (1996). Importance Measures in Global Sensitivity Analysis of Nonlinear Models. *Reliability Engineering and System Safety*, Vol. 52, Issue 1, pp. 1-17, April 1996.

Horta, L.G., Lyle, K.H., & Lessard, W.B. (2003). Evaluation of Singular Value Decomposition Approach for Impact Dynamic Data Correlation. NASA TM 2003-212657, October 2003.

Jackson, K.E., Fasanella, E.L., & Lyle, K.H. (2006). Crash Certification by Analysis-Are We There Yet? *Proceedings of the 62nd Annual Forum of American Helicopter Society*, Phoenix AZ, USA, May 2006.

Kamat M.P. (1976). Survey of Computer Programs for Prediction of Crash Response and of Its Experimental Validation. Measurement and Prediction of Structural and Biodynamic Crash-Impact Response. *Proceedings of the Winter Annual Meeting of the ASME*, pp. 33-48, December 1976.

Kammer, D.C. (1991). Sensor Placement for On-Orbit Modal Identification and Correlation of Large Space Structures. *Journal of Guidance, Control, and Dynamics*, Vol. 14, No. 2, pp. 251-259, April 1991.

Kellas, S. and Jackson, K. E. (2007). Deployable System for Crash-Load Attenuation. *Proceedings of the 63rd American Helicopter Annual Forum*, VA Beach, USA, May 2007.

Kennedy, M.C., & O' Hagan, A. (2001). Bayesian Calibration of Computer Models. *Journal of Royal Statistical Society*, Vol. 63, Part 3, pp. 425-464, January 2002.

McFarland, J., Mahadevan, S., Romero, V., Swiler, L, (2008). Calibration and Uncertainty Analysis for Computer Simulations with Multivariate Output. *AIAA Journal*, Vol. 46, No. 5, pp. 1253-1265, May 2008.

Mullershon, H., & Liebsher, M. (2008). Statistics and Non-Linear Sensitivity Analysis with LS-OPT and DSPEX. *Presented at the10th International LS-DYNA Users Conference*, pp. 4-1-4-13, Dearborn, MI, June 2008.

Mullur, A., & Messac, A. (2005). Extended Radial Basis Functions: More Flexible and Effective Metamodeling. *AIAA Journal*, Vol., 43, No. 6, June 2006.

Mullur, A. & Messac, A. (2006). Metamodeling Using Extended Radial Basis Functions: A Comparative Approach. *Engineering with Computers*, Vol. 21, pp. 203-217, December 2005.

Myers, R.H. & Montgomery, D.C. (2002): *Response Surface Methodology*. 2nd Edition, Copyright © 2002 John Wiley and Sons, Inc. USA.

Oberkampf, W.L. & Barone, M.F. (2006). Measures of Agreement Between Computation and Experiment: Validation Metrics. *Journal of Computational Physics*, 217, pp. 5-56, June 2006.

Oberkampf, *W*.L., Trucano, T.G., & Hirsch, C. (2003). Verification, Validation, and Predictive Capability in Computational Engineering and Physics. SAND 2003-3769, February 2003.

Regis, R.G., & Shoemaker, C.A., (2005). Constrained Global Optimization of Expensive Black Box Functions Using Radial Basis Functions. *Journal of Global Optimization* (2005) 31: pp. 153-171, September 2005.

Roache, P.J. (1998): *Verification and Validation in Computational Science and Engineering*, Hermosa Publishers, Albuquerque, NM, USA.

Schwer, L.E. (2007). Validation Metrics for Response Histories: Perspectives and Case Studies. *Engineering with Computers*, Vol. 23, No. 4, pp. 295-309, June 2007.

Sobol, I.M., Tarantola, S., Gatelli, D., Kucherenko, S.S, & Mauntz, W. (2007). Estimating Approximation Error When Fixing Unessential Factors in Global Sensitivity Analysis. *Reliability Engineering and System Safety,* Vol. 92. pp. 957-960, July 2007.

Sudret, B. (2008). Global Sensitivity Analysis Using Polynomial Chaos Expansion. *Reliability Engineering and System Safety,* Vol. 93, pp. 964-979, July 2008.

Tabiei, AL, Lawrence, C., & Fasanella, E.L. (2009). Validation of Finite Element Crash Test Dummy Models for Predicting Orion Crew Member Injuries During a Simulated Vehicle Landing. NASA TM-2009-215476, March 2009.

Thacker, B.H. (2005). The Role of Non-determinism in Computational Model Verification and Validation. *Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, & Materials Conference*, Austin TX, USA, AIAA 2005-1902, April 2005.

Wittlin, G. & Gamon, M.A. (1975). Experimentally Verified Analytical Techniques for Predicting Vehicle Crash Responses. *Proceedings of the AIAA 11th Annual Meeting and Technical Display*, AIAA Paper No. 75-273, February 1975.

Wilson, G.E., & Boyack, B.E. (1998). The Role of PIRT in Experiments, Code Development and Code Applications Associated with Reactor Safety Assessment, Nuclear Engineering and Design, Vol. 186, pp. 23-37, November 1998.

Part 4

Avionics, Control and Operations

16

An Agile Cost Estimating Methodology for Aerospace Procurement Operations: Genetic Causal Cost CENTRE-ing

R. Curran[1], P. Watson[2] and S. Cowan[3]
[1]*Air Transport and Operations,*
Delft University of Technology, Kluyverweg 1, Delft,
[2]*Queens University Belfast*
[3]*Procurement and Logistics, Bombardier Aerospace Belfast*
[1]*The Netherlands*
[2,3]*UK*

1. Introduction

In an effort to be more competitive, aerospace companies have to embrace a more integrated and concurrent approach to their operational processes. The aim is to meet the key requirements of being more cost effective, lean and agile while delivering consistently high quality performance in their operational practices. This requirement is further set against the backdrop of changeable global events, fluctuating markets, and technological progress in both the commercial and military spheres. Therefore, cost engineering issues are becoming increasingly dominant in Product Lifecycle Management (PLM) and as a consequence, the role of procurement is recognized as evermore influential due to its impact on acquisition cost.
In an effort to address some of the above challenges through practical means, the research presented investigates the development of a methodology and associated tooling for the estimating of supply chain cost management (Pugh et al, 2010a; Pugh et al, 2010b). The main aim is to provide an agile approach to cost estimating that can draw on the in-house engineering experience of an aerospace company, their procurement knowledge, product specification and their knowledge of the procurement market. This is integrated into a methodology that is generic and can therefore assimilate whatever information and relevant knowledge is available in a manner that can be utilized in an agile manner, i.e. dealing with large amounts of historic information in order to provide a agile estimating capability that is based on all of the information (past, present and projected) relating to the acquisition of new supply, parts, and assemblies. The following presents the methodology developed and a number of large case studies undertaken with Bombardier Aerospace Belfast to validate the accuracy and relevancy of the derived tools.

2. Aerospace procurement context

The importance of the procurement function is highlighted by the fact that it is common today for aerospace Original Equipment Manufacturers (OEM) to externally procure as much as 80% of their programmes externally [Flemming, (2003); Dubois, (2003)]]. Momme

(2002) even states that in general any typical industrial company spends 50-85% of its turnover on purchased goods, including raw materials, components and semi-manufactures. This continues to be an increasing trend whereby industrial firms exploit outsourcing for those products and activities deemed to be; (1) performed better by other organizations therefore offering value improvement opportunities or (2) outside the company's core business [Dulmin, (2003)]. Yoon & Naadimuthu (1994) state that the strategic decision to 'make or buy' can often be the major determinant of profitability, making a significant contribution to the financial health of a company.

A recent report from AT Kearney (2004) states that industrial leaders are creating value and gaining competitive advantage through the use of supply markets by focusing on four key areas: 1) Innovation and growth; 2) Value Chain Optimization; 3) Advanced cost-management; (4) Risk management and supply continuity. From the areas offering opportunity for value creation, the wider focus of this current research is that of facilitating improved cost-management for sourcing applications given the 'practical-industrial' constraint of not always having the required degree of cost and financial breakdown data desired (Curran, 2010; Curran et al, 2010c). It is clear that the enhanced significance of the supply chain has made procurement a strategic function [Dubois (2003)] and cost management (Pugh et al, 2010a; Pugh et al, 2010b) and assessment a critical activity for aerospace companies [Ellram, (1996)]. Monozka and Morgan (2002) proposes that increased attention to cost management is a critical factor to the operational control and sustained improvement of the procurement function as it provides a quantifiable basis upon which to assess related activities. Fleming (2003) states that the objective when sourcing is to; "negotiate a contract type and price (or estimated cost and fee) that will result in reasonable contractor risk and provide the contractor with the greatest incentive for efficient and economical performance". The term 'cost' from the AICPA Inventory [Humphreys (1991)] is; "the amount, measured in money or cash expended, property transferred, capital stock issued, services performed, or liability incurred, in consideration of goods or services received". Cost and price are often used interchangeably as parts can be made internally or be externally sourced from the extended supply chain, as shown in Figure 1 [Chen, (2004]; consisting of the internal and external supply chains as depicted. In this sense, the price of the external supplier is equivalent to the cost of internal production, being integrated into some product that is delivered to a customer.

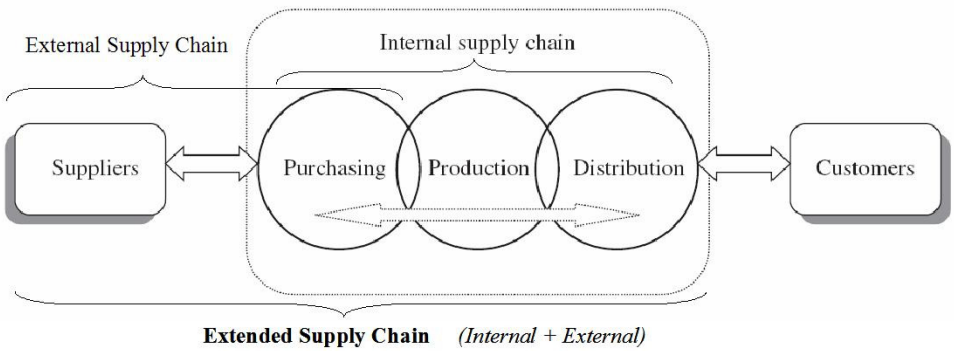


Fig. 1. Elements of extended supply chain (Adapted from Chen, 2004)

Parts that are externally sourced from world-class suppliers operating within a competitive market place often do not exhibit such a discrepancy between the actual manufacturing cost and supplier's selling price; the latter including a fair and reasonable mark-up, as illustrated by Scanlan (2004) in Figure 2. When however orders are placed with suppliers who operate towards the left hand side shown in the Figure for low-efficiency and an uncompetitive market, then a potentially excessive mark-up is likely. It is in the interest of the buyer to understand actual manufacturing cost as well as to have to the ability to assess the quality of potential suppliers before entering into business with them.

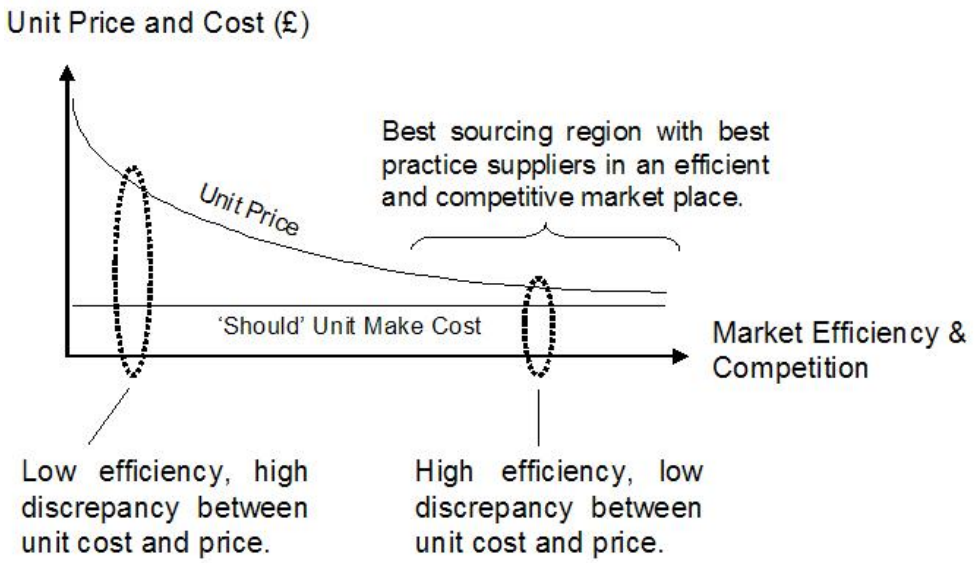


Fig. 2. Cost and price relationship with market efficiency (effects of volume removed), [Scanlan, (2004)].

Figure 3 highlights that unit price is influenced by a number of issues such as; (1) procurement strategy and requirements, (2) the technical requirements which directly influence manufacturing cost, (3) the actual cost basis on which the company operates, and (4) the external forces that determine an acceptable market price. All this is required to actively interface in the activity of negotiation: aimed at identifying mutually satisfactory terms for contract specification and price determination with potential suppliers. Specialist parts for which a buyer is dependent and has little internal knowledge of in terms of design and manufacture tend to result in supplier leverage and a potentially significant difference between cost and price. For standard parts a small difference between unit cost and price is expected. Understanding the costs involved in the production of a part with other specified requirements enables a procurement buyer

to physically negotiate and determine price and contract particulars with potential suppliers; based upon a platform of informed judgment.

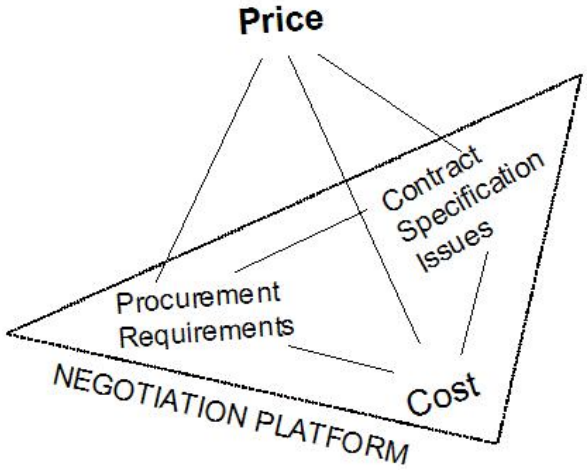


Fig. 3. Underlying components of Unit Price

3. State-of-the-art: Procurement cost analysis

Procurement transactions that occur between companies (buyer to supplier) are characterized by adding value up through the chain and consequent payments down the chain. The procurement function tends to be characterized as exploiting the supply chain in order to develop opportunities for increased profitability. It has been noted by [Hicks, (2000)] that this is envisaged through manipulation of the areas that directly effect asset and resource utilization, as well as profit margins, including: production decisions, outsourcing verses in-house management, supplier relationship type sought, and inventory turnover. The best practice principles that are identified as procedurally correct need to be supported by facilitating tools that provide quantitative measures of cost, time, risk, quality, etc. In particular cost modeling tools can easily be related to the following procurement needs as described in the literature [Fitzgerald, (2002); Handfield, (2002)]:

- Eliciting support from top management,
- Understanding cost drivers in appropriate detail,
- Make / buy opportunity analysis for improved value identification,
- Integrating and modeling the supply chain,
- Developing cooperative supplier relations,
- Measuring the performance of suppliers, systems, and employees,
- Delivering and establishing a culture of continuous improvement,
- Facilitating a cross-functional approach linked through cost,
- Managing and reducing costs across the whole operational structure,
- Developing integrated data management systems and;
- Justifying investment in procurement/supply tooling and management.

The challenge with developing supporting technologies is that of making them as widely applicable as possible and thus providing scope for system integration and cross-functional knowledge extraction. An important factor to consider when discussing best practices in procurement however is that no two companies are exactly alike, and as a result there is no simple generic approach to best practice policy [Handflield, (2002)]. Best practices often depend on people, suppliers, processes, or other operational elements that are specific to a certain situation [Handfield (2002), Fitzgerald (2002). Specifically considering cost analysis; Ellram (1996) states; "there are many cost management tools and techniques and they continue to proliferate. Thus, it is difficult to determine which type of analysis should be used in a given situation, and time pressure may inhibit the purchaser selecting the right tool". Consequently, it is proposed that a methodology be developed to help procurement operators determine what kind of cost assessment technique should be applied to given purchase situations by exploring the following issues:

- How should purchased items be classified into a framework, so that standard procedures can be developed for the analysis of items that fit into certain classifications?
- What cost analysis techniques best support each classification in the framework?
- What cost analysis techniques are more strategic in nature, and can really help purchasing add value to the organization?

The Genetic Causal Approach to cost modelling (Curran et al, 2004) adapted in the presented work addresses these issues directly through the classification into 'genetic part categories' implied in Questions 1 & 2, and through the use of causal models to address Questions 2 & 3.

Probert (1996) recommended the use of sophisticated techniques which offer greater accuracy to those classes of parts that are deemed to be of 'high-importance' and conversely simpler techniques to parts which belong to groups that are thought to be of lesser importance. Fleming noted that purchases may be: (1) big and others small in terms of both quantity and value, (2) some complex whilst others routine, (3) some high risk and others with perhaps no attached risk at all, (4) some requiring a lengthy contract whilst others needing only a short time commitment between the buyer and seller. As procurement needs are different for different purchases, many researchers [Flemming, (2003); Probert, (1996); Ellram, (1996)] recommend categorizing procurements into broad but distinct families before conducting any cost analysis. The old adage of 'not putting all one's eggs in the same basket is known as portfolio theory, which dates back to financial investment analysis in the 1950s [Markowitz, (1952); Olsen, (1997); Kulmala, (2004)]. This in fact can help management to focus more thoroughly on problems or issues specific to each category of procured part [Flemming, (2003)]. Following from this it is thought that optimal analysis approaches may then be identified for application to each particular grouping [Ellram, 1996]. In a similar fashion to that of Flemming (2003) it is acknowledged that before purchasers can choose the right analysis tool for a particular situation, they must understand the nature of the buy (which considers: scale, complexity, duration, contract type, dependency/risk, etc.) and the type of the supplier relationship sought. Ellram (2002) recognizes that this can range potentially from a loose agreement to a strategic alliance which importantly affects the availability of data as well as how much time or additional resources the organization is perhaps willing to devote to both supplier and cost analysis.

Classifying Suppliers / Purchases for Cost Analysis

Nature of Buy	Arm's-length	Strategic Alliance
Ongoing	Items purchased in large quantity that are made to stock with many available sources Items available on commodity exchanges	Items important to distinctive competency Items important to the future success of organisation
	Low Impact	**Critical Projects**
One-time	Most specialised Low price, repetitive buys	Critical project Long-term capital investments

Type of Relationship Sought With Supplier

Fig. 4. Classifying Suppliers / Purchases for Cost Analysis, [Ellram, (1996)].

Figure 4 provides a matrix of buying situations consisting of varying; types of buy and types of supplier relationship sought. Purchases are classified as *low impact, leverage, strategic and critical,* in terms of their cost and impact on the organization and relationship potential. Ellram (1996) acknowledges that, the type of cost analysis techniques used should support the relative importance of the item being purchased, as well as the type of supplier relationship that the organization currently has or desires. Following from Figure 4, Figure 5 highlights potential cost analysis techniques to be used in each of the buying-type situations identified.

Cost Analysis Techniques

Nature of Buy	Arm's-length	Strategic Alliance
Ongoing or Major Impact	**Leverage** **'Cost Analysis Focus'** Estimate cost relationships Value Analysis Analysis of supplier cost breakdowns Cost estimate / 'should cost' Industry analysis Total cost modeling	**Strategic** **'Continuous Improvement Focus'** Open books Target Cost Analysis Competitive Assessment / teardowns Total cost modeling of the supply chain
One-time or Limited Impact	**Low Impact** **'Price Analysis Focus'** Competitive bids Comparison price list / catalogues Comparison to established market Price Indexes Comparison to similar purchases	**Critical Projects** **'Life Cycle Cost Focus'** Total cost modeling and life-cycle costing

Type of Relationship Sought With Supplier

Fig. 5. Cost Analysis Techniques applicable for various types of supplier relationship and types of buy situations, [Ellram, (1996)].

Figure 5 implies that relatively simple analysis techniques are recommended for low impact purchases which focus primarily on *analyzing price*, where competitive bidding is viewed as the most common basic method of analysis. Moving from *low-impact* to *leverage items* it can be seen that greater attention is given to the analysis of *cost* rather than *price* in supplier cost breakdowns. Price analysis is simpler and faster than cost analysis. The simpler price analysis may be satisfactory for *low-impact* items however cost component understanding is desirable for *high-impact* parts. Even though cost analysis requires more processing time to practically employ; it generates a greater breakdown of cost information over that of price analysis and is therefore better able to support informed 'fair-price' negotiation. The technique involving the use of cost estimating relations is similar to that of the price analysis approach [Ellram, (1996)] of comparing similar purchases at price or sub-component cost levels. 'Should-costs' estimates involve attempting to independently construct the current or potential suppliers' product cost structure. Value analysis is a methodology which compares the function of an item or the service it performs to cost, in an attempt to find the best value alternative (Curran et al, 2006; Curran et al, 2010a, Curran, 2010; Curran et al, 2010c) and identify which quality or features that are causing cost but are not required or at least less desirable. Ultimately, total cost modeling or life cycle cost analysis (Curran et al 2003; Curran et al, 2007a; Curran et al, 2007b); goes beyond the focus upon suppliers' cost structures and looks specifically at; "the cost of doing business with a particular supplier for a particular item over the life of that item" [Ellram, (1996)].

The Society of Cost Estimating and Analysis (SCEA) state that cost estimating is: "the art of approximating the probable worth or cost of an activity based on information available at the time" [ISPA (1999)]. The main function of cost estimation is the provision of independent, objective, accurate and reliable capital and cost operating assessments that can be used for investment funding and project control decisions. In particular, accurate cost estimation is important for cost control, successful bidding for jobs and maintaining a competitive position within the marketplace [Ben-Arieh, (2000)]. There are two main approaches towards cost estimation: cost estimation based on past experience variant [Curran, (2004)] and generative cost estimation [Weustink, (2000)]. We can refer to *generative* or *compilational costing* as an approach which seeks to aggregate the various constituent cost elements identified for a given exercise whereas in variant or relational costing, comparative relation of product defining parameters is adopted in order to target/interpret causal reasons for cost differences between similar items, as highlighted in the Genetic Causal Approach Curran (2004). According to Humphreys (1991), variant (analogy) estimating involves identifying a similar part cost and then using this actual cost as a basis for the estimate of the new part. Generative estimating methods can be further divided into explicit (rule-based) cost estimating, Rough-Order-Magnitude (ROM) estimating, parametric and feature based cost estimating as well as detailed estimating potentially using Activity Based Costing (ABC) [La Londe, 1999], all of which are often based upon past experience. ROM or ratio estimating is a factor based technique which is used to arrive at a preliminary cost estimate inexpensively and quickly [Humphreys (1991)]. It is based upon the application of a ratio determined factor, from a previous contract, to a particular variable in order to calculate the value of a second. Parametric estimating is a technique that uses validated Cost Estimating Relationships (CERs) to estimate cost. Parametric cost models [Collopy, 2001] statistically estimate part cost based on the correlation between historical cost data and part properties which are considered

to be related to cost. Parametric models can use a small number of independent variables or in the case of feature based modeling, which is more generative in nature; any number of variables can be used to adequately describe the required detail present in an item. As discussed earlier, Activity Based Costing [La Londe, (1999); Mileham, (1993); Esawi, (2003)] is an accounting practice which specifically aims to identify the activities of an organization and the associated cost of each, using which activity costs are then allocated to cost objects.

Approaches involving the use of Knowledge Based Engineering [Curran et al 2010b; Verhagen et al 2011], artificial intelligence such as fuzzy logic and neural nets [Rush and Roy, (2000); Villareal et al, (1992)] are rapidly developing which mimic the human thought process. Using neural nets for costing involves the training of a computer programme given product-related attributes to cost. A number of researchers are investigating the use of neural nets for cost estimating purposes [Smith and Mason, (1997); Bode, (1998); Cavalieri, (2004); Idri, (2002); Wang, (2005)]. A neural net [Rush and Roy, (2000)] learns which product attributes most influence the associated cost and then approximates the functional relationship between the attribute values and cost during the training. Consequently, when supplied with product attributes describing new parts, the neural net selects the appropriate relationship function and generates the required cost estimate. Neural networks are entirely data driven models which through training iteratively transition from a random state to a final model. Brinke (2002) identifies that both neural nets and regression analysis can be used to determine cost functions based on parametric analysis; whereby parametric analysis is becoming an increasingly employed tool in industry for cost estimating purposes, e.g. SEER software. Both techniques use statistical curve fitting procedures however neural nets do not depend on assumptions about functional form, probability distribution or smoothness and have been proven to universal 'approximators' [Funahashi, (1989); Hornik, (1989)].

The advantages and disadvantages associated with regression analysis and neural nets have been identified by Bode (1998). Brinke (2002) states that when the cost parameters are known and the type of function is unknown or cannot be logically argued then neural networks are suitable to deduce cost functions, however that it is easier to quantify the quality of a result from regression analysis. Bode (1998) demonstrated that neural networks can produce better cost predictions than conventional regression costing methods if a number of conditions are adhered to. Smith and Mason (1997) indicate that in instances where an appropriate CER can be identified, regression models have significant advantages in terms of accuracy, variability, model creation and model examination. Considering the use of such techniques for cost estimating it is desirable that causal relationships are known between cost driving independent variables and cost. This subsequently strengthens one's case when attempting to enforce a cost reduction with a current supplier based upon non-disputable causal logic. Neural nets can sometimes be used to generate more accurate results than those from the use of regression however the challenge associated with the further diffusion and wider implementation of this methodology according to Cavalieri (2004) is that of making the approach more transparent to the analyst and developing tools which reproduce in a comprehensible, easy to use fashion the behaviour of the network. Finally with respect to fuzzy logic, a fuzzy expert system is one that uses a collection of fuzzy membership functions and rules to deal quantitatively with imprecision and uncertainty, and researchers [Gerla, (2001); Kishk, (1999); Ting, (1999); Klir, (1996); Mamdani, (1981)] agree that the major contribution

of fuzzy set theory is the inherent capability of representing vague knowledge. Roy (2003) however states that fuzzy logic applications within the field of cost estimating have not been well established, well researched or published. The impact of uncertainty and sensitivity within cost modelling has been also well researched within aerospace to show that Monte Carlo techniques can be employed to increase the robustness of the analysis (Curran et al, 2009).

It should be noted that each of the estimating methods to varying degrees can be employed in either a 'top-down' or 'bottom-up' fashion. 'Top-down' involves the formulation of an overall estimate to represent the completed project which may then be broken down into subcomponents of cost as required. In contrast, 'bottom-up' estimating [Ting, (1999)] generates sublevel and component costs first which may then be aggregated in order to produce an overall estimate. Elements of each of these methods are more or less applicable at various stages of the product life cycle. Further reviews of these methods are provided by Curran (2004), Roy (2003) and Stewart (1995).

4. Methodology: Cost CENTRE-ing

The purpose of incorporating improved estimating methodologies within Procurement is essentially to provide additional information against which sourcing issues may be more readily considered. The research method presented in this Section gives attention to identifying opportunities for cost reduction from currently outsourced parts based upon unjustifiable cost or price variances amongst similar parts. Control follows estimate generation and usually involves the comparison with actual and other estimates for the purpose of identifying such variances and then attempting to understand their causes with the view to bringing cost to a desired baseline. Three types of cost variance are of interest when comparing cost information of similar items including: 1) comparison of actual cost to actual cost, or indeed lower level actual cost components, 2) comparison of actual costs to cost estimates, at any level of aggregation, and 3) comparison of an estimate to another estimate developed from a different approach.

Figure 6 presents a synthesis of procurement best-practice in unit cost/price analysis, with reference to the authors experience and the literature review in Section 3. It is reflective of the latest cost management research in the area (Pugh et al, 2010a; Pugh et al, 2010b) and involves tailoring cost analysis to given types of purchase situation.

It can be seen that the key elements identified are the roles of Classification, Data mining, Cost/Price Analysis, Supplier Selection and Cost Control. Consequently, the presented work was therefore directed towards the development of a modelling methodology and process that would support the Cost/Price Analysis stage in particular. The resulting methodology was termed (Genetic Causal) Cost CENTRE-ing, as the word 'CENTRE' is an anagram of the 6 key process steps to followed in implementing the methodology. The Genetic Causal basis (Curran et al, 2004) of the methodology refers the decomposition of procurement items into 'genetic' families of similar parts based either on part material, form, function or manufacturing process, so that then, historical costing data can be used to develop 'causal' relations to estimate the part-cost of any instance of an item from that genetic family.

The causality of the costing algorithms is a very significant issue so that the equations are robust and dependable, with the dependant variable as cost being a function of independent variables relating to the part definition, such as part, process or function

information, rather than purely statistical in nature; as we find often in traditional parametric costing (see Curran et al, 2004). In addition, another requirement was that the Cost CENTRE-ing process could provide an agile method for up-to-date analysis, estimation, control and reduction of procurement costs and so it was decided at the outset that it should be able to easily incorporate new cost data and part information in order to upgrade the costing algorithms in an automated manner. As illustrated in Figure 7, the method is broken down into six key steps: (1) Classification, (2) Encircling, (3) Normalization, (4) Trending, (5) Cost Reduction Identification and (6) Enforcement. Steps 1 to 4 involve knowledge discovery incorporating data mining, statistical study (e.g. for variable selection, significance and hypothesis testing, trending and optimization) with scope for sensitivity and likelihood testing, which brings in concepts central to probability.

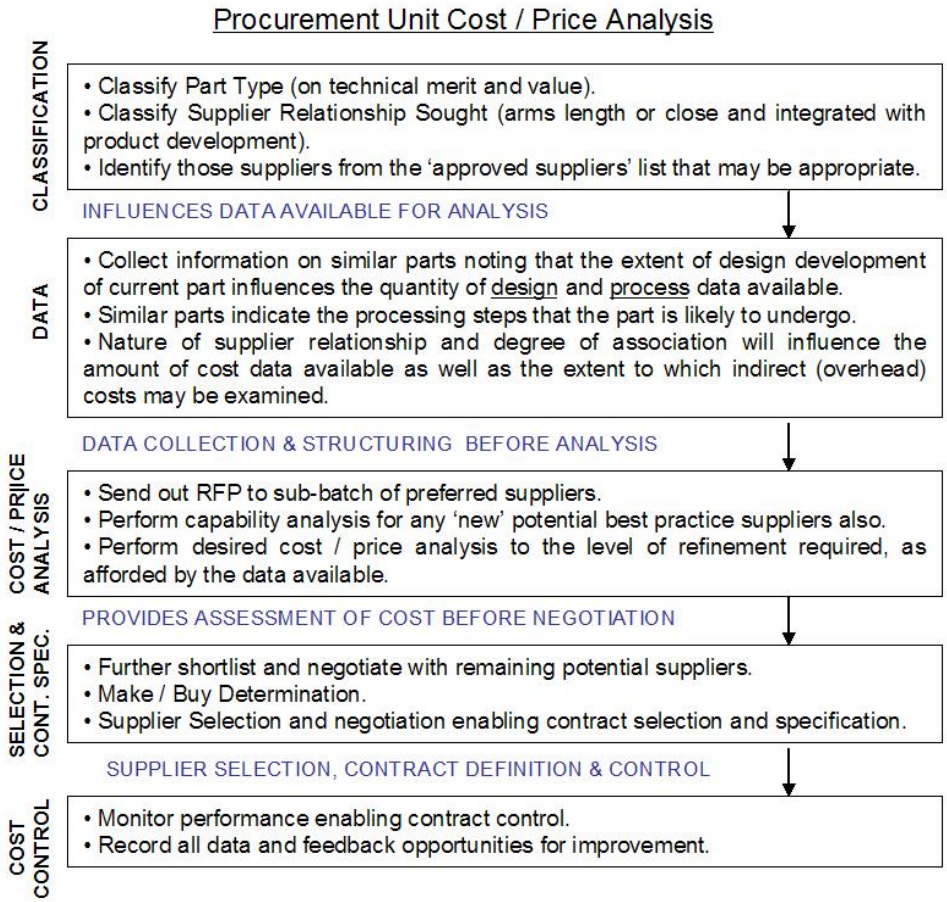


Fig. 6. Procurement best practice in unit cost or price analysis

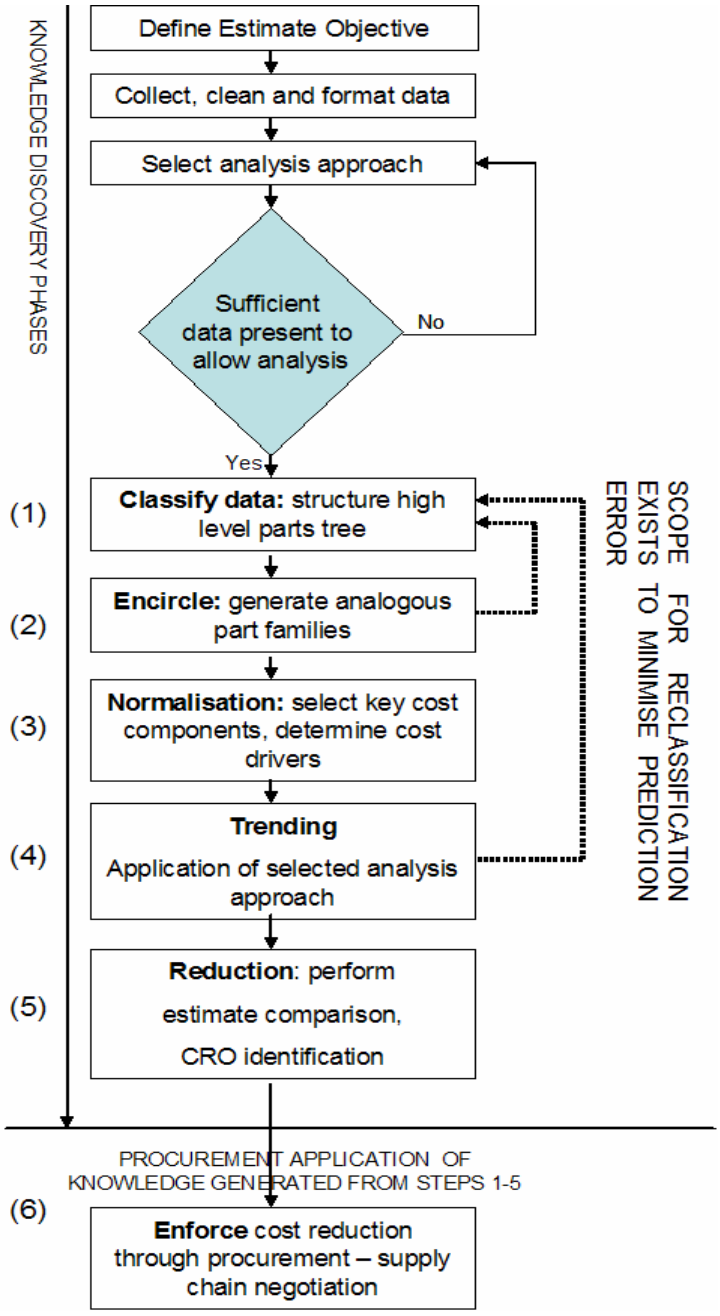


Fig. 7. The Cost CENTRE-ing methodology

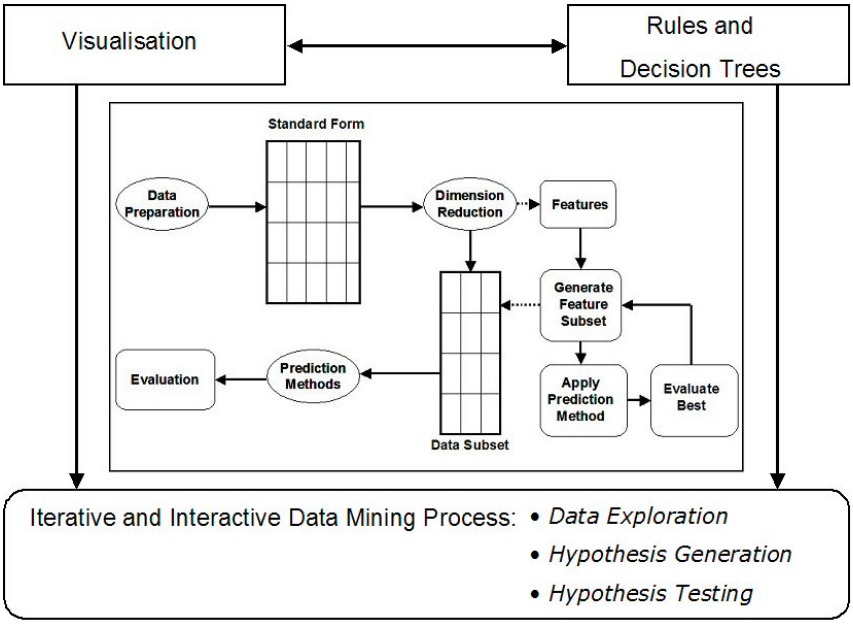


Fig. 8. A hybrid approach to data mining

The steps associated with Cost CENTRE-ing are further expanded below and map equally well to the requirements presented through Figure 6, starting with Classification and finishing with the application to Cost Control:

(1) Classification: as a key aspect of the methodology and was implemented to define families of parts. There is an obvious trade-off in terms of increasing the complexity through the number of Cost Estimating Relationships (CERs) embodied in the eventual methodology. Classification was developed according to the following descriptors as taken from a part's Bill of Material: Procurement Part Type, Aircraft Type, Sub-Level Contract, Process, Material Form and Material.

(2) Encircling: involves analysis of a data set's principal components and allows clusters to be identified in order to improve grouping refinement and proceeds as follows: Machine Type, Part Size and Batch Size. Figure 8 highlights a hybrid data mining approach involving data exploration, standardization, and visualization, reduction with subset generation as well as statistical testing and iterative evaluation (Weiss 1988, Fayyad 2002). Considering this, the process of pattern matching that is being used in the presented approach to data grouping is analogous to having degrees of freedom in a formal statistical test.

(3) Normalization: After surveying the more advanced methods being developed, such as Neural Networks and fuzzy logic etc, it was decided that Multiple Linear Regression would be used to model the link between part attributes, as independent variables, and unit cost, as the dependant variable (Watson et al, 2006). This requires that the data be normalized in order to distil out the key cost drivers to be used in the formulation of parametric relations. There is a trade-off here in terms of the number of drivers, which may be used to optimize a given result and the corresponding actual improvement considering the additional processing time required to generate the result.

(4) Trending: also considering knowledge capture and formalization, this step allows the appropriate trend which describes the mapping relationship of cost to the independent variables to be selected. The most appropriate trend to use may change from case to case although what is common is the means by which the goodness of fit of a relationship may be measured (through the R^2 value that describes the degree of statistical fitting), with the trend that best minimizes random variance or error being selected in each case.
(5) Reduction and (6) Enforcement: these steps are linked to Procurement's use of the relationships and trends developed at this point in the process. 'Reduction' entails application and comparison of prediction trends to current 'actuals' or to results developed by other estimating techniques for the purpose of identifying Opportunities for Cost Reduction either by direct total cost comparison at part level or sub-cost components (e.g. Make, Material, Treatments, etc.). Once identified, the Procurement function must then decide upon the appropriate course of action to be taken in order to attain reductions through 'Enforcement'.

5. Results and validation

The effectiveness of the Cost CENTRE-ing methodology and process was validated on three separate studies (including four specific cases in total) in collaboration with the procurement function at Bombardier Aerospace Belfast. Three studies of a different nature were chosen to represent the range of parts procured within aerospace. This included: 1) a machined parts example with a data set of 850 'Outside Production' aircraft items on one contract and another data set of 117 parts from a different aircraft contract, 2) a vendor-specialized 'systems' part in the form of Thermal Anti-Icing Valves of which there was a typically small data set of 6, and 3) a more common fastening part in the form of a spigot for which there was a data set of 201. The results from these validation studies are presented in the following Sections 5.1 through 5.3, where the methodology is presented according to the six key steps of: (1) Classification, (2) Encircling, (3) Normalization, (4) Trending, (5) Cost Reduction Identification and (6) Enforcement. The machining case study was just one of many carried out on the whole part base of some 7,000 machined parts at Bombardier (Watson et al, 2006).

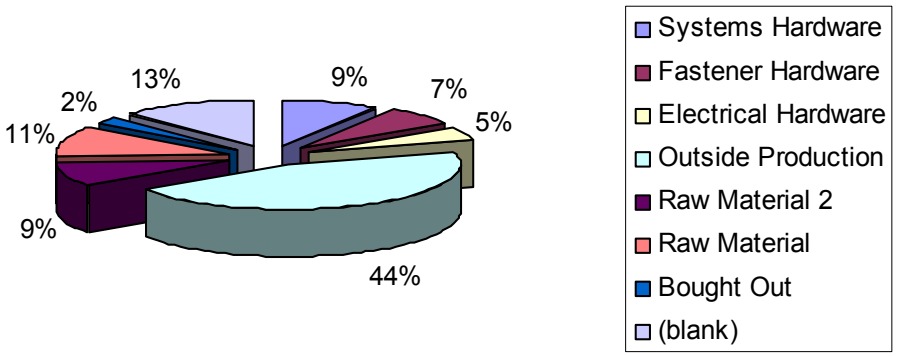


Fig. 9. An example of the procurement spend breakdown for Bombardier Aerospace Belfast

5.1 Validation study 1: Outside-production machined aerospace parts

(1) Classification: Figure 9 presents the general breakdown of procurement spend at Bombardier Aerospace Belfast while Figure 10 further disaggregates the spend on 'Outside Production' parts. Consequently, one can see the opportunity to define and develop families of parts of a similar in nature.

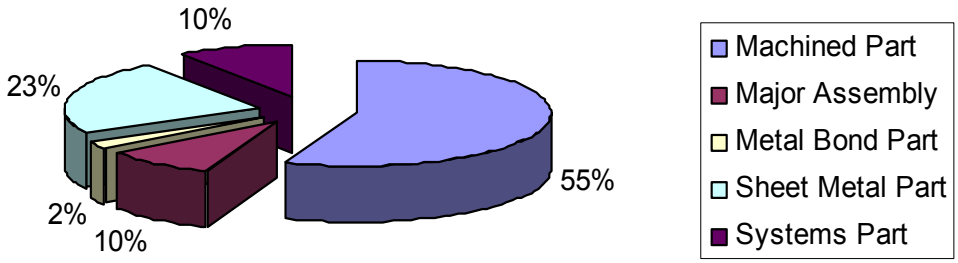


Fig. 10. The breakdown of outside product parts for Bombardier Aerospace Belfast

(2) Encircling: In Figures 9 and 10, it can be seen that the parts have been categorized in order to group parts with a increased degree of commonality. Primarily, at this level of distinction it is paramount to choose associated part attributes that have been identified as driving manufacturing cost, thereby following the principle of causality. For example, weight might be used well as an independent variable for material cost but is less relevant to unit cost (when in aerospace it typically costs money to take weight out of a structure) while other independent variable may be less obvious but still of a causal nature such as using direct as part count as an assembly cost driver. It is also important to choose attributes that are already defined at whatever stage of the product life that the model is to be utilized, and of course that these are also readily available. If the Cost CENTRE-ing implementation is fully coupled to design platforms (Curran et al, 2001; Curran et al, 2007a; Curran, 2010) it is then possible to impose a much greater level of definition, through actual part volume etc, which would increase the accuracy but also the operational complexity of the Model. However, this is more relevant to validation, improvements in the costing algorithms and cost reduction exercises while as procurement costing at the conceptual design phase does not have the design definition one would want for very accurate causal modelling of costs.

(3) Normalization: A simple initial causal parametric relation was generated from the data for machined parts using the Multiple Linear Regression facility within the MS Excel Data Analysis module. The detailed manual cost estimates of the machining times for 850 parts were used as the dependant variables while the readily available independent variables were all based on size attributes (thickness, length and breadth). In terms of driving the parametric relation, the size envelope is primarily linked to the material removal although the relation would be much improved with more detailed attribute data. Work is progressing in also linking part complexity, as driven by key design attributes of the part.

(4) Trending: Trending was carried out using Multiple Linear Regression, where machining time was the estimated time for a given component made from a billet of thickness T, length L and width W; according to three regression coefficients and a constant. It is interesting to also note that the regression in question had a 'Multiple R' value of 0.71, which can be interpreted as the mathematical formulation account for approximately 70% of the variation in the historical data. A Multiple R value of 0.8 would be preferable and could be feasible by improving the range of independent variables used to characterize the parts, e.g. through the additional normalization according to part size and design/machining complexity, as available. However, this machining case study was one of many carried out on the whole part base of some 30,000 parts at Bombardier (Watson et al, 2006).

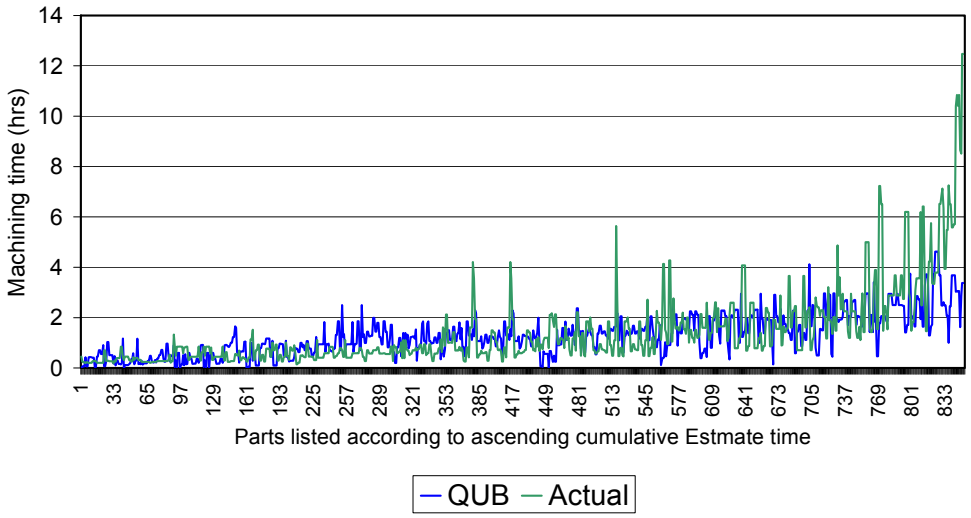


Fig. 11. Cost comparisons of 850 parts using 'actuals' (with more deviation) and the model

The resulting estimates for the 850 parts are presented in Figure 11 where the Cost CENTRE-ing 'QUB' estimate is compared against the actual times. However, the 'Actuals' were not directly available from the suppliers due to the sensitivity of the information and had to be derived from a detailed estimate of the parts using the actual supply price and an averaged machining rate. Anywhere on Figure 11 that there is significant disparity between the two characteristics highlights those parts that require further investigation for potential cost reduction, as presented in the following Section.

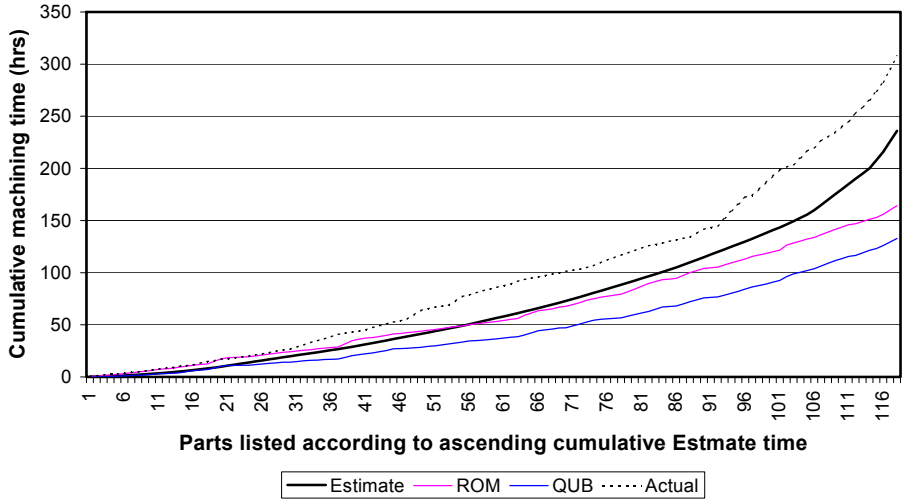


Fig. 12. A detailed comparison for part costs with 'Actuals', the manual ROM and the 'QUB' model values and the current detailed manual estimates (the solid line)

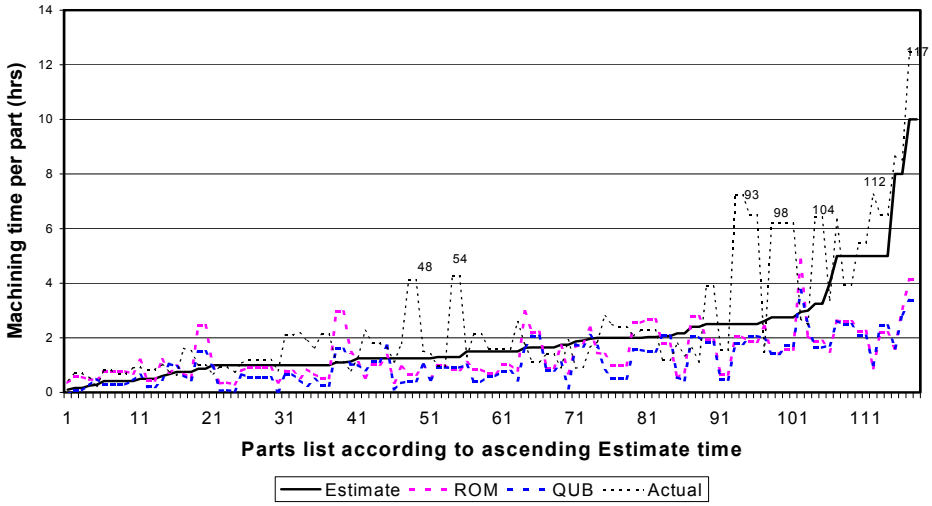


Fig. 13. A comparison of the cumulative cycle times of the parts detailed in Figure 12

(5-6) Reduction/Enforcement: The Cost CENTRE-ing model developed for machined parts was then applied to older 2nd contract where it was believed there might be greater opportunity for cost reduction. Figure 12 presents a direct comparison between all cycle time values for the 117 listed parts associated with the aircraft contract. Four types of estimated values are presented, including: the detailed manual estimate, the Rough Order of Magnitude (ROM)

estimate from an in-house parametric model, the Cost CENTRE-ing 'QUB' estimate and the derived 'Actuals' estimate. It can be seen that a significant number of 'Actuals' are extremely different. Figure 13 provides a cumulative comparison for each of the estimate types in which the cumulative differentials again imply that the 'Actuals' are too high. Consequently, a number of these parts were identified and the differentials calculated to estimate the potential savings if the current suppliers were to reduce their price to the appropriate should cost or else via supplier sourcing. For this case, potential savings of £100,000 were generated through *(6) Enforcement*.

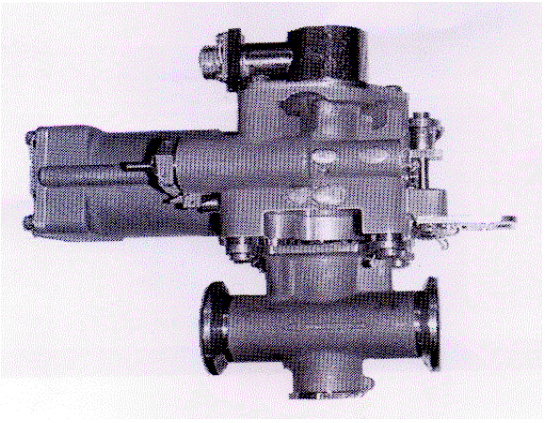


Fig. 14. An example of a typical Off-The-Shelf item used as a case study: an anti-icing valve

5.2 Validation study 2: Off-the-shelf systems items – Aircraft engine anti-icing valves

(1-2) Classification/Encircling: This study considers the procurement of Thermal Anti Icing (TAI) valves as a general off-the-shelf item, relating to the system hardware category in Figure 9 and shown in Figure 14. Ice protection relates to the prevention and removal of ice accumulation (anti-icing and de-icing respectively) on either a wing leading edge or more typically on the Nacelle inlet to an aircraft engine. However, there are a range of pneumatic and electrical systems that supply the required heat from the engine bleed hot air for: wing anti-icing; engine nose cowls and inlets and centre engine inlet duct; the upper VHF antenna; fuel filter de-icing (more under power plant). The case study was undertaken with a view towards determining why there is a cost variation between those TIA valves currently being sourced so that this improved understanding would lead to a better 'Should Cost' estimate; a term commonly used for a target cost or price. As such, the valve was classified within the vendor item group with the valves identified as an encircled grouping of parts with an obvious commonality.

(3) Normalization: The normalization procedure was implemented as set out previously in order to deter-mine the cost drivers that differentiate the cost of one instance of the encircled group from another. It was found that the cost of a valve is dependent for example upon; casing and seal materials, performance specifications, testing and scale of production or order quantities. The valves being examined were particularly challenging as they are vendor-supplied items with little information available over that of the original operational specifications and the actual buying price. Naturally, the implication is that one is dealing

with price as the dependant variable rather than cost, which means that it is less feasible to look for a causal linkage between price and item parameters. Notwithstanding, the more fairly an item is priced the more likely it is that a trend can be established with statistical significance. The initial process followed was that of extracting from the source documents all operational specifications and requirements with a view towards removing any common characteristics and then analyzing the remaining variables, to ascertain their influence on the unit price. It was recognized that there are many attributes that contribute towards any item's overall cost, as well as other environmental factors that affect the part's price, but in such a case with very little or no knowledge of the cost breakdown, basic relationships for those variables considered to be the major performance/functionality cost drivers can be used.

(4) Trending: As previously, the trending relied on Multiple Linear Regression as the means of relating the available cost drivers to the measure of cost, or more accurately price in this case. Figure 15 plots some of the regression findings that were carried out to investigate the relations between performance drivers and the Purchase Order value per part. Some of these initial relations are of use in terms of a Rough Order Magnitude (ROM) estimate and also provide the rationale and negotiating leverage for cost reduction dealt with in the next Section. It should be noted that there is often interaction between such performance parameters so that it is important to use more than one independent variable in calculating a robust estimate.

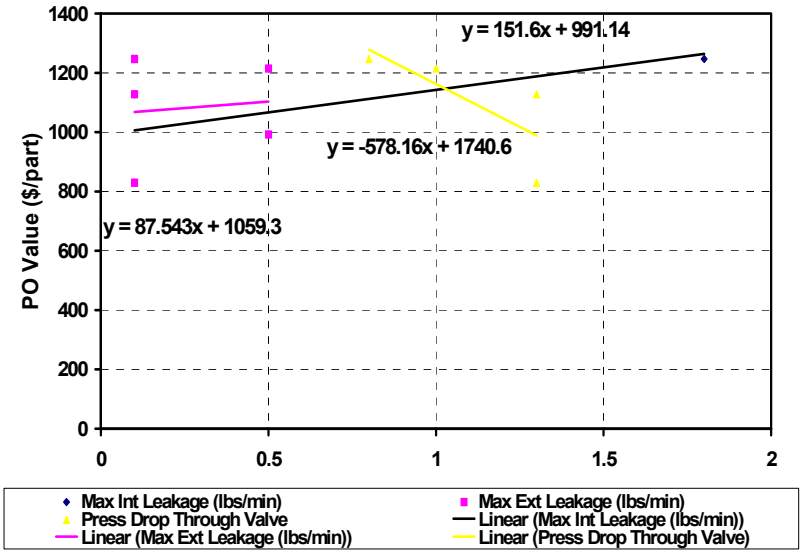


Fig. 15. Indicative cost benefit modeling with regards to performance specification

(5-6) Reduction/Enforcement: It was found from the studies that there was a deviation of almost 50% in the cost of the procurement of these various valves but very little discernable difference in the performance specifications. A more influential parameter

was the order quantity although again there were anomalies in the trending. Ultimately, however, these anomalies were then exploited as the negotiating rationale for cost reduction as part of the *Enforcement* step. Consequently, for these procured parts that are very difficult to cost the Cost CENTRE-ing approach as been used to identify the more likely opportunities for cost reduction due to disparity in the estimates, rather than trying to accurately cost a quite bespoke off-the-self system item, of which there are many on an aircraft.

5.3 Validation study 3: General aerospace supply items - Spigots

(1) Classification: In total Bombardier Aerospace Shorts Methods Procurement currently outsource in region of 34,000 parts across 618 suppliers for use within aircraft sub-assembly build contracts. Of those parts, the overall part list was first classified according to commodity code, for example, 'Machinings' accounting for some 7000 parts. This study focused on what is termed 'General Supply' items, or more minor parts that are used in very large quantities and are directly used typically in fastening and assembly.

(2) Encircling: In encircling a particular cluster of General Supply items for analysis those parts used in engine Nacelle manufacture were considered, reducing the part count down to 840. Of these 840, a further filtering step was carried out to generate a list of those items, which are considered to be similar in nature to a number of other parts within the grouping. This included the main characteristics of a part being present in each item contained within the 'Similar to' part set. The parts list of 840 parts was condensed to a list of 'Similar-to' part sets which contained in total a shortlist of 201 parts. In this instance the encircling was driven more by product orientation and function-role approach, rather than primarily for part family, such as for valves; fuselage panels, Nosecowls etc. One such 'Similar-to' part set related to a particular style of Spigot, which is a member of the 'Round Bar & Tube' part family, as shown in Figure 16.

(3) Normalisation: The individual General Supply items/parts are normalized according to make-cost, material cost and treatments. According to the 'Should Cost' Approach, parts with similar attributes in terms of material, geometry, manufacturing and treatments requirements should approximately have close make, material & treatment costs.

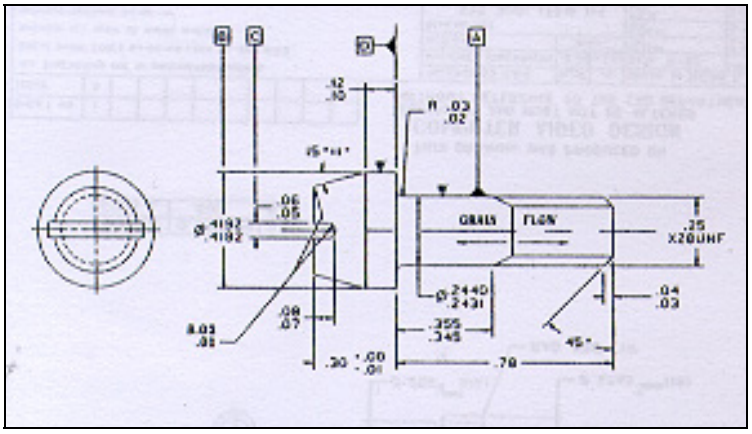


Fig. 16. A example of a General Supply item: a spigot

(4) Trending: Again the procurement information is more price oriented and therefore rather than direct modeling, the lowest component cost for each within the part set is then considered to be an initial baseline value to which the others should be brought in line with, remembering again that the Should Cost target is an estimate of a unit price that accurately reflects reasonably achievable contractor economy and efficiency.
(5) Reduction: For each part set, the opportunities for cost reduction are identified by calculating the differential between each parts' current Make Cost, Treatments Cost & Materials Cost for each of these parts. However, in addition the Should Costs for these Costing components (within each part set) needs to also factor in the quantity of parts per delivery batch, the rate of usage per year and the expected duration of build contracts to which the parts are being used [Marquez and Blanchar, (2004)]. This gives the overall potential for savings for each 'Similar to' Part set.
(6) Enforcement: The projected potential savings across six contracts currently in development with Bombardier Aerospace Belfast are shown below in Figure 17 for the spigots. It is interesting to note that there is a greater potential for savings in three particular projects. This can be accounted for by the fact that Contracts D, E & F had been focused on for some time with the application of the Should Cost philosophy, hence less opportunity for cost reduction. If the other parts in the set have been sourced via the one supplier then procurement contacts the supplier to discuss the cost drivers for the set of parts to establish why each are not currently being supplied at Should Cost and ultimately look to renegotiate the part costs. If sourced via a few different suppliers then this process is more complicated but in essence the same as the cost drivers will indicate the true unit cost for an item so that through mutually beneficial discussion (supply and demand) it should be possible to bring the items to an agreed Should Cost. It should be noted that an activity that requires and develops increased understanding of the cost drivers is beneficial for both the supplier and customer and *Enforcement* is not carried out in order 'to eat unfairly into supplier profit margins' but to establish a profitable and sustainable relationship between the two based upon enhanced efficiency and best practice driven initiatives.

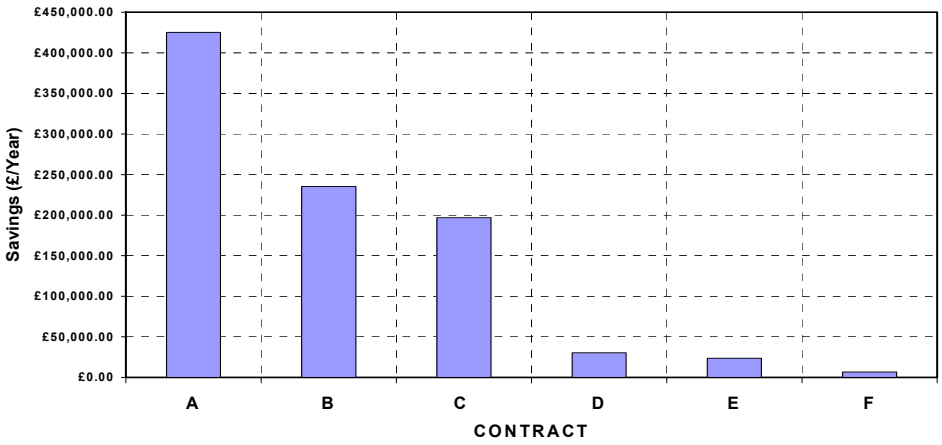


Fig. 17. Enforced savings for the spigot General Supply case study across a number of contracts

6. Discussion

In terms of key insights and contribution, Genetic Causal Cost CENTRE-ing utilizes part or product attribute information to build families of causal cost estimating relations that are based on rationale, rather than simply using market forces in procurement cost control and the traditional practice of buyer-purchasing based on part numbers without any insight into what is being purchased. Furthermore, the methodology has been applied to categorize very large quantities of parts in order to provide an agile and responsive tool for supply chain cost management. This provides the buyer with a stronger rationale in negotiating price reductions, ideally to be used in conjunction with some gaming theory for example and the more traditional assessment of market forces.

The application and relevance to real-world industrial situations has been validated in collaboration with Bombardier Aerospace Belfast and is synthesized into the model presented in Figure 7, the application of which was described in detail in Section 5. Essentially, this is encapsulated in the six procedural steps of: item Classification; data Encircling; cost driver Normalization; parameter Trending; cost Reduction identification; negotiated Enforcement; termed Cost CENTRE-ing. Following the Genetic Causal approach, this entails the categorization of part and product families stored in large data banks of cost information, the generation of associated causal 'Should Cost' estimation algorithms, and the application to current procurement operations through price negotiation. A tool was developed and is being used by Bombardier Aerospace Belfast which has automated the rapid formulation of the cost estimating functions, based on the most up to date data available, so that the buyer can select the generic type of part to be procured and then generate a 'Should Cost' range with associated limits of confidence relative to an ideal cost estimate.

It is envisaged that practitioners will extend the work to improve the gathering of more extensive data, including quantitative and qualitative knowledge capture, and that this would entail more effective integration within the companies' Design and Manufacturing functions; in collecting and utilizing key part and product information. Ultimately, the modeling capability could also explicitly facilitate the Design to Cost procedure to help drive the design process towards more effective design solutions that exploit key supply chain and procurement knowledge. However, in terms of a pure procurement tool, it is envisaged that the application can be developed and exploited more fully as a web-based technology that is more responsive in the identification and control of Lean suppliers who operate within an optimal cost basis.

7. Conclusion

This Chapter presents an agile cost estimating methodology to be deployed in a procurement operations tool for enabling more cost effective procurement control and cost reduction. The method is agile in being able to easily include the latest market data to generate its own costing algorithms that are established using the Genetic Casual Cost CENTRE-ing approach: item Classification; data Encircling; cost driver Normalization; parameter Trending; cost Reduction identification; negotiated Enforcement. It is shown that the Cost CENTRE-ing method provides an agile method for responsive cost analysis, estimation, control and reduction of procured aerospace parts. The methodology is based on the structuring of parts into product families and utilized both manufacturing and performance cost drivers to establish causal cost estimating relationships, according to the

Genetic Causal approach. Case studies have been presented to test the generic relevance and validity of the method. A 'machined part' example representing out-side production used both specific design and cost data while a General Supply spigot example used analogy applied to comparison of sub-cost components. An off-the-shelf Thermal Anti-Icing valve study relied exclusively on broad contract based information (not specific to the part) with purchase order value as the dependent variable and performance specifications as the independent variable. IN particular the latter was shown to be inherently difficult due to differing suppliers using alternative cost stack up and allocation policies, as well as profit margins, which makes it difficult to identify causal drivers that affect the cost differentials. However, once again the Genetic Causal method forces the use of causal cost drivers (performance related in the latter study) that can be clustered according to the cost family under consideration, while being facilitated by the Cost CENTRE-ing process. The Cost CENTRE-ing method uses 'comparison' in early data grouping and refinement but is also the basis of normalization and trend selection. It does this by selecting those drivers with the smallest measure of random error and which can be linked causally to cost.
The proposed methodology was applied to the three validation studies to show that it is effective in a wide range of applications (generic), has been used to significantly reduce the cost of supplied items (accurate), and is being adopted by a leading aerospace manufacturer (relevant). It is concluded that the proposed Genetic Causal Cost CENTRE-ing methodology exhibits all the above because it is based on an improved understanding of procurement operations and supply chain costing; thereby contributing to the body of knowledge in terms of process understanding; the importance of a causal relations in estimating; and identifying inheritance and family commonality in groups of products. It is envisaged that the application can be further developed into a web-based technology that is more responsive in the identification and control of Lean suppliers who operate within an optimal cost basis

8. Acknowledgements

The presented work was carried out in collaboration between Queens University Belfast and Bombardier Aerospace Belfast. The authors were involved in a range of cost modeling research projects within the Design, Manufacturing and Procurement domains, but the presented work was carried out as part of a Bombardier/DEL funded 'Cast Award' that resulted in the successful PhD work of Dr. Paul Watson, in association with that Department of Education and Learning (DEL), NI initiative. Within Bombardier Aerospace Belfast, acknowledgment and thanks must also be noted for the expert contributions of Mr. Neil Watson and Mr. Paddy Hawthorne, under the strategic direction of the Director of Procurement, Mr. Steven Cowan.

9. References

AT Kearney Report, (2004), "Creating value through strategic supply management – 2004 Assessment of Excellence in Procurement".

Ben-Arieh, D., (2000), "Cost estimation system for machined parts", International Journal of Production Research, Vol. 38, No. 17, pp. 4481 – 4494.

Brinke, E.T., (2002), "Costing Support and Cost Control in Manufacturing: A Cost Estimation Tool applied in the sheet metal domain", PhD Thesis, Printed by PrintPartners Ipskamp, Enschede, The Netherlands, ISBN 90-365-1726-5.

Cavalieri, S., Maccarrone, P. and Pinto, R., (2004), "Parametric vs. neural network models for the estimation of production costs: A case study in the automotive industry", International Journal of Production Economics 91, pp. 165 – 177.

Collopy, P.D., (2001), "Aerospace Manufacturing Cost Prediction from a Measure of Part Definition Information", Copyright © 2001 Society of Automotive Engineers, InC.

Chen, I.J., and Paulraj, A., (2004), "Understanding supply chain management: critical research and a theoretical framework", International Journal of Production Research, Vol. 42, No. 1, pp. 131 – 163.

Curran, R, Kundu, A., Raghunathan, & Eakin D (2001), "Costing Tools for Decision Making within Integrated Aerospace Design", Journal of Concurrent Engineering Research, 9(4), 327-338.

Curran, R, Kundu, AK, Raghunathan, S and R McFadden (2003). "Influence of Manufacturing Tolerance on Aircraft Direct Operating Cost (DOC)", Journal of Materials Processing Technology, Elsevier Science BV, ISSN; 0924-0136, 239-247.

Curran, R., Raghunathan, S. and Price, M., (2004), "Review of aerospace engineering cost modeling: The genetic causal approach", Progress in Aerospace Sciences 40, pp. 487 – 534.

Curran, R., and Watson, P., Cowan, S., Hawthorne, P., Watson, N., "Cost CENTRE-ing: A Procurement Methodology for Cost Reduction within Aerospace", Proceedings of Concurrent Engineering Conference, CE2004, Beijing, China, 2004.

Curran, R., Rothwell, A., and Castagne, S., "A Numerical Method for Cost-Weight Optimization of Stringer-Skin Panels", Journal of Aircraft, Vol. 43, No. 1, 2006, pp. 264-274.

Curran, R, Gomis, G, Castagne, S, Butterfield, J, Edgar, T, Higgins, C and C McKeever (2007a), "Integrated Digital Design for Manufacture for Reduced Life Cycle Cost", International Journal of Production Economics (IJPE), 27-40.

Curran, R, Castagne, S, Early, J, Price, M, Raghunathan, S, Butterfield J & A. Gibson (2007b) "Aircraft Cost Modelling using the Genetic Causal Technique within a Systems Engineering Approach", The Aeronautical Journal, Royal Aeronautical Society, 409-420.

Curran, R, Castagne, S, Rothwell, A, Price, M, Murphy, A. & Raghunathan, s. (2009). Uncertainty and sensitivity analysis in aircraft operating costs in structural design optimization. Journal of aircraft, 46(6), 2145-2155.

Curran, R. (2010), Value Driven Design: Value Integration, Encyclopedia of Aerospace Engineering, ISBN: 978-0-470-75440-5, Volume 6, Part 31, Chapter 322.

Curran, R, T. Abu-Kias , M.J.F Repco, Y.L.J. Sprengers, P.N.C.S van der Zwet and W. Beelearts, A (2010a), Value Operations Methodology for Value Driven Design: Medium Range Passenger Airliner Validation, Proceedings of AIAA General Meeting, Orlando.

Curran, R., Verhagen, W.J.C., Van Tooren, M.J.L. & Van der Laan, A. (2010b). A multidisciplinary implementation methodology for knowledge based engineering: KNOMAD. Expert Systems with Applications, Volume 37, Issue 11, November 2010, Pages 7336-7350

Curran, R., Butterfield, J., Jin, Y., & Burke, R., Collins, R. (2010c), Value Driven Manufacture: Digital Lean Manufacture, Encyclopedia of Aerospace Engineering, ISBN: 978-0-470-75440-5, Volume 6, Part 31, Chapter 313.

Dubois, A., (2003), "Strategic cost management across boundaries of firms", Journal of Industrial Marketing Management, 32, pp. 365 – 374.

Dulmin, R., and Mininno, V., (2003), "Supplier selection using a multi-criteria decision aid method", Journal of Purchasing and Supply Management 9, pp. 177 – 187.

Ellram, L.M., (1996), "A structured method for applying purchasing cost management tools", The Journal of Supply Chain Management – A Global Review of Purchasing and Supply, Volume 32, No.1, pp. 11 – 19.

Ellram, L.M., (2002), "Strategic Cost Management in the Supply Chain: A purchasing and Supply Management Perspective", Copyright © 2002 by CAPS Research.

Esawi, A.M.K., and Ashby, M.F.*, "Cost Estimates to Guide Pre-selection of Processes", Mechanical Engineering Department, The American University in Cairo, *Engineering Design Centre, Cambridge University Engineering Department, January 2003.

Fitzgerald, K.R., (2002), "Best Practices in Procurement", Ascet Volume 4, Ascet – Achieving Supply Chain Excellence Through Technology.

Fleming, Q., (2003), "Project procurement management: contracting, subcontracting, teaming", Fmc Publisher, ISBN: 0974391204.

Funahashi, K., (1989), "On the approximate realisation of continuous mappings by neural networks", Neural Networks, Vol. 2, pp. 183 – 192.

Gerla, G., "Fuzzy Logic, Mathematical Tools for Approximate Reasoning", Kluwer Academic Publishers, ISBN 0-7923-6941-6, 2001.

Handfield, R.B., Krause, D.R., Scannell, T.V., and Monczka, R.M., (2002), "Avoid the pitfalls in Supplier Management", Sloan Management Review, January 2000.

Hicks, C., McGovern, T., and Earl, C.F., (2000), "Supply chain management: A strategic issue in engineer to order manufacturing", International Journal of Production Economics, Volume 65, Issue 2, 20 April 2000, pp. 179-190.

Hornik, K., Stinchcombe, M., and White, H., (1989), "Multilayer feedforward networks are universal approximators", Neural Netwroks, Vol. 2, pp. 359 – 366.

Humphreys, K.K., (1991), "Jelen's Cost and Optimisation Engineering", Third Edition, McGraw Hill Publishers, ISBN 0-07-053646-5, pp. 538.

Idri, A., Khoshgoftaar, T.M. and Abran, A., (2002), "Can neural nets be easily interpreted in software cost estimation?", 2002 World Congress on Computational Intelligence, Honolulu, Hawaii, May 12 – 17.

Kishk, M., and Al-Hajj, A., "An integrated framework for life cycle costings in buildings", RICS Research Foundation , ISBN 0-85406-968-2, 1999.

Klir, G.J., and Ruan, D.A., "Fuzzy logic foundations and industrial applications", Kluwer Academic Publishers, International Series in Intelligent Technologies, 1996.

Kulmala, H.I., (2004), "Developing cost management in customer-supplier relationships: three case studies", Journal of Purchasing and Supply Management 10, pp. 65 – 77.

La Londe, B.J., and Ginter, J.L. (1999), "A Summary of Activity-Based Costing Best Practices, The Ohio State University's Supply Chain Management Research Group", Ohio State Publications.

Mamdani, E.H., and Gaines, B.R., "Fuzzy reasoning and its applications", Academic Press, 1981.

Marquez, A. and Blanchar, C., (2004), "The procurement of strategic parts. Analysis of a portfolio of contracts with suppliers using a system dynamics simulation model", International Journal of Production Economics 88, pp. 29 – 49.

Markowitz, H., (1952), "Portfolio Selection", Journal of Finance 7 (1), pp. 77 – 91.

Momme, J., (2002), "Framework for outsourcing manufacturing: strategic and operational implications", Computers in Industry", Vol. 49, pp. 59 – 75.

Mileham, A.R., (1993), "A Parametric approach to Cost Estimating at the Conceptual Stage of Design", Journal of Engineering Design, Volume 4, Part 2, pp. 117 – 125.

Monczka, R. and Morgan, J., (2000), "Competitive supply strategies for the 21st century" Purchasing, January 13. pp. 48-59.

Olsen, R.F., Ellram, L.M., (1997), "A portfolio approach to supplier relationships", Industrial Marketing Management 26, (2), pp. 101 – 113.

ISPE Parametric Estimating Handbook, (1999), "Appendix A – Definitions of Estimating Terminology", 2nEdition, International Society of Parametric Analysis.

Probert, D. (1996), "The practical development of a make or buy strategy: The issue of process positioning", Integrated Manufacturing Systems, 7 (2), pp. 44-51.

Pugh, P., Faddy, D., & Curran, R. (2010a), Project Management: Cost Analysis, Encyclopedia of Aerospace Engineering, ISBN: 978-0-470-68665-2, Volume 8, Part 37, Chapter 386.

Pugh, P., Faddy, D., & Curran R. (2010b), Project Management: Cost Forecasting, Encyclopedia of Aerospace Engineering, ISBN: 978-0-470-68665-2, Volume 8, Part 37, Chapter 387.

Roy, R., (2003), "Cost Engineering: Why, What and How?", Decision Engineering Report Series, Cranfield University, ISBN 1-861940-96-3.

Rush, C. and Roy, R., (2000), "Analysis of Cost Estimating Processes used within a Concurrent Engineering Environment throughout a Product Life Cycle", Proceedings of CE2000, Lyon, France, July 17 – 21, pp. 58 – 67.

Scanlan, J.P., (2004), "DATUM (Design Analysis Tool for Unit Cost Modeling): a tool for unit cost estimation of gas turbine design within Rolls-Royce", The Cost Engineer, No. 4, November 2004, pp. 8 – 10.

Smith, A.E. and Mason, A.K., (1997), "Cost Estimation Predictive Modelling: Regression versus Neural Network", The Engineering Economist, Volume 42, Winter, No. 2, pp. 137 – 138.

Stewart, R., Wyskidsa, R., Johannes, J., (1995), "Cost Estimator's Reference Manual", 2nd Edition, Wiley Interscience.

Ting, P.,-K., Zhang, C., Wang, B., and Deshmukh, A., "Product and process cost estimation with fuzzy multi-attribute utility theory", The Engineering Economist, Volume 44, Number 4, 1999.

Verhagen, W.J.C., Bermell Garcia, P., Mariot, P., Cotton, J.P., Ruiz, D., Redon, R., Curran, R., (2011), "Knowledge-Based Cost Modelling of Composite Wing Structures", International Journal of Computer Integrated Manufacturing [in press].

Villarreal, J.A.; Lea, R.N. & Savely, R.T., (1992), "Fuzzy logic and neural network technologies", 30th Aerospace Sciences Meeting and Exhibit, Houston, Texas, January 6-9.

Yoon, K., and Naadimuthu, G., (1994), "A make-or-buy decision analysis involving imprecise data", International Journal of Operations and Production Management, 14 (2), pp. 62 – 69.

Wang, Q. and Maropoulos, P. G., (2005), "Artificial neural networks as a cost engineering method in a collaborative manufacturing environment", Next Generation Concurrent Engineering © 2005 ISPE Inc., ISBN 0-9768246-0-4, pp. 605 – 610.

Watson, P, Curran, R, Murphy, A and S Cowan, (2006). "Cost Estimation of Machined Parts within an Aerospace Supply Chain", Journal Concurrent Engineering, Vol. 14, No. 1, 17-26.

Weustink, I.F., Ten Brinke, E., Streppel, A.H., and Kals, H.J.J., (2000), "A Generic Framework for Cost Estimation and Cost Control in Product Design", Laboratory of Production and Design Engineering, University of Twente, The Netherlands.

17

Developing Risk Models for Aviation Inspection and Maintenance Tasks

Lee T. Ostrom and Cheryl A. Wilhelmsen
University of Idaho
USA

1. Introduction

Risk assessment has been used to analyze a wide range of industries to determine vulnerabilities with the ultimate purpose of eliminating the sources of risk or reducing them to a reasonable level. The purpose of this chapter is to show how risk assessment tools can be used to develop risk models of aviation maintenance tasks. Two tools will be discussed in this chapter, though many other methods exist. The tools discussed in this chapter are:

- Failure Mode and Effect Analysis (FMEA)
- Event and Fault Tree Analysis

Ostrom and Wilhelmsen (2011) discuss a wide range of risk assessment tools and this book provides many examples of how these tools are used to analyze various industries.

2. Failure mode and effect analysis

An FMEA is a detailed document that identifies ways in which a process or product can fail to meet critical requirements. It is a living document that lists all the possible causes of failure from which a list of items can be generated to determine types of controls or where changes in the procedures should be made to reduce or mitigate risk. The FMEA also allows procedure developers to prioritize and track procedure changes (Mil Std 882B, C, 1984 and 1993). The process is effective because it provides a very systematic process for evaluating a system or a procedure, in this instance. It provides a means for identifying and documenting:

1. Potential areas of failure in process, system, component, or procedure.
2. Potential effects of the process, system, component, or procedure failing.
3. Potential failure causes.
4. Methods of reducing the probability of failure.
5. Methods of improving the means of detecting the causes of failure.
6. Risk ranking of failures, allowing risk informed decisions by those responsible.
7. A starting point from which the control plan can be created.

FMEA can be used to analyze:

1. Process: Documents and addresses failure modes associated with the manufacturing and assembly process.
2. Procedure: Documents and addresses failure points and modes in procedures.

3. Software: Documents and addresses failure modes associated with software functions.
4. Design: Documents and addresses failure modes of products and components long before they are manufactured and should always be completed well in advance of prototype build.
5. System: Documents and addresses failure modes for system and subsystem level functions early in the product concept stage.
6. Project: Documents and addresses failures that could happen during a major program.

A procedure analysis will be used to demonstrate how an FMEA can be conducted. An FMEA is conducted on a step-by-step basis. Table 1 shows an example of an FMEA table. The following constitutes the steps of an FMEA. These steps will be illustrated by use of an example.

Item	Potential Failure Mode	Cause of Failure	Possible Effects	Probability	Criticality (Optional)	Prevention
Step in procedure, part, or component	How it can fail: –pump not working –stuck valve –no money in a checking account –broken wire –software error –system down –reactor melting down	What caused the failure: Broken part Electrical failure Human error Explosion Bug in software	Outcome of the failures: Nothing System crash Explosion Fire Accident Environmental release	How possible is it: Can use numeric values: 0.1, 0.01, or 1E-5 Can use a qualitative measure: Negligible, low probability, high probability.	How bad are the results: Can use dollar value: $10., $1,000., or $1,000,000 Can use a qualitative measure: Nil, Minimal problems, major problems.	What can be done to prevent either failures or results of the failures?

Table 1. Example FMEA Table

The first step is to create a flow diagram of the procedure. This is a relatively simple process in which a table or block diagram is constructed that shows the steps in the procedure. Table 2 shows the simple steps checking an engine chip detector. Note that this is a simple example and not an exhaustive analysis. Table 3 lists the major, credible failures associated with each step in the process. Table 4 shows the effect of the potential failures. Table 5 shows the complete FMEA for the task.

Inspecting Chip Detector	
Step Number	Process Steps
1	Cut and Remove Lock Wire from Oil Drain Plug
2	Remove Oil Drain Plug
3	Drain Oil
4	Cut and Remove Lock Wire from Chip Detector
5	Remove Chip Detector
6	Examine Chip Detector
7	Clean Chip Detector
8	Replace Chip Detector
9	Lock Wire Chip Detector
10	Replace Oil Drain Plug
11	Lock Wire Oil Drain Plug
12	Replace Oil

Table 2. Process Steps for checking a chip detector

FMEA is a relatively simple, but powerful tool and has a wide range of applicability for analyzing aircraft maintenance tasks.

3. Event tree and fault tree analysis

An event tree is a graphical representation of a series of possible events in an accident sequence (Vesely, William; et. al., 2002). Using this approach assumes that as each event occurs there are only two outcomes, failure or success. A success ends the accident sequence and the postulated outcome is either that the accident sequence terminated successfully or was mitigated successfully. For instance, a fire starts in an engine. This is the initiating event. Then the automated system closes fuel feed. If the lack of fuel does not extinguish the fire, the next step is that that the fire suppression system is challenged. If the system actuates the fire suppression system the fire is suppressed and the event sequence ends. If the fire suppression system fails the fire is not suppressed then the accident sequence progresses. Table 6 shows this postulated accident sequence. Figure 1 shows this accident sequence in an event tree.

As in most of the risk assessment techniques, probabilities can be assigned to the events and combined using the appropriate Boolean Logic to develop an overall probability for the various paths in the event. Using our example from above, we will now add probabilities to the events and show how the probabilities combine for each path. Figure 2 shows the addition of path probability to the event tree.

Inspecting Chip Detector	
Process Steps	Major Failures
Cut and Remove Lock Wire from Oil Drain Plug	No major failures that affect process outcome
Remove Oil Drain Plug	No major failures that affect process outcome
Drain Oil	No major failures that affect process outcome
Cut and Remove Lock Wire from Chip Detector	No major failures that affect process outcome
Remove Chip Detector	Improper removal can remove debris from chip detector and cause false reading. Chip detector can be damaged if improperly removed.
Examine Chip Detector	Aircraft Maintenance Technician (AMT) fails to notice debris on chip detector.
Clean Chip Detector	AMT fails to properly clean chip detector
Replace Chip Detector	AMT fails to properly install chip detector
Lock Wire Chip Detector	AMT fails to properly lock wire chip detector
Replace Oil Drain Plug	AMT fails to properly install oil drain plug
Lock Wire Oil Drain Plug	AMT fails to properly lock oil drain plug
Replace Oil	AMT fails to properly replace oil

Table 3. Failures Associated with Each Step

Inspecting Chip Detector		
Process Steps	Potential Failure Modes	Potential Failure Effects
Remove Chip Detector	Improper removal can remove debris from chip detector and cause false reading. Chip detector can be damaged if improperly removed.	Engine could fail if chips are not properly detected. Added cost to replace damaged chip detector.
Examine Chip Detector	Aircraft Maintenance Technician (AMT) fails to notice debris on chip detector.	Engine could fail if chips are not properly detected.
Clean Chip Detector	AMT fails to properly clean chip detector	Debris could be placed back into engine.
Replace Chip Detector	AMT fails to properly install chip detector	Oil could leak past chip detector. Threads of chip detector could be damaged.
Lock Wire Chip Detector	AMT fails to properly lock wire chip detector	Chip detector could become lose and fall out, leading to loss of engine oil.
Replace Oil Drain Plug	AMT fails to properly install oil drain plug	Engine oil could leak out. Oil drain plug could become damaged.
Lock Wire Oil Drain Plug	AMT fails to properly lock oil drain plug	Oil drain plug could become loose and fall out. Oil drain plug could become damaged.
Replace Oil	AMT fails to properly replace oil	Engine could fail.

Table 4. Effect of Potential Failures

Procedure Step	Potential Failure Mode	Cause of Failure	Possible Effects	Probability	Criticality	Prevention
Cut and Remove Lock Wire from Oil Drain Plug	No major failures that affect process outcome	AMT Fails to Perform Task	Delay in performing task.	Very Low	Not Critical	Ensure AMTs follow work schedule
Remove Oil Drain Plug	No major failures that affect process outcome	AMT Fails to Perform Task	Delay in performing task.	Very Low	Not Critical	Ensure AMTs follow work schedule
Drain Oil	No major failures that affect process outcome	AMT Fails to Perform Task	Delay in performing task.	Very Low	Not Critical	Ensure AMTs follow work schedule
Cut and Remove Lock Wire from Chip Detector	No major failures that affect process outcome	AMT Fails to Perform Task	Delay in performing task.	Very Low	Not Critical	Ensure AMTs follow work schedule
Examine Chip Detector	AMT fails to notice debris on chip detector.	AMT Fails to Properly Perform Task	Engine could fail if chips are not properly detected. Added cost to replace damaged chip detector.	Moderate	Critical	Training, procedures, and inspection oversight
Clean Chip Detector	AMT fails to properly clean chip detector	AMT Fails to Properly Perform Task	Engine could fail if chips are not properly detected.	Moderate	Critical	Training, procedures, and inspection oversight
Replace Chip Detector	AMT fails to properly install chip detector	AMT Fails to Properly Perform Task	Debris could be placed back into engine.	Moderate	Critical	Training, procedures, and inspection oversight
Lock Wire Chip Detector	AMT fails to properly lock wire chip detector	AMT Fails to Properly Perform Task	Oil could leak past chip detector. Threads of chip detector could be damaged.	Moderate	Critical	Training, procedures, and inspection oversight
Replace Oil Drain Plug	AMT fails to properly install oil drain plug	AMT Fails to Properly Perform Task	Chip detector could become lose and fall out, leading to loss of engine oil.	Moderate	Critical	Training, procedures, and inspection oversight
Lock Wire Oil Drain Plug	AMT fails to properly lock oil drain plug	AMT Fails to Properly Perform Task	Engine oil could leak out. Oil drain plug could become damaged.	Moderate	Critical	Training, procedures, and inspection oversight
Replace Oil	AMT fails to properly replace oil	AMT Fails to Properly Perform Task	Oil drain plug could become loose and fall out. Oil drain plug could become damaged.	Low	Critical	Training, procedures, and inspection oversight
			Engine could fail.			

Table 5. Complete FMEA for Chip Detector Task

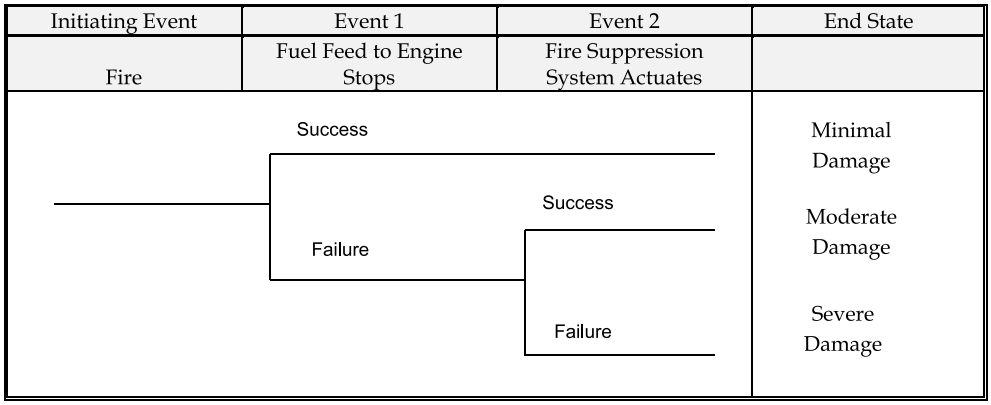


Fig. 1. Event Tree

Event	Description	Possible Outcomes
Fire	This is the initiating event.	
Fuel feed is stopped	The lack of fuel causes the fire to stop.	Success – the fire stops
		Failure – the fire continues
Fire suppression system actuates	The fire suppression system detects the fire and it actuates.	Success – system actuates and controls the fire
		Failure – fire destroys the engine

Table 6. Accident Sequence

Event	Description	Possible Outcomes	Probability
Fire	This is the initiating event.		0.001
Fuel feed stops	The automatic controls stops fuel flow to the engine	Success – stopping fuel flow stops fire	0.999
		Failure – fire continues	0.001
Fire suppression system actuates	The fire suppression system detects the fire and it actuates.	Success – system actuates and controls the fire	0.99
		Failure - system fails to control the fire	0.01

Table 7. Event Sequence with Probabilities

This result of this analysis tells us that the probability derived for a fire in which the fuel feed system stops fuel supply to engine actuates and the consequence in minimal damage is approximately 1/1000 or $1X10^{-3}$. The probability derived for a fire in which the fuel feed system fails to actuate, but the fire suppression system successfully extinguishes the fire and there is only moderate damage is 1E-6 or $1X10^{-6}$. Finally, the probability that a fire occurs and both the fuel feed system fails and fire suppression system fails and severe damage occurs is 1E-8 or $5X10^{-8}$.

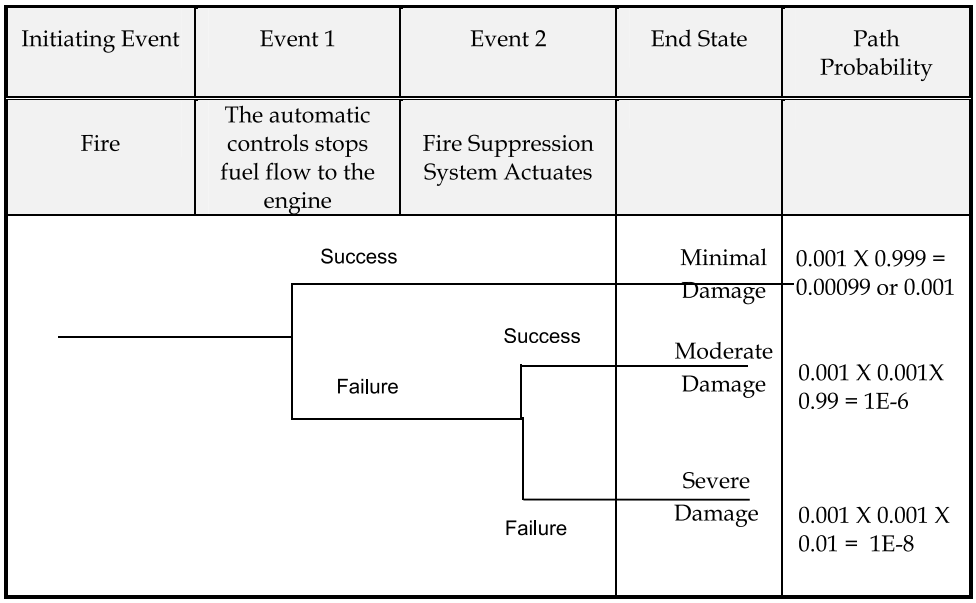


Fig. 2. Event Tree with Path Probabilities

This approach is considered inductive in nature. Meaning the system uses forward logic. A fault tree, discussed below, is considered deductive because usually the analyst starts at the top event and works down to the initiating event. In complex risk analyses event trees are used to describe the major events in the accident sequence and each event can then be further analyzed using a technique most likely being a fault tree (Modarres, M., 2006).

As indicated, the fault tree begins at the end, so to speak. This top-down approach starts by supposing that an accident takes place (Vesely, William; et. al., 2002). It then considers the possible direct causes that could lead to this accident. Next it looks for the origins of these causes. Finally it looks for ways to avoid these origins and causes. The resulting diagram resembles a tree, thus the name.

Fault trees can also be used to model success paths as well. In this regard they are modeled with the success at the top and the basic events are the entry level success that put the system on the path to success.

The goal of fault tree construction is to model the system conditions that can result in the undesired event. Before construction of a fault tree, the analyst must acquire a thorough understanding of the system. A system description should be part of the analysis. The analysis must be bounded, both spatially and temporally, in order to define a beginning and

endpoint for the analysis. The fault tree is a model that graphically and logically represents the various combinations of possible events, both fault and normal, occurring in a system leading to the top event. The term "event" denotes a dynamic change of state that occurs to a system element. System elements include hardware, software, human, and environmental factors (Vesely, William; et. al. 2002).
Table 8 shows the most common fault tree symbols. These symbols represent specific types of fault and normal events in fault tree analysis. In many simple trees only the Basic Event, Undeveloped Event and Output Event are used.

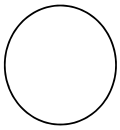
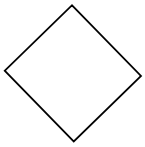
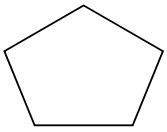
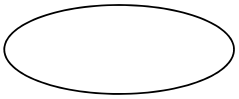
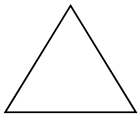
Symbol Name	Symbol	Description
Basic Event		A basic initiating fault (or failure event).
Undeveloped Event		An event which is no further developed. It is a basic event that does not need further resolution.
Output Event		An event that is dependent on the logic of the input events.
External Event (House Event)		An event that is normally expected to occur. In general, these events can be set to occur or not occur,
Conditioning Event		A specific condition or restriction that can apply to any gate.
Transfer		Indicates a transfer to a sub tree or continuation to another location.

Table 8. Common Fault Tree Symbols

Events representing failures of equipment or humans (components) can be divided into failures and faults. A component failure is a malfunction that requires the component to be repaired before it can successfully function again. For example, when a turbine blade in an engine breaks, it is classified as a component failure. A component fault is a malfunction that will "heal" itself once the condition causing the malfunction is corrected. An example of a component fault is a switch whose contacts fail to operate because they are wet. Once they are dried, they will operate properly.

Output events include the top event, or ultimate outcome, and intermediate events, usually groupings of events. Basic events are used at the ends of branches since they are events that cannot be further analyzed. A basic event cannot be broken down without losing its identity. The undeveloped event is also used only at the ends of event branches. The undeveloped event represents an event that is not further analyzed either because there is insufficient data to analyze or because it has no importance to the analysis.

Logic gates are used to connect events. The two fundamental gates are the "AND" and "OR" gates. Table 9 describes the gate functions and also provides insight to their applicability.

There are four steps to performing a Fault Tree Analysis:

1. Defining the problem
2. Constructing the fault tree
3. Analyzing the fault tree qualitatively
4. Documenting the results

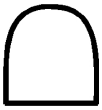
Description	Symbol
AND Gate. The AND gate indicates that the output occurs if and only if all of the input events occur.	
OR Gate. The OR gate indicates that the output occurs if and only if at least one of the input events occur.	

Table 9. Logic Gates

A top event and boundary conditions must be determined when defining the problem. Boundary conditions include:

- System physical boundaries
- Level of resolution
- Initial Conditions
- Not allowed events
- Existing Conditions
- Other Assumptions

Top events should be precisely defined for the system being evaluated. A poorly defined top event can lead to an inefficient analysis.

Construction begins at the top event and continues, level by level, until all fault events have been broken into their basic events. Several basic rules have been developed to promote consistency and completeness in the fault tree construction process. These rules, as listed in Table 10, are used to ensure systematic fault tree construction (American Institute of Chemical Engineers, 1992).

Item	Description
Fault Tree Statements	Write the statements that are entered in the event boxes and circles as malfunctions. State precisely a description of the component and the failure mode of the component. The "where" and "what" portions specify the equipment and its relevant failed state. The "why" condition describes the state of the system with respect to the equipment, thus telling why the equipment state is considered a fault.
Fault Event Evaluation	When evaluating a fault event, ask the question "Can this fault consist of an equipment failure?" If the answer is yes, classify the fault event as a "state-of-equipment" fault. If the answer is no, classify the fault event as a "state-of-system" fault. This classification aids in the continued development of the fault event.
No Miracles	If the normal functioning of equipment propagates a fault sequence, assume that the equipment functions normally. Never assume that the miraculous and totally unexpected failure of some equipment interrupts or prevents an accident from occurring.
Complete Each Gate	All inputs to a particular gate should be completely defined before further analysis of any other gate. For simple models, the fault tree should be completed in levels, and each level should be completed before beginning the next level. This rule may be unwieldy when constructing a large fault tree.
No Gate-to-Gate	Gate inputs should be properly defined fault events; that is, gates should not be directly connected to other gates. Shortcutting the fault tree development leads to confusion because the outputs of the gate are not specified.

Table 10. Rules for Constructing Fault Trees

Many times it is difficult to identify all of the possible combinations of failures that may lead to an accident by directly looking at the fault tree. One method for determining these failure paths is the development of "minimal cut sets." Minimal cut sets are all of the combinations of failures that can result in the top event. The cut sets are useful for ranking the ways the accident may occur and are useful for quantifying the events, if the data is available. Large fault trees require computer analysis to derive the minimal cut sets, but some basic steps can be applied for simpler fault trees:

1. Uniquely identify all gates and events in the fault tree.

2. If a basic event appears more than once, it must be labeled with the same identifier each time. Resolve all gates into basic events.
3. Gates are resolved by placing them in a matrix with their events.
4. Remove duplicate events within each set of basic events identified.
5. Delete all supersets that appear in the sets of basic events.

By evaluating the minimal cut sets, an analyst may efficiently evaluate areas for improved system safety. The analyst should provide a description of the system analyzed, a well as a discussion of the problem definition, a list of the assumptions, the fault tree model(s), lists of minimal cut sets, and an evaluation of the significance of the minimal cut sets. Any recommendations should also be presented. An example fault tree for the engine fire example is shown in Figure 3.

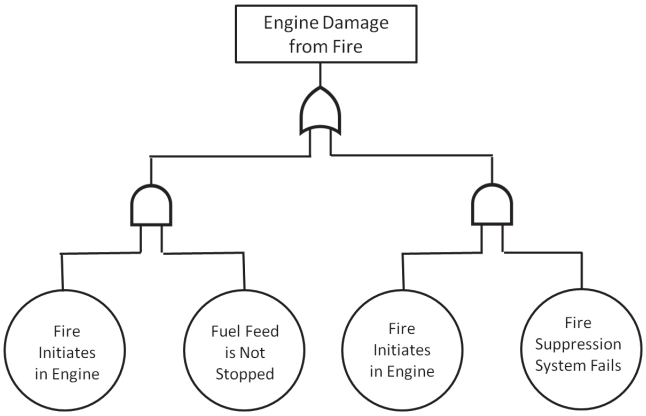


Fig. 3. Example Fault Tree

4. Summary

This chapter discussed how common risk assessment techniques could be used to perform risk assessments of aviation related activities. As discussed in the very beginning paragraph of this chapter, Ostrom and Wilhelmsen (2011) discuss in depth how to use risk assessment techniques to analyze a wide variety of systems, tasks, and activities.

5. References

American Institute of Chemical Engineers., (1992) *Guidelines for Hazard Evaluation Procedures,* New York.

Mil Std 882B, C (1984, 1993)

Modarres, M., (2006) Risk Analysis in Engineering: Techniques, Tools, and Trends, CRC Press; 1 edition, ISBN: 1574447947

Ostrom, L. & Wilhelmsen, C., (Summer 2011) *Risk Assessment Tools and techniques and Their Application, in Process.*

Vesely, William; et. al. (2002) (pdf). *Fault Tree Handbook with Aerospace Applications*. National Aeronautics and Space Administration.
http://www.hq.nasa.gov/office/codeq/doctree/fthb.pdf. Retrieved 2010-01-17.

18

Novel Digital Magnetometer for Atmospheric and Space Studies (DIMAGORAS)

George Dekoulis
Space Plasma Environment and Radio Science Group, Lancaster University
United Kingdom

1. Introduction

This chapter describes the design and scientific results from a new tri-axial magnetometer for atmospheric and space studies. The new instrumentation features multi-frequency (1 KHz – 1.5 MHz), multi-bandwidth (1 Hz – 10 KHz), multi magnetic field measurement range, low-power and re-configurable operation. This novel architecture is based on producing a programmable receiver independent of the fluxgate sensor used. Any ring-core, race-track or parallel type fluxgate sensor may be connected to the input and a working system is produced within hours. The current implementation is based on a novel shielded 3-dimensional sensor optimised for previously unseen low-power consumption compared to existing tri-axial macro-sensors. As a fully digital system the sensor's output is sampled after the amplification stage.

The full extent of the complex solar wind – magnetospheric - ionospheric system is measured by the magnetometer, due to the receiver's programmable dynamic range (DR) and the programmable cross-correlator integration time. The field-programmable gate array (FPGA) hardware implementation yields to a programmable filtering scheme that selects different centre frequencies and desired signal bandwidths. The output is directly in nT to avoid any unnecessary software post-processing calibration procedures and to ensure compatibility of results amongst users. The user selects either the magnetic field's vector components or the total intensity to be present at either the universal asynchronous receiver and transmitter (UART) or the fast 10/100 Mbps Ethernet communications ports.

The development of the magnetometer's sensor is based on the implementation of a novel design methodology for producing optimised fluxgate sensors for low-power consumption. The sensor is qualified for both ground-based and spaceborne observations. The results are of interest to the aerospace and defence industries, since fluxgate sensors exist in airborne and spaceborne systems for decades. Lowering power consumption is followed by simplification of the accompanying electronic circuits.

The system is constantly being enriched by research results from an ongoing collaboration with NASA's Jet Propulsion Laboratory (JPL) on a different project for future missions (Dekoulis & Murphy, 2008). The presented system is complementing existing and under development state-of-the-art systems, such as, scalar/vector magnetometers. The system acts as a pathfinder for future high-resolution planetary exploration missions.

2. Sensors and magnetometers

Fluxgate sensors and magnetometers are used to measure the Earth's magnetic field. Currently, the research conducted on sensors is based on miniaturisation by using new materials. Non-semiconductor sensors, such as fluxgates, induction sensors etc. are already using micro-technologies. Micro-coils and micro-relays are using modern micromachining processes (Seidemann et al., 2000). Amorphous materials such as wires and tapes are applied to sensors (Meydan, 1995).

Within the Space Physics context, magnetic fields of interest are the galactic magnetic field ($B \approx 0.2$ nT), Earth's magnetic field ($B \approx 60$ uT), white dwarf's ($B \approx 1$ KT) and pulsar's ($B \approx 100$ MT). At the poles the field has a 90-degree inclination and magnitude of 60 uT. At the equator the field has 0 degrees inclination and magnitude of 30 uT. There are anomalous cases where the field is 180 uT vertical (Kursk, Russia) and 360 uT vertical (Kiruna, Sweden). Other anomalies are created due to the magnetisation of the rocks and human ferromagnetic structures that could affect communications. During a day the magnetic field fluctuates between 10 – 100 nT, due to solar radiation and the induced ionisation of the ionosphere. The observed micro-pulsations have periods of 10 ms – 1 h and amplitudes up to 10γ. Magnetic storms happen frequently within a month, last for couple of days and exhibit amplitudes of few hundreds of nT.

After extensive analysis of the available technologies, the choice was between the ring-core and race-track sensors. Race track sensors have lower demagnetisation factor, higher directional sensitivity, less sensitivity to orthogonal fields (interference). Race-track sensors exhibit large unbalanced spurious signals and problems from higher tape pressure in the corners. Ring-core sensors have an anuloid excitation coil and a solenoid-sensing coil. Although they have low sensitivity due to the demagnetisation, ring-core designs have many advantages and produce low-noise sensors. Rotating the core with respect to the sensing coil permits precision balancing of the core symmetry. Ring-core sensors exhibit uniform distribution of any mechanical stress. The increased noise associated with open-ended rods is absent. Tape ends are a insignificant source of noise. Sensitivity is proportional to the sensor diameter. For a given diameter, a trial-by-error procedure was followed to determine the optimum for the other dimensions.

In geophysical measurements the scalar resonance magnetometers are used. Overhauser magnetometers have replaced the classic proton magnetometers. Optical magnetometers are used for airborne applications. Similarly to this project, if 3-axial vector measurements are required, fluxgate magnetometers are used.

3. Magnetometers' design study

A fluxgate magnetometer with a ring-core sensor is to be designed. Comparison of existing sensor and magnetometer designs determined the optimum specifications of the novel digital magnetometer (Dekoulis & Honary, 2007).

A review on analogue fluxgates is in (Dekoulis, 2007). The major components of the second harmonic fluxgate magnetometer are presented in the block diagram of Fig. 1. A frequency generator generates the f and 2f frequencies. The f frequency is a sine-wave or a square-wave between 400 Hz – 100 KHz and excites the sensor. For crystalline core materials a 5 KHz square-wave is used. The power amplifier is a totem-pole pair of hexfet transistors. The 2f signal switches the phase-sensitive detector (PSD). The sensor output is amplitude modulated by the Earth's magnetic field and the PSD demodulates it to dc. The analogue

feedback has a large gain, so the sensor functions as a zero indicator. The current output of the voltage-to-current (V/I) converter is used to increase the DR of the instrument and is the current into the compensation coil. The feedback gain controls the sensor's nonlinearity and sensitivity. The sensor coil has roughly 2000 turns. Pre-amplification and bandpass filtering prior to PSD is required to filter the first harmonic and other spurious signals present at the sensor's output. The integrator is used to provide sufficient amplification (Pallas-Areny et al., 1991). The first digital fluxgate magnetometer (Primdahl et al., 1994) verifies that digital technology can be employed into magnetometry. The first real-time fluxgate magnetometer using FPGAs was made by Max-Planck Institute (Auster et al., 1995). A digital fluxgate magnetometer was added to the instrumentation of the Swedish satellite Astrid 2 (Pedersen et al., 1999).

Another solution is presented by (Kawahito et al., 1999). An analogue switching type synchronous detector was used connecting to an analogue integrator and a second order delta sigma modulator. A one-bit digital-to analogue converter was used to close the magnetic feedback loop. The 1-bit DAC guarantees linearity. The output of the DAC is connected to an analogue low-pass filter. The disadvantage with this system is the excessive noise of the device because the magnetic circuit was also implemented on the same device where the digital signal processing electronics were also implemented.

High-performance amorphous ferromagnetic ribbons were integrated on silicon wafers using CMOS manufacturing techniques and batch integration post-process of the ferromagnetic cores to create a 2- D fluxgate for electronic compass applications (Chiezi et al., 2000). The two cores are placed diagonally above the single square driving coil and the sensor is equivalent to a parallel type. However, the significant error of 1.5° (0.5 uT) makes the system inaccurate even for electronic compass applications.

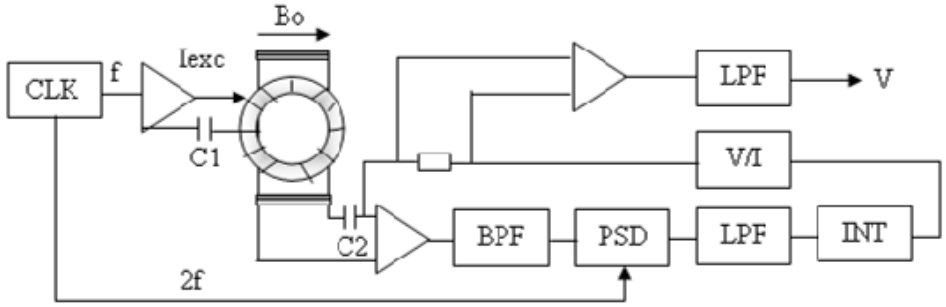


Fig. 1. Second Harmonic Fluxgate Magnetometer.

Space mission results are in the range of 5 pT – 2 mT, with a typical accuracy of 1°. High accuracy is required for mapping of planetary magnetic fields. The integration time for the correlated data is from 1 s to hundreds of samples per second (Ness, 1970). Most magnetometers are of the vector type. Scalar magnetometers measure only the magnitude the ambient field, not the direction. Scalar magnetometers were not considered for this study due to their different architecture, specifications and measurement results. Tri-axial vector magnetometers are widely used on balloons, sounding rockets and spacecrafts. The calibration process is based on known magnetic fields both in amplitude and phase. The specifications include their output for zero-field, scale factor, temperature stability, time drift, weight, power consumption, operating temperature range and radiation hardness. The

Earth's magnetic field has been mapped using vector magnetometers with a resolution of 5 nT and 3 arc-seconds.
Ground based magnetometers, such as the INTERMAGNET, EISCAT, SAMNET and IMAGE networks, couple the operation of spaceborne systems. The Sub-Auroral Magnetometer Network (SAMNET) is operated by Lancaster University. It consists of 13 tri-axial analogue fluxgate magnetometer stations. They use the double-rod parallel type sensors, encapsulated into epoxy material for increased temperature stability, mechanical stress relief and maintenance of the sensors orthogonal relation. Typical measurements of SAMNET include the determination of Pi2 pulsations.
The magnetometer design is rapidly changing and two categories of evolving designs are identified. The first category covers design of a new sensor and the receiver architecture is adjusted to the sensor to obtain acceptable results. In many cases, due to the nature of the sensor the complexity of the receiver is amazing. It seems that there is only one chance to get the receiver working and a high possibility that the results are inferior even compared to a conventional analogue magnetometer.
The second category encompasses designs using existing sensors that improve the receiver's architecture. However, these solutions are accompanied with extra complexity on the digital side, (e.g. three DSP devices for a vector magnetometer) which means lots of hardware, design time, testing time, unknown propagation delays, synchronisation problems, complicated calibration procedures, decreased reliability factors, higher power consumption, which may lead to extra temperature and noise increase, and offset drifts. Three independent communication ports to the host data acquisition system are required.
In both cases the systems are considered to be hardwired, since they perform a specific function. This is a term used in the past for analogue electronics specifically and based on the research results it applies to both the recent digital systems, since there is no flexibility in terms of signal processing. The systems are tied to one sensor, one operating frequency and one specific bandwidth. Most of them are tied to one integration time, although there are cases where a variable integration time scheme is operated. This is justified for the measurement of the Earth's magnetic field, but not when the magnetometer is, also, used for the combined study of the complex solar wind – magnetospheric - ionospheric system. To perform a parametric experiment corresponds to more than one human's effort, significant design time, and possibility of new errors, extra cost and the system is still hardwired. Another disadvantage of the existing analogue and digital systems is the finite DR. In recent digital systems with ADC resolution of 16 bits, the DR is tied to a specific value (maximum 96 dB for a top design), since there is no control on the hardwired analogue electronics. Thus, magnetospheric events and the variance of magnetic fields at the DR edges are lost.
The proposed system implements an architecture where both the analogue and digital counterparts are programmable and controlled by a single controller embedded into the chosen hardware target. The next section is dedicated to the selection of the appropriate hardware implementation technology.

4. Novel digital magnetometer DIMAGORAS

The system presented in Fig. 2 features the following:

- Single chip solution.
- Any ring-core or parallel type fluxgate sensor can be connected to the receiver input.

- Multi-frequency (1 KHz–1.5 MHz).
- 16-bit ADC resolution.
- 3 MSPS maximum sampling rate.
- Multi-bandwidth (1 Hz – 10 KHz).
- Output in nT.
- Selection of the 3 magnetic fields components in magnetic coordinates or the total intensity.
- UART and Ethernet interfaces for data logging.
- Target axis alignment orthogonality is <0.5%.
- Programmable field range of +/- 100 uT.
- Programmable sensitivity of 150uV/nT.
- 96 dB DR expandable to 176 dB.
- Programmable Integration Time. Default is 1 s.
- Re-configurability within ms.
- GPS referenced clock generation.
- GPS UT data timestamping.
- Lengthy data logging or DMA transfers.
- Automatic calculation of the Earth's magnetic field, based on the current GPS location.
- System operation -55 ° -+85 ° C.
- Power consumption <790mW.
- Weight <400 g including the three-axis sensor.
- Digital output at zero field < +/-0.02V.
- Temperature stability per sensor < 0.1 nT/deg.
- Long-term drift: < 10 pT/day.
- Noise < 7 pT/ $\sqrt{Hz}$ @ 1Hz.
- 20-year life span at maximum ratings.
- The FPGA failure rate is less than 1/1,000,000. FPGAs are used for space magnetometers, due to their efficient radiation hardness properties
- Flexible prototyping platform for other space centre ground based magnetometers.
- INTERMAGNET and SAMNET compatibility.
- Low cost lightweight standalone unit.

4.1 System description

The tri-axial sensor is made of high permeability supermalloy material for increased sensitivity and low power consumption. The excitation coil has roughly 28 turns, while each pickup coil 24 turns. The user controlled sensitivity of the magnetometer is 163 μV/nT to cover the +/- 100,000 nT range of the Earth's magnetic field. For a system located away from the poles the Earth's field variations are smaller, so a smaller range, e.g. +/- 65,000 nT, may be selected.

The excitation waveform is FPGA controlled and set to 5 KHz. A transistor network is used to drive the excitation coil to saturation with a square-wave of peak-to-peak current of only 60 mA.

The pre-amplifier (PA) and programmable gain controller are set by default to a gain of 20 dB. The AGC also provides the necessary filtering of the input to keep the noise to low levels. The master controller of the FPGA accurately controls the AGC. The 16-bit AD7621

ADC samples the three signals. The ADC accepts a differential voltage input of +/- 2.5 V. 16 bits provide maximum 96 dB of dynamic range. The AGC functionality is employed when specific thresholds are exceeded to further increase the resolution to a maximum value of 176 dB.

The coil's sensitivity established that no more than 90 dB of gain are needed. The signal from any sensor between 1 KHz – 1.5 MHz may be sampled. Most of the existing sensors reviewed are tuned between 8 KHz – 100 KHz. The ADC samples at 0.9 MHz for a sensor tuned at 100 KHz. Exceeding oversampling increases severely the internal noise.

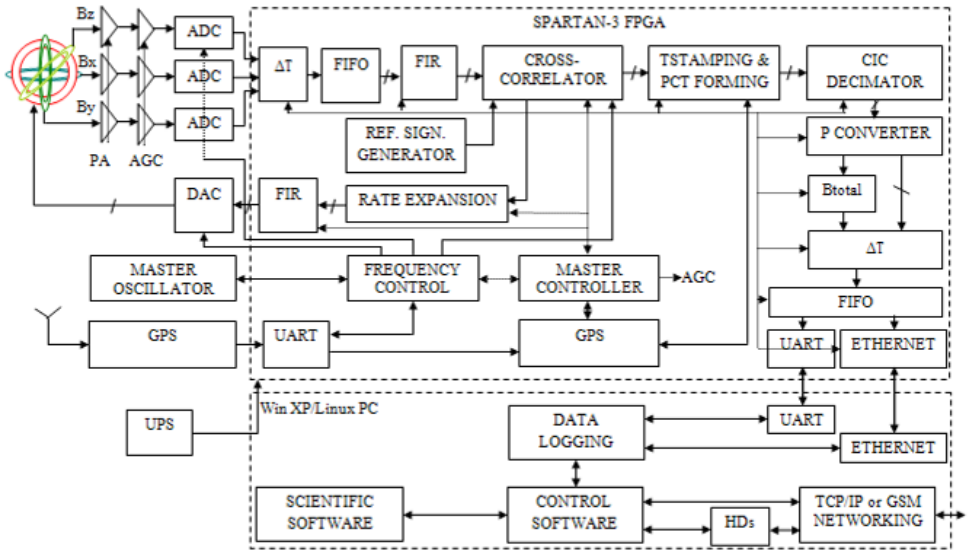


Fig. 2. DIMAGORAS Block Diagram.

The digital finite impulse response (FIR) filter extracts the second harmonic of the output, since all odd harmonics are eliminated due to the designed core structure. It partially controls the bandwidth by low-pass filtering the input signal at the third harmonic located at 15 KHz. The bandwidth is set from 1 Hz – 10 KHz. Most systems use a bandwidth of 10 Hz, which significantly increases the complexity of the FIR filter, compared to a 15 KHz filter.

More than one filtering stages are required to adjust the required bandwidth to a relative to the centre frequency low value. For a bandwidth of 10 KHz a single stage FIR filter is sufficient. For a bandwidth of 10 Hz, more than one stage is required. A set of four fully customisable CIC filters is used. A new bandwidth is set by reconfiguring the FPGA. The classical analogue phase sensitive detector has been implemented digitally to improve its performance.

A reference waveform is generated and multiplied with the output of the FIR filter. The reference waveform is similar to the pulse used to switch on and off the analogue PSD circuit. Values other than 0 and 1 are also used to recreate the sensor's characteristic.

The results of the cross-correlation are integrated for a default value of 1 s, similar to ARIES, IRIS and SAMNET. The outputs of the three integrators are timestamped, packeted and transmitted to the host. The GPS decoder extracts the accurate timing information from the GPS receiver and timestamps the data at the exact place where the measurements are taken. The interpolator and the FIR filter at the feedback loop are used to reduce the quantisation noise of the DAC.

The FIR filtering is effectively low-pass filtering the noise and limiting it to the system's bandwidth. The output of the decimator expresses a relative power measurement, which depends on the system's design characteristics. An internal calibration procedure assigns a fundamental relative power value to a magnetic field of known value, based on the sensitivity, gain, resolution and integration time. In this way, each axis output is translated into nT.

The user selects between the total intensity and individual constituents of the vector field. Decoding software extracts the field intensity and the timing information for each taken measurement. The PC runs the automation software for the production of the theoretical Earth's magnetic and the measured data is co-plotted for direct comparison.

4.2 Sensor design

The novel sensor is in Fig. 3. Each xyz sensor is winded around two unused sides.

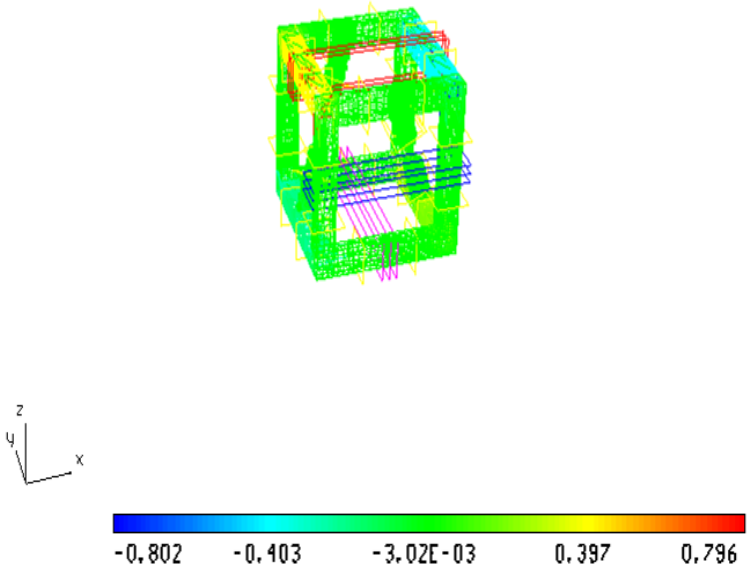


Fig. 3. The 3-D Dimagoras sensor and B Distribution (Bsat = 0.8 T, Iexc = 60 mA).

After following an extensive optimisation technique, the middle sections of the bottom right side are joined with the front topside and the bottom left side with the back topside of the structure using a suitable cone-link section (Dekoulis & Honary, 2008).

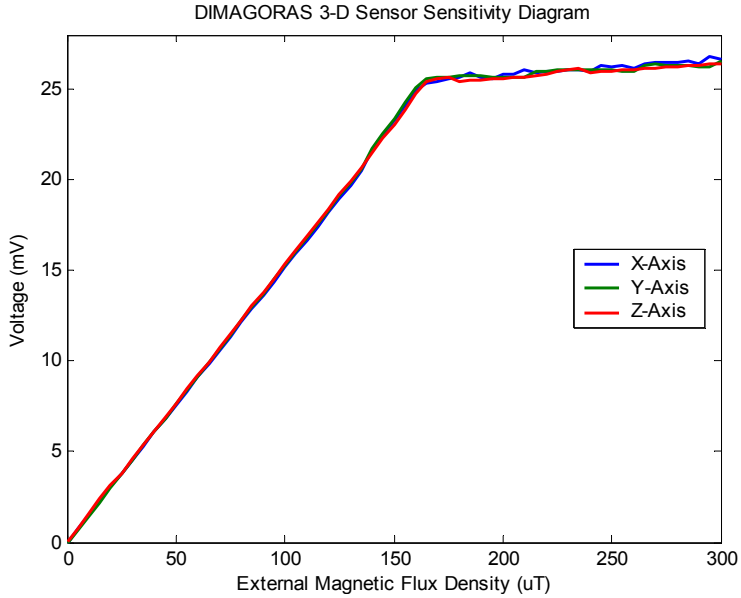


Fig. 4. 3-D Sensor's Sensitivity Diagram.

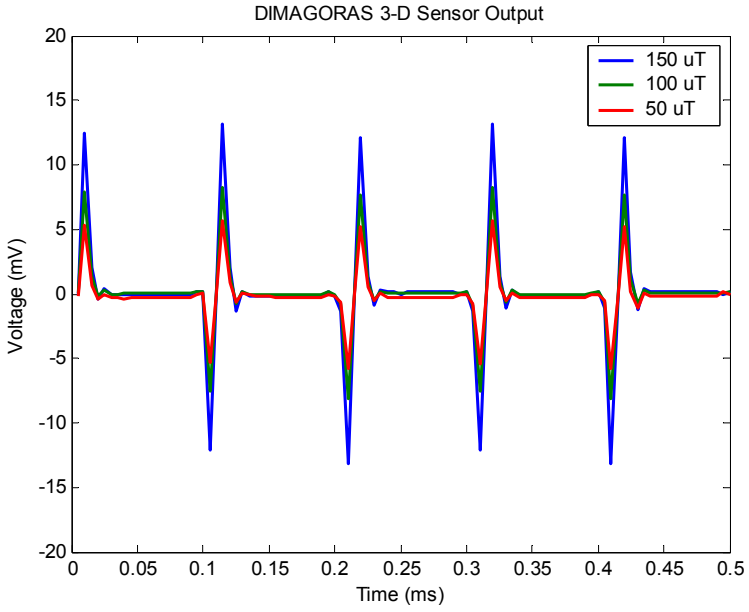


Fig. 5. 3-D Sensor's Output Response for B = 50, 100 and 150 uT.

The flux density distribution in the y-direction is shown in Fig. 3. Using 28 turns of excitation coil wrapped around the structure, the core is saturated at 0.8 T using a square-wave excitation current of +/- 60 mA at 5 KHz frequency. This is in complete harmony with the outcome of a similar optimisation procedure for the single-axis sensor.
The sensitivity diagram of the tri-axial sensor is plotted in Fig. 4. The measured sensitivity is 163 uV/nT. The sensor's y-axis output response for an external field of 50, 100 and 150 uT is in Fig. 5. Vpp = 25.2 mV for B = 150 uT. The characteristic curve for the tri-axial sensor is in Fig. 6.

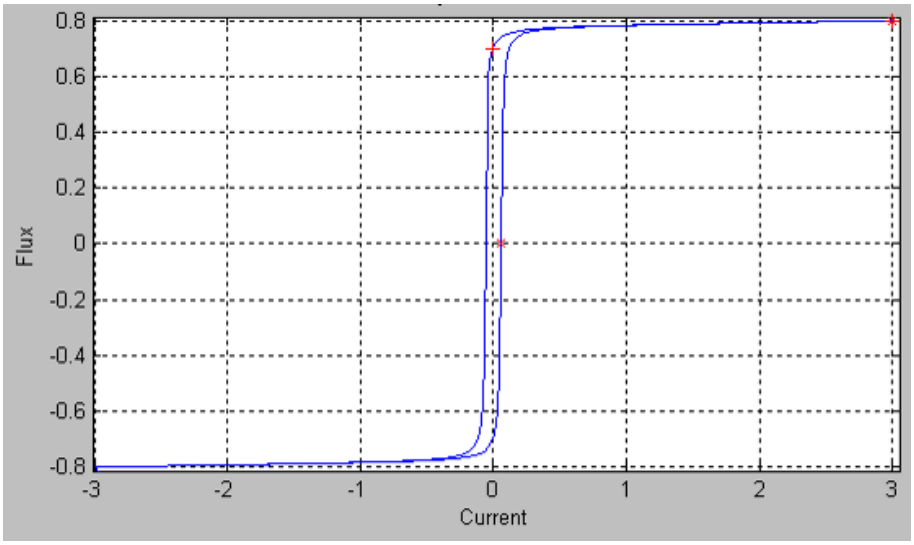


Fig. 6. Sensor's Magnetisation B(H) Curve (T(A/m)).

4.3 Software design

The hardware prototype outputs the timestamped values in nT. This yielded to a reduction in software development time. The user selects the preferred interface, runs the logging software and the results are plotted in nT. This eliminates any calibration and scale adjustments. The power to nT curves are plotted in Fig. 7.
The circuit may be implemented as a decoder or the values may be stored directly into the Block RAM (BRAM) of the FPGA. Both solutions do not account for gain variations and adjustments are required before transmitting the results to the host. The final solution is based on a new real-time numerical approach. A fundamental curve has been produced by averaging the three calibration curves. The error of the derived curve with the original is less than 0.1 % and maximises for values within the saturation region. The error for the linear region is safely assumed to be 0. The curve is approximated by a square function in the linear region. The saturation region is approximated by a linear function. The circuit adapts the weights to pre-known intermediate or maximum values stored in the BRAM.
The total intensity B_T of the magnetic field (Macmillan et al., 1997) is calculated by eq. (1).

$$B_T = \sqrt{B_x^2 + B_y^2 + B_z^2} \tag{1}$$

where, B_x, B_y and B_z are the external magnetic field intensities towards the north, east and zenith directions.

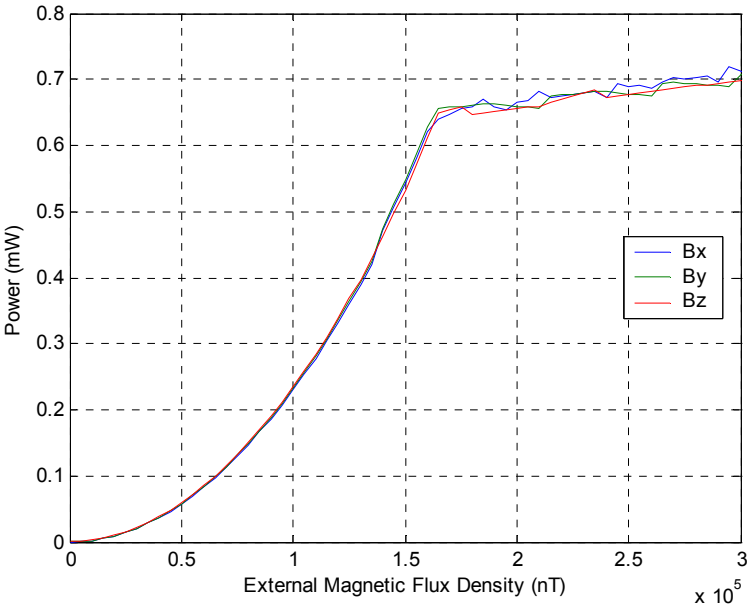


Fig. 7. Power to nT Curves for the 3-D Sensor.

The azimuth field intensity B_ϕ, included in eq. (1), is given by eq. (2):

$$B_\phi = \sqrt{B_x^2 + B_y^2} \tag{2}$$

The inclination I and declination D are given by:

$$I = \arctan\frac{B_z}{B_\phi} \tag{3}$$

$$D = \arctan\frac{B_y}{B_x} \tag{4}$$

The final system implements eq. (1).
The system is SAMNET compatible (---, 2003). Data are recorded using one hour long files. Due to the improved capabilities of the digital magnetometer no further processing is required by the base, since the data are already in nT. The base may remotely choose

between the total intensity and the vector components, already expressed in magnetic coordinates.
The system was initially tested at Lancaster University 54.01° N latitude and 2.77° W longitude, as shown in Fig. 8.

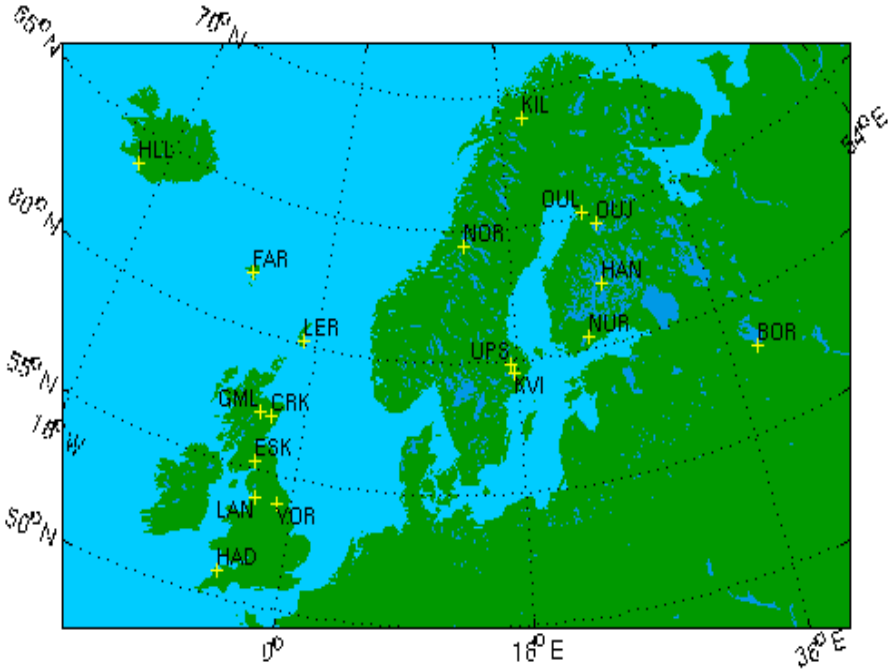


Fig. 8. SAMNET Locations.

Equations (5-11) apply for this geographical location on 28/5/2009.

$$D = -4.294° \quad \text{changing by } 0.171°/\text{year} \tag{5}$$

$$I = 70.635° \quad \text{changing by } -0.005°/\text{year} \tag{6}$$

$$B_x = 16520.85 \text{ nT changing by } 18.63 \text{ nT/year} \tag{7}$$

$$B_y = 1240.43 \text{ nT changing by } 48.3 \text{ nT/year} \tag{8}$$

$$B_z = 47137.77 \text{ nT changing by } 30.25 \text{ nT/year} \tag{9}$$

$$B_\phi = 16567.35 \text{ nT changing by } 15.04 \text{ nT/year} \tag{10}$$

$$B_T = 49964.46 \text{ nT changing by } 33.52 \text{ nT/year} \tag{11}$$

The H component is plotted for all SAMNET stations in Fig. 9.

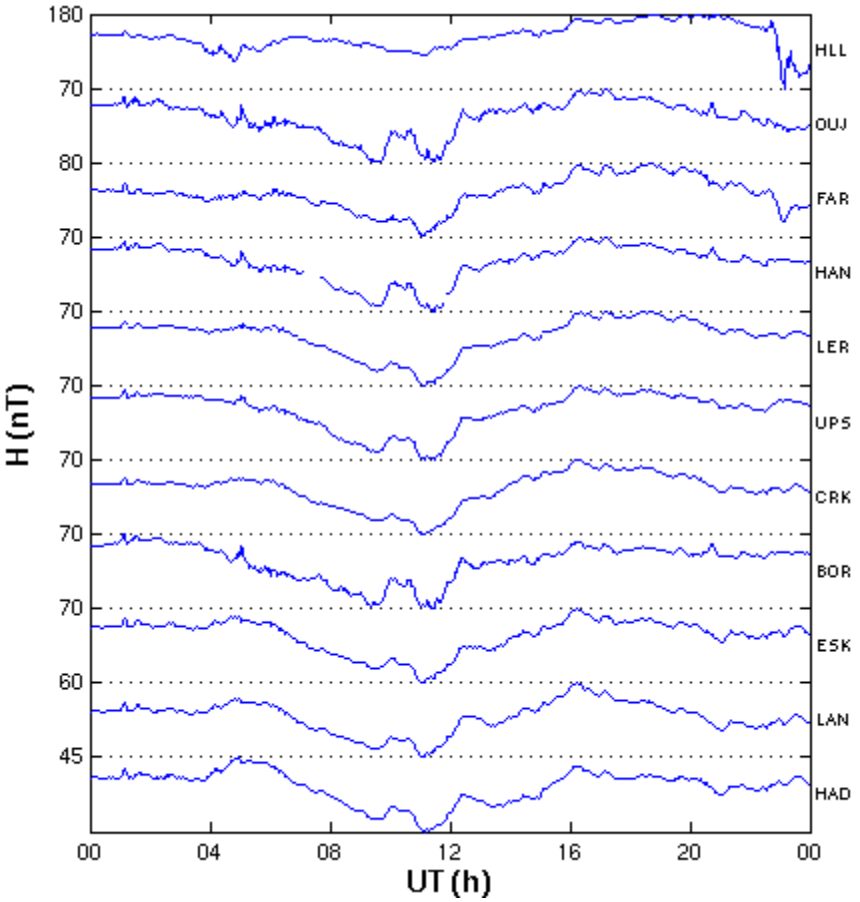


Fig. 9. H-Component SAMNET Magnetogram.

The D component is similarly plotted for all SAMNET stations for the same day in Fig. 10.

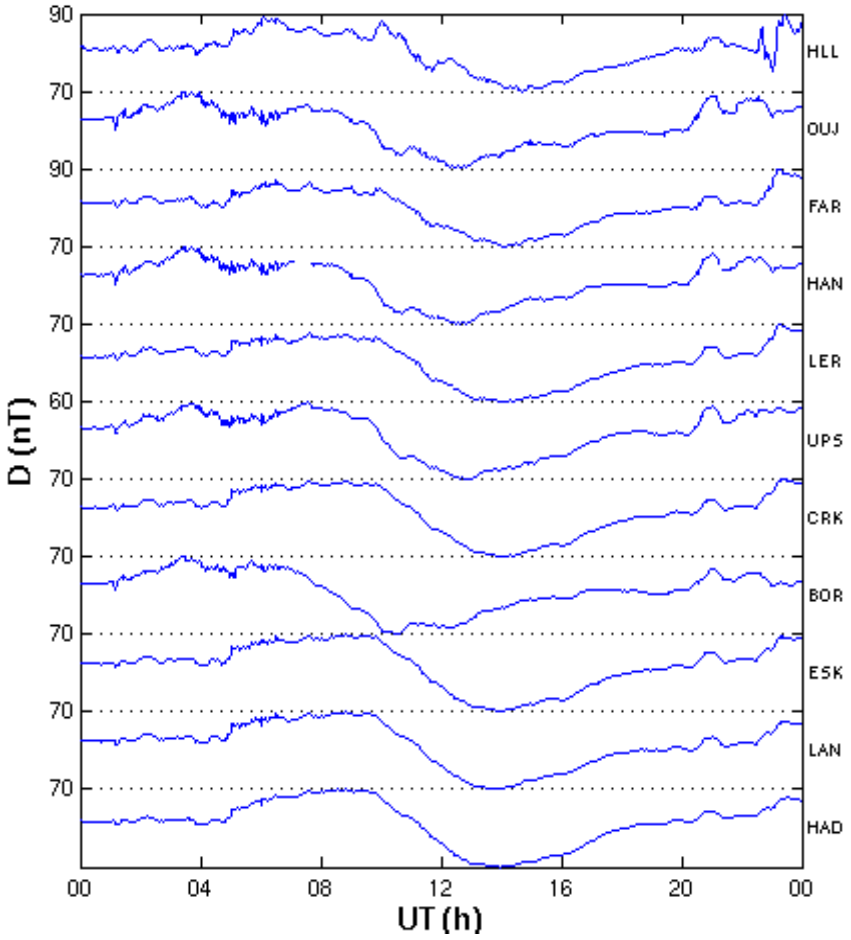


Fig. 10. D-Component SAMNET Magnetogram.

The Z component is plotted for all SAMNET stations for the particular day of measurement in Fig. 11.

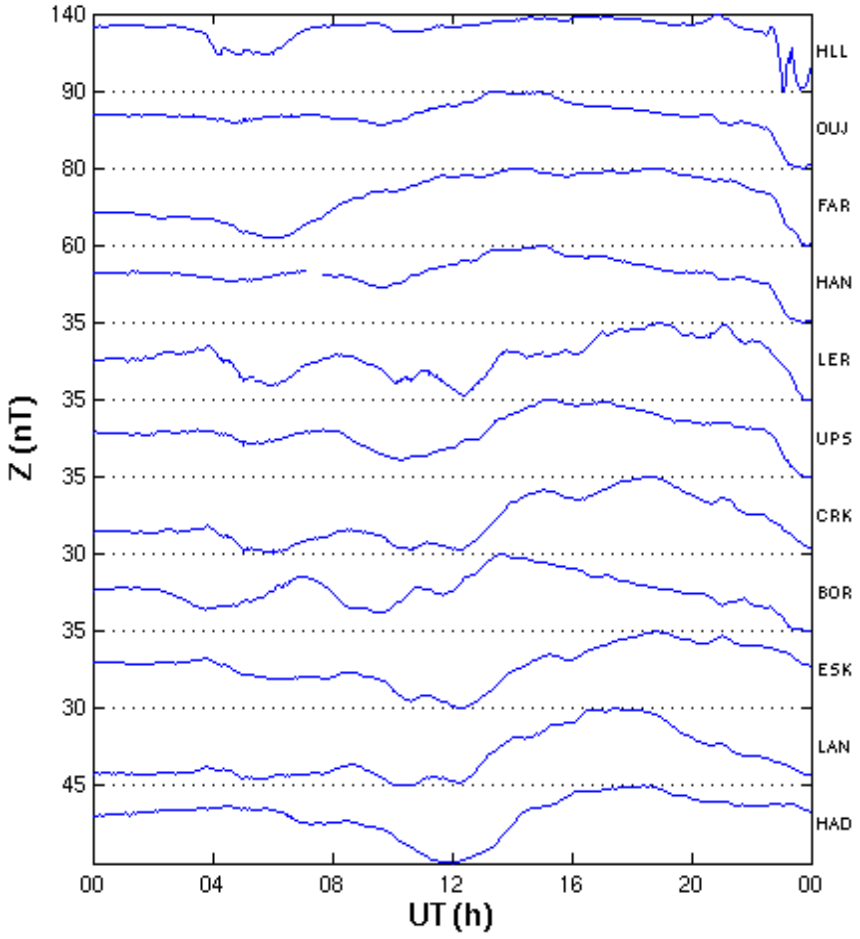


Fig. 11. Z-Component SAMNET Magnetogram.

The H, D and Z components are similarly plotted for DIMAGORAS in Fig. 12.

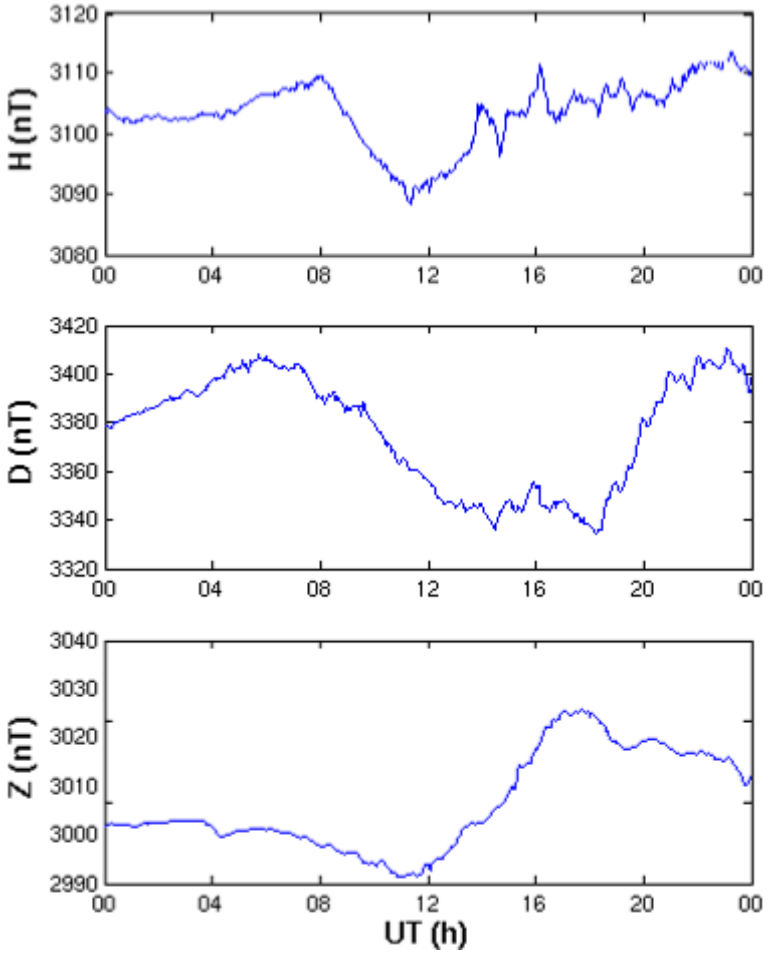


Fig. 12. H, D and Z-Components DIMAGORAS Magnetogram.

For a distant installation, the results are transferred to the central database in an automatic and unsupervised way. Automation software retrieves, at a specific time every day, the last day's data. Various methods have been tested, such as, PPP modem connection, FTP and e-mail.

5. Conclusion

The chapter presents a new reconfigurable magnetometer for measuring planetary fields. The scale is programmable for space field measurements. The modular design allows similar sensors' instrumentations to be quickly evaluated. The all-digital computer architecture implemented allows full control in both the analogue and digital domains. Almost all hardware functions are controlled and occasionally reprogrammed by the FPGA. The FPGA may be reconfigured approximately 20,000,000 times without any problems. 370,000 gates are required for basic operation, which is increased to 640,000 gates for optimum results. This great variation depends on the filters and DSP implementation. The minimum frequency of internal operation is 60 MHz. The system acts as a pathfinder for future space missions, since it is a replacement to existing magnetometers found in every spacecraft.

6. References

Auster, H. et al. (1995). Concept and First Results of a Digital Fluxgate Magnetometer. *Measurements Science & Technology,* Vol. 7, 477-481

Chiezi, L.; Kejik, P.; Jannosy B. & Popovic R. S. (2000). CMOS Planar 2-D Microfluxgate Sensor. *Sensors and Actuators A*, Vol. 82 , 174-180

Dekoulis, G. (2007). Novel Digital Systems Designs for Space Physics Instrumentation, *Ph.D. Thesis*, Lancaster University

Dekoulis, G. & Honary, F. (2007). Novel Low-Power Fluxgate Sensor Using a Macroscale Optimisation Technique for Space Physics Instrumentation. *SPIE, Smart Sensors, Actuators, and MEMS III,* Vol. 6589, 65890G-1 – 65890G-8

Dekoulis, G. & Honary, F. (2008). Novel Sensor Design Methodology for Measurements of the Complex Solar Wind – Magnetospheric - Ionospheric System. Journal *of Microsystem Technologies,* Vol. 14, No. 4-5, 475-482

Dekoulis, G. & Murphy, N. (2008). New Digital Systems Designs for Validating the JPL Scalar Helium Magnetometer for the Juno Mission. *NASA JPL Research Report*

Kawahito, S. et al. (1999). A Delta-Sigma Sensor Interface Technique with Third Order Noise Shaping. *Transducers Conference,* Sendai, Japan, 824-827

Macmillan S.; Barraclough, D. R.; Quinn, J. M. & Coleman, R. J. (1997). The 1995 Revision of the Joint US/UK Geomagnetic Field Models - I. Secular Variation. *Journal of Geomagnetism & Geoelectrism,* Vol. 49, 229 – 243

Meydan, T. (1995). Application of Amorphous Materials to Sensors. *Journal of Magnetic Materials*, Vol. 133, 525-532

Ness, N. F. (1970). Magnetometers for Space Research. *Space Science Review*, Vol. 11, 459-554

Pallas-Areny, R. & Webster, J. G. (1991). Sensors and Signal Conditioning. New York: Wiley

Pedersen, E. B. et al. (1999). Digital Fluxgate Magnetometer for the Astrid-2 Satellite. *Measurements Science & Technology,* Vol. 10, N124-N129

Primdahl, F. et al. (1994). Digital Detection of the Fluxgate Sensor Output Signal. *Measurements Science & Technology,* Vol. 5, 359-362

Seidemann, V.; Ohnmacht, M.; & Buttgenback, S. (2000). Microcoils and Microrelays- An Optimised Multilayer Fabrication Process, *Sensors and Actuators A*, Vol. 83, 124-129

--- (2003). SAMNET Data Collection and Processing. *Lancaster University Technical Report*

19

Aeronautical Data Networks

Mustafa Cenk Erturk, Wilfrido Moreno, Jamal Haque and Huseyin Arslan
University of South Florida
USA

1. Introduction

The wireless connectivity is becoming an integral part of our society. The advances in signal processing, rapid prototyping and an insatiable consumer demand for wireless connectivity is opening a new paradigm of data service, "Aeronautical Data Networks (ADN)". Programs lead by National Aeronautics and Space Administration (NASA), Federal Aviation Administration (FAA) [NASA/CR-2008], EUROCONTROL and Networking the Sky for Civil Aeronautical Communications (NEWSKY) [Newsky] are all including the aeronautical platform as part of their network. The objective is to provide a low delay and cost effective data network for an aeronautical platform, as well as use it as a relay for ground and airborne nodes [Sakhaee], [Medina]. Most of current systems use a satellite for connecting to an aeronautical platform. Satellite resources are limited, expensive and offer limited data throughput as compared to a terrestrial networks. Moreover, frequency spectrum is a valuable estate and needs to be used efficiently. Hence, advance spectrum efficient techniques needs to be evaluated for this environment.

The book chapter will explore the challenges of aeronautical environment to provide connectivity at all times. A detail analysis with mathematical equations will be presented to show the aeronautical channel impairments. The impact of Doppler on the channel that limits the use of a highly efficient modulation scheme, such as orthogonal frequency division multiplexing (OFDM), will be presented. Doppler has a major impact on OFDM based systems. In addition, Doppler spread in ADN depicts rather different characteristics compared to terrestrial networks, i.e., multiple Doppler shifts in the channel and profound delays. Results of parametric spectrum estimation methods for extracting the Doppler shifts will be presented.

OFDM in combination with dense encoding, offers a robust communication and spectrum compression, however its usage is limited to terrestrial domain due to Doppler. OFDM sensitivity to frequency shifts results in intercarrier interference (ICI) and degrades spectral efficiency. High mobility platform, such as train and aircraft offer a challenging environment for OFDM. OFDM ICI and frequency shift caused by the high mobility of the platform is investigated and potential methods are proposed.

ADN's can provide a critical service for various situations, such as: public safety communications, denial of service (DoS), disaster situations, in-flight Internet, as well as mobile communication on the ground such as providing services for highways, trains etc. The network connectivity of ADN will be explored. Current and future prospects of ADN will be discussed in terms of cross interoperability with a terrestrial backbone. The result of

a notional network capacity analysis is presented. Connectivity and robustness of an Aeronautical based Network, both as a relay for terrestrial networks and to provide in-flight internet will be presented.
Finally the chapter will explore the system and architecture requirements for a cognitive driven reconfigurable hardware for an aeronautical platform, such as commercial aircraft or high altitude platforms. The scope of such a system would provide an intelligent configurable radio system, provide connectivity for a changing geographical, political and regulatory environments that an aircraft experiences. Such a system will take advantage of opportunistic services available for today and future. With advances in components and processing hardware, mobile platforms such as those mentioned above are ideal candidates to have configurable hardware that can morph itself, given the location and available wireless service. The global movement of the aeronautical system can take advantage of emerging wireless services and standards. This section of the chapter will propose a system for an intelligent self-configurable software and hardware solution for an aeronautical system, Cognitive Aeronautical Software Defined Radio (CASDR).

2. Motivation and challenges

The ever-changing geographical environment of an aircraft and an increasing availability of different wireless services make's one wonder, what if such services can be accessed in real time.
This provided the motivation to develop a concept system and its hardware that would accommodate to the rapid changes, not just due to the aircraft location, but also to support the growth of services and industry evolution. Fig. 1 depicts the notional framework of opportunistic wireless data service that may be available for an aircraft in flight. At higher altitude the services may be more traditional and fixed, however on ground, the growing WiMAX and local area network services may be available to be accessed from the aircraft. The high-speed mobility of an aircraft adds additional challenges to the design of system physical layer, such as path loss and multi-Doppler spread.

3. Literature review

The desire for a universal and a reconfigurable terminal first appeared in the military area. The need for mobility and accessibility was the driving requirement. One of the early concept was a reconfigurable system appeared as an equipment called "SPEAKeasy". The Software Communications Architecture (SCA) developed by the Joint Tactical Radio System (JTRS) program of the U.S. Department of Defense (DoD) further fueled the growth of SDR. JTRS aims to provide a family of digital, programmable, multiband, multimode, modular radios to alleviate communications interoperability problems. Finally the work of J. Mitola [Mitola_1], there is now a growing interest in reconfigurable terminals.
The increase in air traffic is resulting in the surge of commercial airborne communication system [Eurocontrol]. Aircell and AeroSat have developed the ground based hardware and now offer in flight Internet service. Aircell uses a concept of air-to-ground link [Bluemenstein] and provides the in-flight Internet service called 'gogo' on aircrafts. GOGO service works of cellular phone base stations in the continental US, which act as access points for an en route flight. A recent flight from Tampa, Florida to Detroit, Ohio USA, a user using GOGO service experienced an average upload speed of 0.27 Mbits/s and an

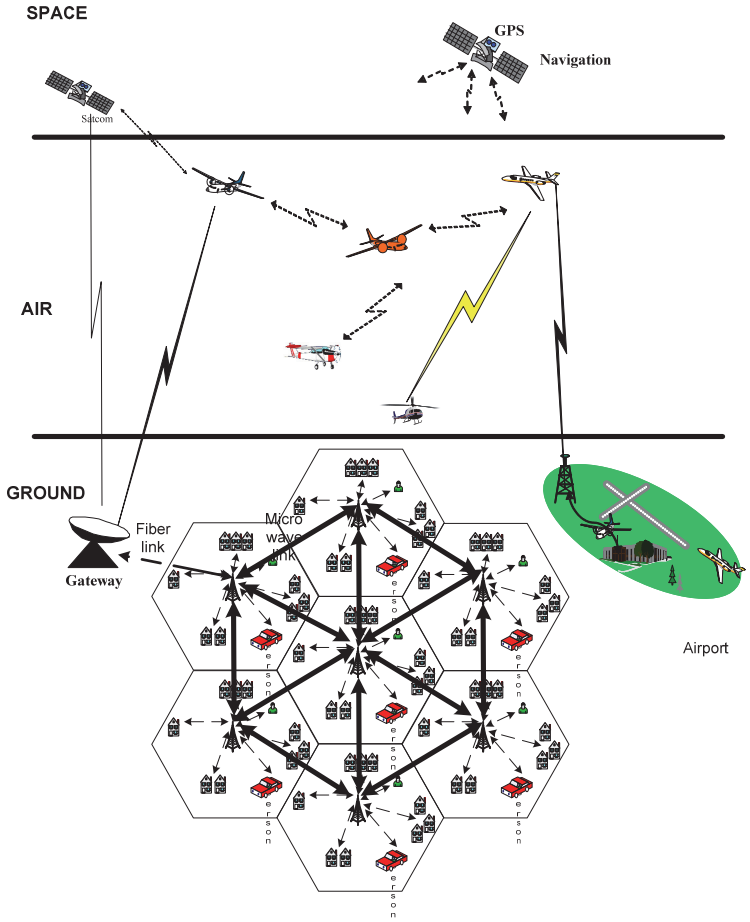


Fig. 1. Aeronautical System

average download speed of 0.33 Mbits/s with latency of 233ms. However, the ground based service is limited to flight coverage over land only. For the oceanic flight satellite based connectivity is required. AeroSat developed satellite communication (SATCOM) Ku band for commercial airliners [AeroSat]. This offers broad connectivity, however the cost and data throughput of satellite based service is not conducive to user demand.

The growth in SDR has been enabled by advances in semiconductor, which has led to the development of programmable multi-core General Purpose Processor (GPP), Digital Signal Processor (DSP), Field Programmable Gate Array (FPGA) and Analog to Digital Converter (ADC). GPP, DSPs and FPGAs provide the programmability and processing capability to realize such a system. Hence, the processing chain starting from digital intermediate frequency (IF) down to demodulation can be implemented in digital signal processing [Srikanteswara], [Mohebbi]. Another key enabler is the high speed ADC that bridges the analog and digital world [Zanikopoulos], [Salkintzis]. Advance algorithms that require intense processing can now be implemented in the combination of these moderate size, weight and power processing

engines. FPGA's, with their ability to parallelize, can implement intense processing algorithms that may be difficult to implement in a DSP or GPP.
Therefore the maturity of; SDR algorithm's, high bandwidth processing engines, development of tunable antenna and availability of high speed ADC makes the implementation of CASDR a possibility. The global mobility of an Aeronautical platform is the ideal implementation of a CASDR concept. A CASDR will learn and configure itself in order to provide multi standard/service modem's as it traverses continents, countries and cities.

4. Aeronautical system

4.1 ASDR system scope

The scope of this system would be to provide an intelligent configurable radio system, provide connectivity for a changing geographical, political and regulatory environments that an aircraft experiences. Such a system will take advantage of opportunistic services available today and planned in future.
The communication design is beginning to converge on standard building blocks, or systems, which form the basic building block of a communication system, i.e., Read Solomon, Turbo Encoder, Modulations, Viterbi etc. Whether a communication link is being developed for short range, long range, line of sight (LOS) or non line of sight (NLOS) the basic building blocks of communication system are the same. If available in software they can be stitched together to build a radio transceiver. Aeronautical Networks (ANs) could be an important application of such systems, since different regions or countries assign different frequency bands based on their needs and spectrum allocation policies.

4.2 Aeronautical network geometry

Geometric relations are observed between an aircraft station (AS) or an aircraft's altitude (h1) with a Ground Station (GS). The LOS communication distance (without considering Fresnel and other parameters) from AS to GS can be calculated using the Pythagoras theorem as follows:

$$d_1 = \left(h_1 \times \left[2R + h_1\right]\right)^{0.5} \cong \left(2Rh_1\right)^{0.5} \tag{1}$$

where, R is the radius of the Earth which varies from 6336 km to 6399 km, but assumed 6370 km (for the purpose of calculations). For distances between the two nodes above the sea level, the above formula needs additional steps for calculating the communication distance. The formula is calibrated by a statistically measured parameter by International Telecommunication Union (ITU), i.e., 'k'.

$$d_1 \cong \left(2Rkh_1\right)^{0.5} \tag{2}$$

Figure 2 shows the maximum communication distances that can be achieved between AS and AS/GS. The jump in the first 2 km altitudes for GS communications can be considered a very low orbit AS which can reach a communication zone of D=120 km. Many commercial planes flying at the altitude of 9 km can potentially create communication zones about D=250 km with a very conservative approach (k=0.5). On the other hand, considering the communication distance between two ASs, it can be inferred that it could reach up to D=480 km with k=1/2.

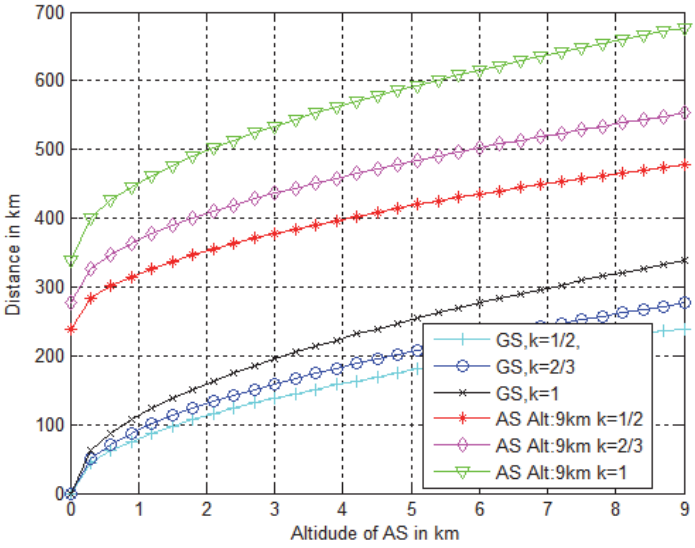


Fig. 2. Communication zone of an AS

Figure 2 shows that ASs could be used as a backhaul or relay for wireless infrastructures, since they have the capability of communicating long distances as compared to wireless ground backhauls. Aeronautical Network (AN) will have a substantial lower round trip delay, which will allow for a low delay telephone and voice over IP services.

4.3 Aeronautical network scenarios and data access

Aeronautical Networks can provide critical services for various situations, such as; public safety communications, Denial of Service (DoS), disaster situations, in-flight Internet, as well as mobile communication on the ground, providing services for highways, trains, etc. The network structure that is being proposed in this paper is as follows: Given a particular region to be covered, initially Service Gateway Ground Stations (SGGS) should be built according to the communication distance, see Fig. 3. Assuming that a GS can communicate to an AS within the distance of 200 km, roughly 8 SGGS will be able to provide service for an area of 1600 km by 800 km.

Data access in an AN can be defined as follows: When a GS or AS has data to send, the flow of the data should be from/to SGGS so the connection with other networks such as public switched telephone network (PSTN), cellular networks and Internet Protocol (IP) could be established.

To provide in-flight services, a centralized configured network should be considered; SGGSs act like Base Stations (BS), covering a particular region where Subscriber Stations (SS) are simply ASs. Scheduling is done by the SGGS and in this structure, AS's are not communicating with each other, except with SSGS's. However, if an AS is not able to register to a SGGS, which could be a case of oceanic flights, then data of that particular AS should be routed to an AS which was already registered to a SGGS with ad-hoc networking strategies.

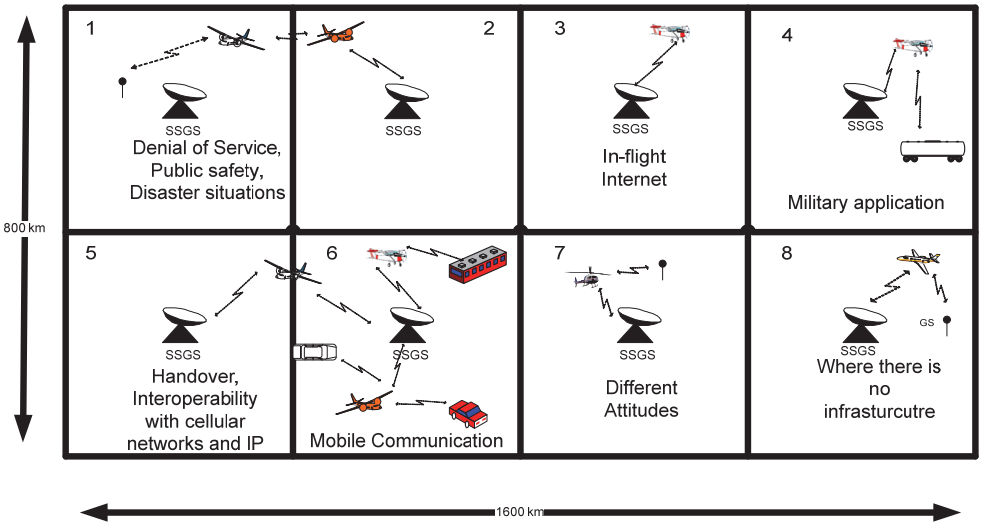


Fig. 3. Aeronautical Network Scenarios

For an AN network the use of AS as a base station used for cellular network is also discussed. In this case, SS's are the GS's, which can be fixed or mobile. When a GS has data to send, it sends its data to an AS. This can be considered as a relay, reflecting the data to its associated SGGS to finalize the establishment. This structure is feasible to provide public safety services in disaster scenarios, provide backhaul option for terrestrial networks and military communication applications. Moreover, in this structure, since both AS's and GS's are not fixed, the handover of a GS between multiple AS is also another challenging issue. It is important to note that the handover process in this structure is somehow different, since GS are doing handovers not only because of their own mobility, but also due to mobility of AS's.

One of the main issues in AN's is the topology estimation. Since there are many mobile stations, in terms of GS and AS, the scheduling and routing of data would differentiate from time to time. In these cases the topology estimation of the network should be done properly, so that the data can be routed and scheduled in mesh and centralized networks strategies respectively.

4.4 Physical layer

In a wireless system design, understanding the limits and bounds of a channel impairments theoretically and empirically are critical to the design of the system. An aeronautical environment poses a daunting task to cover a huge area for any system designer. Global channel characteristics need to be understood to establish model parameters. However, this would lean toward statistical average and will result in inefficient system parameters. Current system based on 'gogo' service, uses a ground based link and provides a limited data rates. A data connectivity sample was taken for a Delta flight traversed between Tampa, Florida to Detroit, Ohio USA, using 'Speed Test' (www.speedtest.net). Different

global servers were pinged periodically duration the flight to measure download, upload and latency. Fig. 4 and Fig. 5 are the global data rates and latency experienced during the flight.

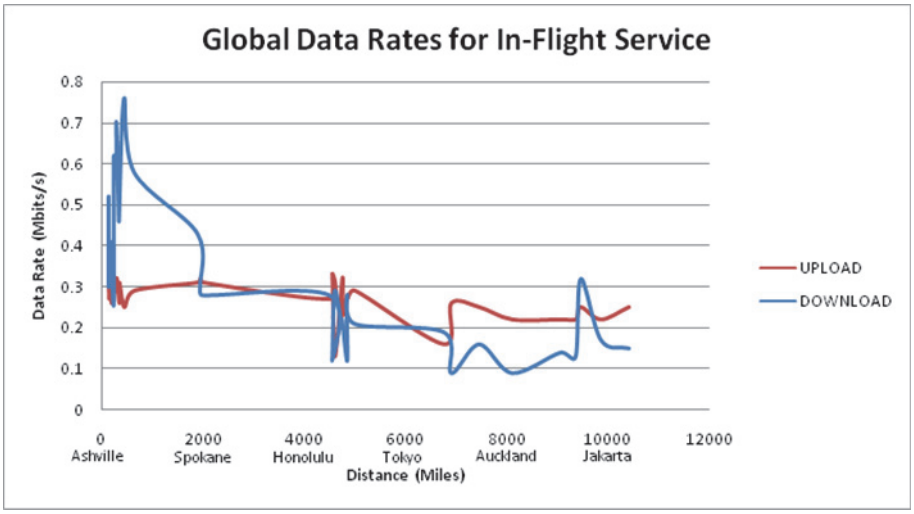


Fig. 4. Global In-Flight Data Rates

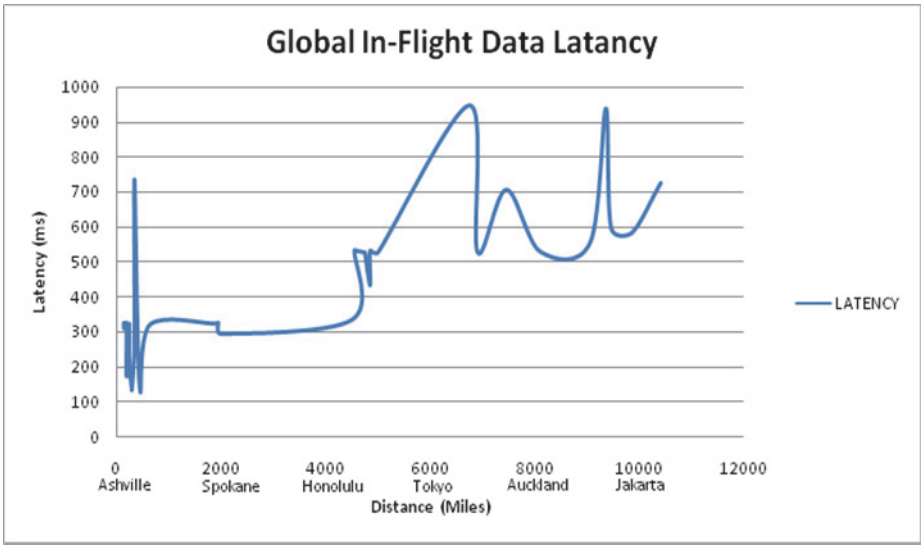


Fig. 5. Global In-Flight Latency

Most of the current system, assumes a line of sight (LOS). This is also the case for the aeronautical platforms connectivity modeling. However, an intelligent CASDR will allow

for the ability to configure the system and learn the channel condition over the flight route and establish history, hence establish accurate channel parameters for a given location, altitude and speed. Since the aircraft traverses pre-planned route, over time this channel parameters will provide accurate characteristics [Bello]. This will allow higher order spectral efficient modulations and multi-carrier system to be used and provide higher data throughput. Details of this cognitive channel sensing behavior are discussed in section 4.

A time varying wireless impulse response is represented by equation (3), where the signal is impaired by amplitude, phase, Doppler and time delay.

$$h(n) = \frac{1}{\sqrt{N}} \sum a_k \exp\left(\frac{j2\pi\theta}{N}\right) \exp\left(\frac{j2\pi f_d n}{N}\right) \delta(t - \tau_k) \tag{3}$$

where a_k , θ , f_d and τ_k are amplitude, phase, Doppler shift and delay for each path respectively.

For a LOS, the effect of number of paths is significantly less, $\tau \approx 0$, f_d, Doppler shift based on platform would be fixed and a limited variation of phase will lead mostly to amplitude degradation due to path loss. For the diffused path, according to Bello [Bello], it represents a wide-sense stationary uncorrelated (WSSUS) channel, which emulates a small area characterization for multipath channel. The effect of Doppler to the line of sight is mostly frequency shift; however the diffused and specular reflections will have a spread due to Doppler. This Doppler spread for an aeronautical communications depicts a bandwidth less than 360º [Hoeher], [Haas], [Elnoubi]. Most of the current research assumes a two-ray model as the channel model for flat surface areas. In an extremely mountainous terrain environment, the channel model results in an intermittent loss of LOS along with increasing angle spread that could match the Jakes Doppler spread. In the two limiting cases; the angular spread at the receiver depicts either a two ray model or Jakes spectrum. Therefore, a modified Doppler spread model needs to be developed, that will go from a narrow beam width to 360º, as the mobile moves from flat to rough environments. Hence, the use of D_f factor from 0 to 1 for the growing Doppler spread, due to beam width, represents going from flat to rough terrain environment:

$$D_f = \frac{\theta_H - \theta_L}{2\pi}, \quad \text{where} \quad (\theta_H - \theta_L) \le 2\pi \tag{4}$$

$$D = \begin{Bmatrix} 2D_f & ,\theta_H > \theta > \theta_L \\ \infty & ,else \end{Bmatrix}, \quad \text{where} \quad \theta = \cos^{-1}\left(\frac{f_d \lambda}{v}\right), \; |f_d| \le f_{d\max} \tag{5}$$

Doppler density going from non-isotropic to isotropic:

$$p_{fd}(f_d) = \begin{Bmatrix} \dfrac{1}{\pi D f_{d\max} \sqrt{1 - (\dfrac{f_d}{f_{d\max}})^2}} & if \; |f_d| \le f_{d\max} \\ 0 & else \end{Bmatrix} \tag{6}$$

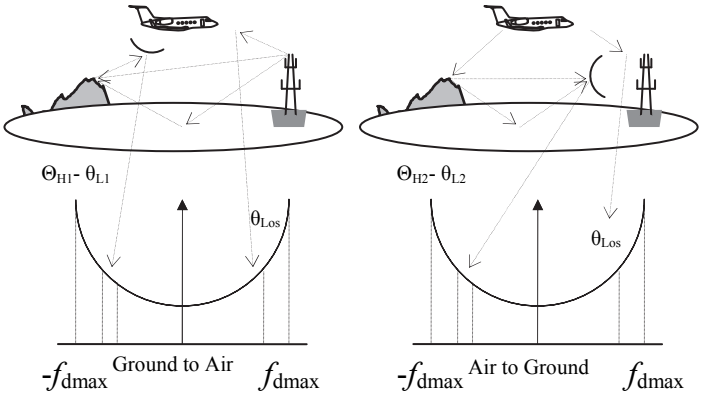


Fig. 6. Doppler Power Spectrums for ADNs

Fig. 6 shows the Air to Ground (A/G) and Ground to Air (G/A) aeronautical communications in an en-route scenario and their corresponding Doppler spectrum. The arrival/take-off, taxi and parking scenarios depicts different multipath and received angle spreads [Haas].The spectrum in an en-route scenario depicts a Doppler shift with a narrow beam Doppler spread, where it can be assumed as another Doppler shift. Among carrier and modulation systems, orthogonal frequency division multiplexing (OFDM) is the most sensitive to Doppler. OFDM based systems has been adopted/proposed for several current/future communication systems all over the world, i.e., asymmetric digital subscriber line (ADSL) services, IEEE 802.11a/g/n, IEEE 802.16, IEEE 802.20, digital audio broadcast (DAB), digital terrestrial television broadcast (DVD) in Europe, ISDB in Japan and fourth generation (4G) cellular systems. Therefore, it is reasonable to assume that any SDR application will also need to support OFDM in the ADN network. In an OFDM based systems, a serial symbol stream is converted into parallel streams and each symbol is modulated with different orthogonal sub-carriers. With the usage of cyclic prefix (CP), since OFDM based systems have already relatively longer symbol durations compared to single carrier systems, they are known for their robustness against frequency selectivity of the channel, i.e., delay spread. However, longer symbol durations lead to weakness of the OFDM systems to time variation of the channel, i.e., Doppler shift/spread which is a challenging issue in ADN.
The two Doppler shifts affecting the system can be described as follows:

$$y(n) = x(n) * h(n) + w(n) \tag{7}$$

where, $w(n)$ is noise and $h(n)$ is the channel impulse response defined as:

$$h(n) = \sum_{i=1}^{2} a_i \exp\left(\frac{j2\pi f_{\Delta i}(n-\tau_i)}{N}\right)\delta(n-\tau_i) \tag{8}$$

where a_i is the attenuation value, N is the number of FFT bins, τ_i and $f_{\Delta i}$ are the delay and the normalized Doppler frequency shift (NDF) for the first and second ray respectively where $f_{\Delta i} = \frac{f_{Di}}{\Delta_f}$.

For the ADN, Figure 4 presents the two-path channel model. In OFDM, as long as the carrier's orthogonality is maintained, then there is no bleeding of energy. Intercarrier interference (ICI) is related to sub-carrier bandwidth and their proportional interference due to Doppler offset. As an example, an estimation of ICI interference for a system with 64 point FFT OFDM symbol, which has a 312 kHz subcarrier bandwidth, is plotted in Fig. 7.
The Fig. 7 shows the ICI error vector contribution due to frequency shift caused by Doppler in a two ray model. At 0.1 fraction of sub carrier frequency the ICI error contribution approaches -10dB. To support higher spectral efficiency generally ICI should remain within or less than 0.02 fraction of sub-carrier frequency. This will allow ICI interfering energy to remain well below -25 dB allowing higher spectral efficiency.

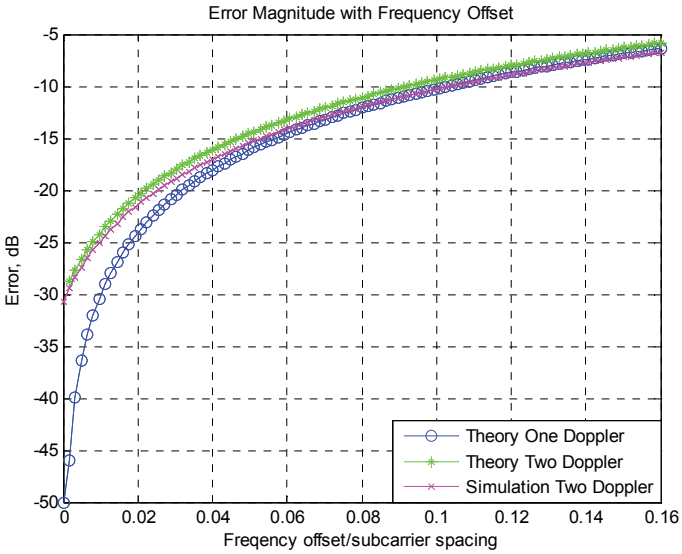


Fig. 7. Doppler ICI vs. Sub-carrier BW

4.5 Cognitive route based physical layer estimates

The aircraft routes driven by FAA for various segments are ideal to establish a history of wireless channel conditions for the route. Once a route is traversed, its history of channel impairments are stored with associated coordinates and aircraft attitude information. This data is downloaded to a central database to be shared with another aircraft. For new routes, the cognitive channel estimator will try to understand the channel condition. Over time, the channel history collected from different aircraft will create a channel map for each route. The ASDR will then be able to download this data and prior to a flight adjust the physical layer parameters for the route. For a mobile platform that has a predetermined route, such as AN, the channel estimation is broken down to static and dynamic components. The static components effecting the channel would be large objects i.e., mountains, buildings, etc. The averaging over multiple routes will provide of stable static channel estimates. The dynamic components will be due to time varying objects.

5. Aeronautical software define radio

The advances in components and signal processing techniques are the leading enabler for a configurable hardware and intelligent software. Software defined radio emerges from the desire of single radio hardware that molds its feature to different radio schemes [Apostolis]. The artificial intelligence needed for the smarts of such configurable hardware is emerging into what is known as cognitive radio [Mitola_2].

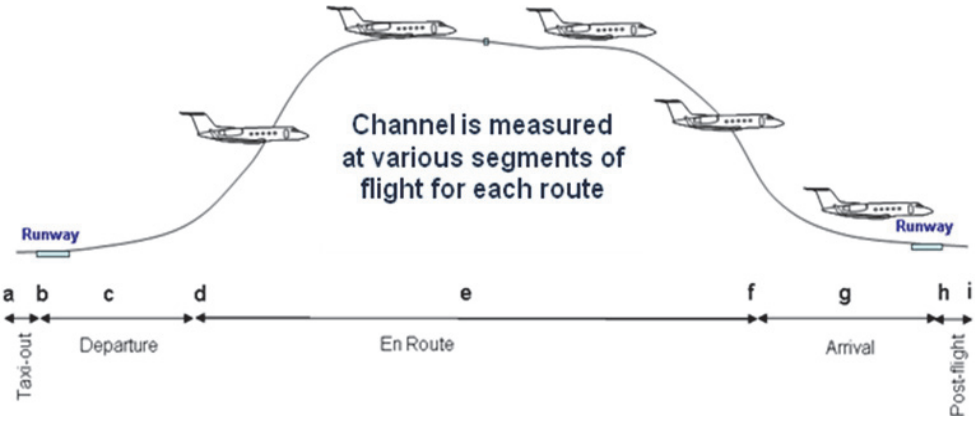


Fig. 8. Route based Channel Sensing

Cognitive algorithms combined with configurable hardware can take full advantage of varying location of an aircraft, whether that is in the air, en route and lends themselves to take advantage of opportunistic spectrum and network connectivity.
A system with the ability to morph to accommodate the aeronautical changing environments, channels conditions across domestics and international boundaries is required. Aeronautical software defined radio (ASDR) platform will also allow the flexibility to comply with countries regulations governing the spectrum usage and interference.

5.1 Spectrum coverage

The spectrum bandwidth use and frequency band allocation for different systems is one of the challenges to overcome for truly building an ASDR capable of accommodating itself for different regions. For a given region or country, the standard may be the same but the frequency band used may be different. For example, the 802.16 specification applies across a wide range of radio frequency spectrum and WiMAX could function on any various frequencies i.e., 2.5 GHz is predominantly being used in the USA, elsewhere in the world 2.3 GHz used in Asia and some countries are using 3.5 GHz.
The Analog TV bands (700 MHz) may become available for WiMAX usage, but currently it is being used for digital TV, however different countries might choose to use the spectrum that best suits their needs. Table 1 below lists opportunistic frequency data network available [Zhang], [Peter].

Band (GHz)	BW (MHz)	Standard	Region	Service
2.4	20	802.11b/g	US	Wi-Fi
5	20	802.11a	US	Wi-Fi
2.5	20	802.16	US	Fixed WiMAX
3.5,2.5	20	802.16a	Can	Fixed WiMAX
2.3	20	802.16e	Aus	Fixed WiMAX
1.616 - 1.6265	10.5	Custom	Global	Iridium Down Link
19.4 -19.6	20	Custom	Global	Iridium Up Link
2,4		Sirius/XM	US	QPSK, OFDM
1.9, 0.85	1.23, 5	W/CDMA	US	3G Cellular
1.8, 0.9	1.23, 5	W/CDMA	EU	3G Cellular
0.5 – 0.8	n/a	n/a		Analog TV

Table 1. Wireless Standards

Another feature that will be necessary in a SDR application is a tunable RF front end capable of locking on the various bands.
Frequency bands and bandwidths for future wireless communication studies in terms of aeronautical communications are discussed at the World Radio communication Conference (WRC) 2007. This international body maintains and agrees to abide by the use of spectrum by international treaty. Aeronautical Mobile (Route) Service (AM(R)S) communication is defined as a safety system requiring high reliability and rapid response. Safety and security applications together with, Air Traffic Control (ATC) and Air Traffic Management (ATM) communications are considered to be AM(R)S. To accommodate the future growth of aeronautical communication, new band allocations are being made in AM(R)S rather than VHF band in L and C. L band (960-1164 MHz) and C band (5091-5150 MHz) allocations are discussed in the meeting. L band is suggested as a suitable band for future aeronautical communication studies. C band is considered to be used in Airport surface network systems, since it is thought to be useful for short range, high data throughput.

5.2 Critical system parameters

Cognitive radio system will require optimization of system performance. Algorithms capable of real time optimizing the system performance as well as pre/post flight will create pre-flight configuration Table 2;

Aeronautical Optimization Parameters
Customer Usage
DQ: Quantity of data transferred at various flight segments. *DT:* Duration of data transfer per segmented route. *TC:* Traffic classes: multi-media, navigation, system health & safety. *BER:* Required Bit error rate per Traffic Classes.
Network & Data Access Layer
Protocol Selection, Routing configuration, Forward error selection given the customer driven BER, Available to provide relay service, Packet error rate
Physical Layer
T_M :Multipath delay spread: Characterizes channel smearing due to arrival of multi-signal reflection arrivals. $f_{DS:}$ Doppler spread or Doppler bandwidth. $f_{d:}$ Doppler Shift. *A:* Attenuation: power loss, function of frequency and distance. *L:* Impulse Response Length: length, in signal elements, of CIR. *Band:* Carrier frequency Band. *BW:* Available bandwidth. *SWP:* Standard waveform performance.

Table 2. Parameters

5.3 Aeronautical cognitive radio

The term cognitive comes from psychology meaning "brains" the ability to learn and understand. The aeronautical environment is ideal application for an intelligent radio,

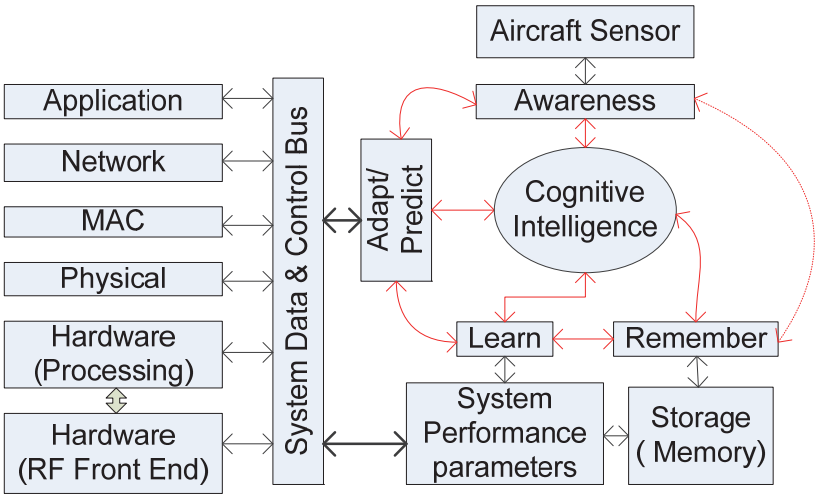


Fig. 9. Aeronautical SDR and CE

which is capable of learning the environment for various locations and altitudes, see Figure 7 and 8. Over time, each aircraft flying over certain route will store the data on board storage devices. This data shall contain the route the airline/aircraft traversed, the opportunist wireless links available, frequency band, bandwidth, data rate, wireless standard, signal quality for the route, etc. Upon arrival at the destination, data is then downloaded to a centralized flight communication data bank. This data is then available for flights heading on the same route.

5.4 Aeronautical cognitive intellegence

The brain of the aeronautical cognitive engine would be to work of its constant awareness of aircraft geographical location and RF channel. It will sense weather conditions that may affect the radio transmission and available services available.

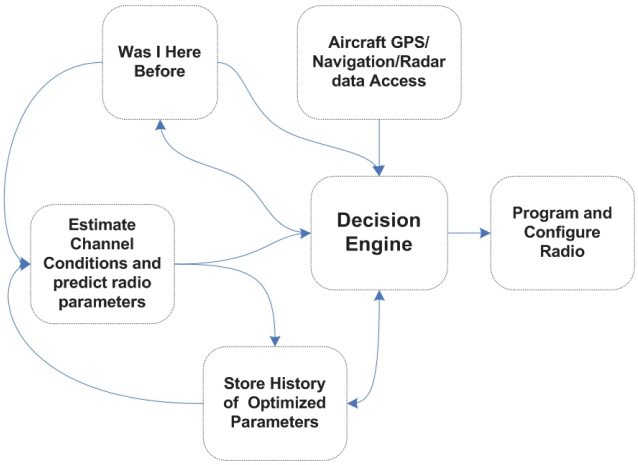


Fig. 10. Cognitive Engine

- *Awareness:*

The aircraft navigation and radar systems will provide the sensing stimulation to the cognitive engine. The inertial measurement unit (IMU) used for flight navigation will provide aircraft speed, altitude, and attitude. Advance forward looking radar will provide the weather conditions that may affect the radio transmission performance. Global position system (GPS) will provide location of the aircraft with respect to global geography. Furthermore, the awareness engine will have the ability to estimate the data requirements based on past data use and flight profile, before accessing the spectrum for services.

- *Learn:*

The cognitive awareness provides an opportunity for CASDR to learn the spectrum usage, data demand and system throughput based flight route during day or night. Such statistics will allow a constant learning and developing statistics profile that is stored for each route. This allows cognitive radio of other airlines that have not travelled that particular route to have a priori knowledge and schedule services accordingly. The system parameters available at particular location can be configured for that country or location.

The channel sensing and estimation for the flight route will serve to establish channel statistics, as shown in Fig. 11. Aeronautical Channel Sensing. The CASDR cognitive channel awareness can configure the system to establish channel impairments for the flight route.

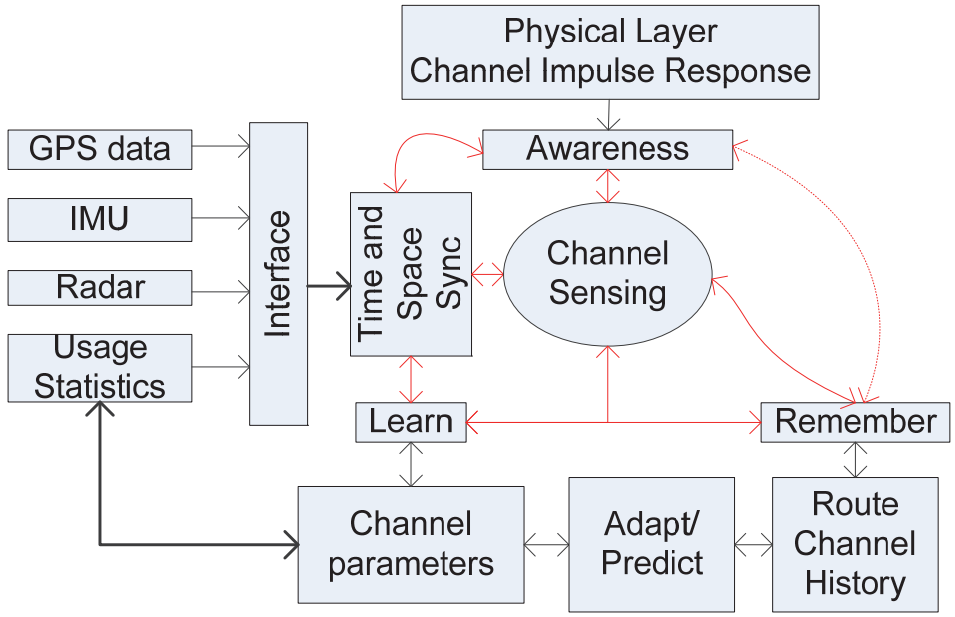


Fig. 11. Aeronautical Channel Sensing

- *Remember:*

System performance data gathered for different flight routes through different airlines will serve as means to remember these flight parameters, exchanged through a centralized data archives. Such data will grow in time and averaging over time will provide a reliable statistics for configuring the SDR radio parameters.

- *Adapt and Predict:*

The cognitive engine learning and sensing ability with an aircraft system will allow ability to predict system configuration parameters and adapt them to data gathered through flights travelled by other carriers.

5.5 Aeronautical configurable hardware

The key to a configurable system for an aeronautical system is to design hardware with minimum analog front-end, access different antenna system, digitize the signal and a scalable architecture**.** Fig. 12. Aeronautical Software Defined Radio presents such a system. The RF front-end board will support multiple bands with varying gain amplifiers. Closely coupled A/D boards with FPGAs are required for high-speed data connectivity and processing. A technique such as under sampling for demodulations is used to reduce the front-end components. The advances in ADC devices as well as non-compliance feature of Nyquist sampling theory is an enabler for an ASDR application. Violation of Nyquist theory will create signals aliased at integer multiple of sampling frequency (N* fs). This put the

challenge on front-end processing system. The advances in programmable Digital Signal Processing (DSP) and Field Programmable Gate Array (FPGA) are ideal for such processing [Susaki]. FPGA offers ability to parallelize processing hence, allow a high-end processing throughput. The Virtex-6 FPGA family by Xilinx provides up to 2,016 DSP48 slices that deliver up to 1000 Giga MAC/s of DSP processing performance. Xilinx offers solutions for evolving standards such as WCDMA, WiMAX, TD_SCDMA and LTE. Texas instrument DSP products are now offering six DSP processors in a single package with processing capability 4000 million MACS (16-Bits) at 500 MHz.

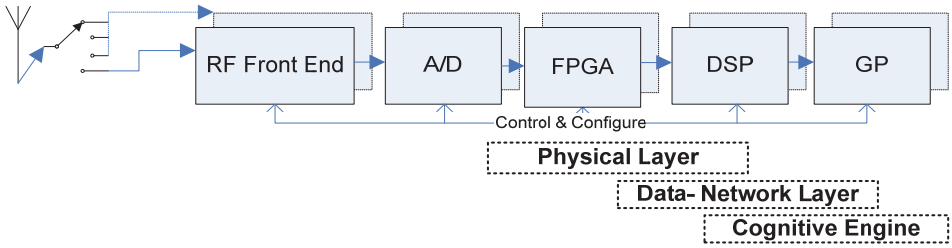


Fig. 12. Aeronautical Software Defined Radio

6. Conclusions

The advances in component technology, evolution in communication services, and increase in data demand and aircraft mobility creates an ideal application for CASDR to support the aeronautical system. Current deployed systems are beginning to form shape, i.e., gogo, however they are now adding another hardware box to provide connectivity. Since the system is hardwired for a particular modem, the evolution will require hardware modification to keep up with growth in the telecommunication growth. Accurate measurements of channel characteristics, such as Doppler, will allow spectral efficient modulation to be used for higher data rates. Advance algorithms along with processing capabilities can resolve the impact Doppler, due to aircraft high mobility. The novel cognitive channel measurement and estimation for each route will increase spectrum efficiency and in return provide high data throughput. An optimum combination of bandwidth, subcarrier bandwidth, acceptable Doppler frequency and multipath immunity system can be developed for ADN. This will result in an efficient use of the spectrum and provide a high data rate for the global connectivity.

7. References

AeroSat Corporation. (2008). About Airborne SatCom [Online]. Available: www.aerosat.com/about/ about_airborne_satcom.asp

Apostolis K. S., Hong N., & P. Takis Mathiopoulos, ADC and DSP Challenges in the Development of Software Radio Base Stations.

Bello P. A., "Aeronautical channel characterization," *IEEE Trans. communication*., vol. COM-21, pp. 548–563, May 1973.

Blumenstein J. (2007). Aircell: Inflight Wi-Fi Built for the Airline Business [Online]. Available: blog.aircell.com.

Elnoubi, S.M., "A simplified stochastic model for the aeronautical mobile radio channel," Vehicular *Technology Conference, 1992, IEEE 42nd* , vol., no., pp.960-963 vol.2, 10-13 May 1992.

Eurocontrol, STATFOR - Air Traffic Statistics and Forecasts, http://www.eurocontrol.int/statfor/

Haas, E., "Aeronautical channel modeling," *Vehicular Technology, IEEE Transactions on*, vol.51, no.2, pp.254-264, Mar 2002.

Hoeher, P., & Haas, E., "Aeronautical channel modeling at VHF-band," *Vehicular Technology Conference, 1999. VTC 1999 - Fall. IEEE VTS 50th*, no., pp.1961-1966 vol.4, 1999.

Medina D., F. Hoffman, Ayaz, S., Rokitansky, C. H., Feasibilty of an Aeronautical Mobile Ad-Hoc Network Over the North Atlantic Corridor, *Proceedings of IEEE Secon* 2008.

Mohebbi B., E. C. Filho, R. Maestre, M. Davies, & F. J. Kurdahi, "A case study of mapping a software-defined radio (SDR) application on a reconfigurable DSP core," *in Proceedings of 1st IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*, pp. 103–108,Newport Beach, Calif, USA, October 2003.

Mitola_1 J. and Gerald Q. Maguire, Jr, Cognitive Radio:Making Software Radios More Personal.

Mitola_2 J., The software Radio architecture, *IEEE Commun. Mag*, vol. 33, no. 5, pp. 26--38, May 1995.

NASA/CR—2008-215144, Future Aeronautical Communication Infrastructure Technology Investigation.

Newsky Project website: http://www.newsky-fp6.eu/.

Peter W., Glenister, S M., & Alan W. Miller *Motorola IEEE AES Systems Magazine*, November 1999.

Salkintzis, A.K., Hong Nie, & Mathiopoulos, P.T., "ADC and DSP challenges in the development of software radio base stations," *IEEE Personal Communications*, vol. 6, no. 4, pp. 47–55, Aug. 1999.

Sakhaee, E., & Jamalipour, A., "The Global In-Flight Inter-net," Selected Areas in Communications, *IEEE Journal on* , vol.24, no.9, pp.1748-1757, Sept. 2006.

Srikanteswara R., Chembil Palat, R., Reed J. H., & Athanas P., An Overview of Configurable Computing Machines for Software Radio Handsets, *IEEE comms Mag*., Vol.41 No.7, pp.134-141, July 2003.

Susaki H., "A Fast Algorithm for High-Accuracy Frequency Measurement: Application to Ultrasonic Doppler Sonar," *in IEEE journal of oceanic engineering*, vol. 27, no. 1, January 2002.

Zhang, Y., & Nirwan Ansari, Wireless Telemedicine service over integrated IEEE802.11/Wlan and 802.16/Wimax, *IEEE Wireless Communications*, February 2010.

Zanikopoulos A. Hegt, & Van Roermund H., A. Programmable/Reconfigurable ADCs For Multi-standard Wireless Terminals, *IEEE Conferences in Communications, Circuits and Systems Proceedings*, 2006 International Conference, Volume: 2, 2006 , Page(s): 1337 – 1341.

20

Air Traffic Control Decision Support for Integrated Community Noise Management

Sander J. Hebly[1] and Hendrikus G. Visser[2]
[1]*National Aerospace Laboratory*
[2]*Delft University of Technology*
The Netherlands

1. Introduction

The noise resulting from flight operations at major airports is a continuing source of annoyance in nearby residential communities. This being recognized by the industry, new aircraft generations continue to be less noisy than their predecessors, but this development in itself does not solve the problem in a fast growing market. Therefore a range of mitigation measures has been implemented at airports located close to sensitive communities. Some of these measures, like (night) curfews, restrictions on flight numbers and noise pricing tend to control and/or shape the demand from the airport's point of view. A second range of measures, including the use of noise preferential runways, noise abatement routes and the use of low noise procedures aims at a reduction of noise impact without interfering with the supply of airport capacity or the demand for air traffic.

The implementation of noise abatement procedures at the side of Air Traffic Control (ATC) authorities is not always straight forward, as it may interfere with respect to safety and efficiency requirements. This can be observed when considering the current implementation of the Continuous Descent Approach (CDA). The trade-off for this procedure is either to accept a less than ideal continuous descent, or to accept a reduced arrival capacity (Davison Reynolds et al., 2006; Kershaw et al., 2000; Weitz et al., 2005). However, contradictory requirements are not the only problem that air traffic controllers face with respect to reducing community noise impact. Taking noise beneficial decisions can also be difficult because of a lack of noise-related information. Usually, controllers have access to 'static' information, like the preferred use of certain routes and runways. However, they are not provided with information on the continuously developing situation with respect to community noise exposure. This means for example that they cannot respond to developments in the noise exposure in the past or expected developments in the near future. Nor can they evaluate the environmental effects of a tactical or operational decision they are about to take.

This paper presents a concept for integrated community noise management in the form of a decision support system (DSS) for air traffic controllers. It should assist controllers in guiding arriving and departing traffic near airports in a safe and efficient matter, making use of the future concept of four-dimensional trajectory-based operations and future technology currently under development. The system should be able to create conflict-free

or de-conflicted, individually customized and optimized trajectories for all arrivals and departures. While doing so, the system minimizes the negative environmental effects of the flight operations and manages their spatial allocation, both for individual movements and cumulative exposure.

2. Towards trajectory based operations

Trajectory based planning is a relatively novel concept. Both the European Single European Sky ATM Research (SESAR) program, as well as the US Next Generation Air Transportation System (NextGen) program envision the transition to trajectory-based operations (TBO), based on the four-dimensional trajectories (4DT) of the aircraft. The TBO concept, including digital data exchange between aircraft and the Air Navigation Service Provider (ANSP) is expected to replace the current way of operating based on flight plans, resulting in a greatly reduced uncertainty with respect to the future (forecasted) position of an aircraft in flight (Joint Planning and Development Office, 2007). This is not only true for the spatial position in three dimensions, but also for the expected time along the different positions of the trajectory. For the NextGen program, this is achieved through the concept of Controlled Time of Arrival (CTA), a time window in which the aircraft is expected to cross a certain waypoint.

The NextGen Concept of Operations identifies that there will be different types of operations in TBO airspace. For example, oceanic airspace operations are managed in a different way than operations into or out of an airport. The airspace around the airports is expected to be managed by the ANSP, taking responsibility for both trajectory management as well as separation management. Ideally, arriving aircraft are assigned a 4DT trajectory at the top of descent that does not employ the current practice of low-altitude path stretching and holding.

The primary reason for shifting towards TBO is to increase efficiency and airspace and airport capacity. This can be achieved because it allows for removing additional separation that is the result of the current lack of control precision and behaviour predictability. The result of the increased predictability is that the tasks of the air traffic controllers can shift from a more controlling or operating task to a more supervisory, planning-oriented task, supported by sophisticated automation tools. Together, the accurate position forecasting possibility and the planning-oriented task of the air traffic controller will also greatly improve the possibility to manage the geographical allocation of environmental effects with respect to individual movements. This means that the already foreseen transition towards TBO will provide a unique opportunity to combine environmental management with the traditional responsibilities of air traffic control. Especially if such a system would enable aircraft to fly individually customized and optimized trajectories, multiple benefits can be identified, as discussed in the next section.

3. Concept and benefits of integrated environmental management

In the concept of integrated environmental management, the meaning of the phrase ‘integrated’ is twofold. First, it is used to indicate that all efforts towards environmental impact reduction are managed concurrently and consistently, yielding a more effective approach (Clairbois, 2005). Second, it refers to the integration of environmental management

into the air traffic control processes through the use of advanced decision support systems. Both concepts and their advantages are discussed in this section.

Noise mitigation efforts can be categorized into different levels of aggregation. The design of noise preferential routes and/or procedures is at a lower level than the actual use of them in an operational environment. Even higher are activities in the strategic range such as noise allocation efforts in relation to noise zoning and land use planning, for example through changing runway use preferences (Hebly & Wijnen, 2005; Galis et al., 2004). Currently, these efforts do not always take place simultaneously and are often not managed by a single party. For example, departure procedure design is largely an international affair, aimed at the development of standardized operating procedures, such as the ICAO-A and ICAO-B departures. Subsequently, it is the ANSP that is responsible for selecting one of the procedures to be used for a particular airport. However, if the chosen procedure is aimed at reducing community noise exposure, it should match the land use planning policy, which is governmental responsibility. For this specific example, it turns out that three different stakeholder groups are involved, which may each have different objectives. Second, although the reasoning behind standardization is clear in today's operational environment, it also means that these procedures are not optimized with respect to the local demographic situation.

In the ideal situation, all environmental impact mitigation efforts at all levels should be managed concurrently. When using such a form of integrated environmental management, it can be ensured that all actions taken to minimize the nuisance caused by aircraft noise and emissions will be consistent, complement each other, and make use of synergy benefits. At the same time, it helps avoiding that a certain decision (partly) reduces the effectiveness of another measure at a different level or made by a different stakeholder.

When making use of a DSS for consistent environmental management, the basis for all decisions involving trade-offs with respect to noise and emissions should be the actual situation around the airport. This requires a detailed model of the surrounding areas, preferably not limited to static population density only. A more dynamic representation of the location and activities of people should be used, as a lot of people do not spend their day at home. If desired, this information can be combined with building data, also including (estimated) sound transfer loss, allowing for a much more accurate estimation of actual noise exposure. Sound proofing programs directly influence the indoor noise exposure and can in turn influence noise allocation considerations. The model should keep track of previous noise exposure allowing it to take this into consideration during future noise allocation decisions, and it should be aware of current local air quality. Optionally, noise from other sources (other traffic modes and industry) could be regarded as well, if this is desired.

On top of the model of the surrounding areas, the responsible governmental body should set the policies with respect to their environmental objectives. Without such objectives, decision making is often hampered by occurring conflicting interests. For example, changes in aircraft routings can be beneficial for a lot of people if a certain residential area is avoided. However, such a change typically comes at the cost of increased exposure in other areas. Even if the area experiencing increased exposure would be completely uninhabited, still conflicts may arise because of the commercial, recreational or wildlife preservation functions that that area may have. The government imposed policies are required to settle the conflict in these situations.

Together, the airport surroundings model and the government policy model function as an additional input for the trajectory synthesis process. This leads to the situation where environmental considerations are directly present at the operational level of air traffic control. The trajectory synthesis process should eventually be capable of handling both arrival and departure traffic simultaneously, while applying a multi-objective trajectory optimization algorithm searching for conflict-free trajectories, optimized with respect to efficiency, fuel burn, environmental and possibly other objectives. At the same time, the system is also responsible for sequencing of traffic, runway assignment and managing cumulative noise exposure in the area around the airport.

The use of TBO alone, thus without an interrelated environmental management system, can also provide several environmental benefits. For example, it should enable the possibility to perform high navigation precision, Continuous Descent Approaches (CDA) during all traffic demand situations. However, integrating the environmental management into TBO air traffic management also gives several benefits on top of the advantages that can be accredited to TBO alone. Most of the advantages stem from the possibility to decrease uniformity and increase flexibility by designing a trajectory for each individual flight.

Different aircraft have different performances, not only concerning flight performance, but also concerning noise and emissions. Even two aircraft of the same type and with the same systems installed may show different behaviour, mostly because of different instantaneous weights and atmospheric conditions. This is inevitable and does not have to be a problem, but it currently results in two disadvantages. First of all, for most airports, the procedures for arriving and departing traffic are designed such that at least the great majority of visiting aircraft should be able to adhere to the procedures under a wide range of weather and wind conditions. Basically, this means uniform design for the weakest link, possibly inhibiting better performing aircraft (in any sense) to exploit their capabilities. Second, differences in flight performance may reduce airport capacity. This is most evident when considering two consecutive aircraft flying the same trajectory with a different speed profile. The difference in speed will at some point result in an unnecessary large gap, which is basically a waste of capacity. A situation where each aircraft would be flying its own customized and optimized trajectory can eliminate both disadvantages (Vormer et al., 2006). It allows for optimizing for individual performance, and it can prevent aircraft with different speed behaviour to fly the same trajectory. This may not be necessary and/or desirable all of the time, but could be employed during peak hours if capacity is critical for the airport under consideration

Flexible use of airspace (FUA) is currently being implemented in ECAC states, including sharing airspace between civil and military users. When used, segregation of traffic is temporary, based on real-time usage within a specific time period. The concept of individually designed and assigned trajectories matches very well with the concept of flexible use of airspace. Areas can be closed down on a rather short notice by no longer issuing trajectories through that area, or even updating already issued trajectories to clear an area as fast as possible. Although this is in fact again an advantage of TBO itself and has nothing to do with environmental considerations, the same principle can be used in that sense. In the Netherlands, several temporary restricted areas exist in order not to disturb memorial ceremonies with loud aircraft noise. Since the proposed DSS can perform trajectory synthesis with access to noise information, it can easily take such temporary restrictions into account. Please note that also offers the opportunity to transforms the

current spatial restriction into real noise restrictions, which means that individual noise performance can be taken into consideration.
When addressing both departing and approaching traffic simultaneously in the trajectory planning process, it should be possible to apply a less strict procedural segregation between these two flows. Based on previous research, it is expected that this results in fewer altitude restrictions for departing traffic (Jung & Isaacson, 2002). This can in turn reduce noise exposure, emissions and fuel burn, which is not only beneficial for the residential community, but of course also for the airline itself.
With respect to gaseous emissions, the actual spatial allocation currently receives less attention than total airport related emissions. However, from a health perspective, local air quality is far more important than total airport related emissions. When using flexible routing, this also offers the possibility to influence the air quality to some extent. Although one should consider that aviation as a source only has a limited share in the resulting air quality, it may for example be possible to avoid certain areas in the trajectory synthesis process if that area is experiencing air quality problems at that time. Similarly, departure procedures can be chosen such that the emission of a particular substance is minimized if the concentration of that substance if critical at that time. Both options are dynamic air quality measures and are comparable to concepts such as adapting highway speed limits based on actual measured or predicted air quality (Spit & Sluis, 2006).
Finally, depending on the development of new noise models and the availability of more accurate real-time atmospheric condition information, it might be possible in the future to take current conditions such as wind and temperature gradient into consideration in the noise propagation modelling, resulting in more accurate noise predictions. If it is possible to perform such calculations in a timely manner in an operational environment, this would allow for a trajectory synthesis process based on more realistic noise modelling. In other words, trajectories may be adjusted not only for actual atmospheric conditions with respect to flight performance of the aircraft, but also with respect to the actual weather related noise propagation properties.

4. Arrival management as interim concept

It is important to realize that the ideal environment for the DSS is currently not in place. First of all, high accuracy 4D navigation needs to be available to all or at least the great majority of aircraft. The same is true for the required 4DT exchange functionality. There are, however, research projects that look into elements of the proposed system without relying on the 4D technology. These projects look into modifying currently existing or planned automation tools for arrival management. These tools help controllers in creating an efficient flow of aircraft towards the runway, eliminating delay as much as possible. Often, the resulting trajectories from these tools are basically small variations to existing arrival patterns, in order to achieve a certain amount of delay required for a safe and efficient flow.
The Center-TRACON Automation System (CTAS) terminal area tools used at some airports in the US include tools used for arrival management. These tools are capable of generating advisories that respect separation requirements and minimize delay. A suggestion has been made to inject noise related information into the current constraint resolution and scheduling logic, to create a system that is capable of generating advisories with respect to both delay and noise (Capozzi et al., 2002, 2003). The resulting concept is called the Noise

Avoidance Planner (NAP). However, because of the current CTAS constraint resolution architecture, noise considerations and efficiency cannot be addressed simultaneously.
Although only altering arrival management tools cannot offer all of the benefits as previously identified, it can be seen as a step in the right direction in the absence of the technology required for 4DT. The remainder of this paper will focus on research done concerning an arrival management model where noise and efficiency are considered simultaneously (Hebly, 2007). Please note that it is not designed or build to function as an operational arrival manager, but purely to study the effects of adding noise objectives to the otherwise delay driven support tools for sequencing and scheduling, and to identify the interaction between the noise and efficiency objectives.

4.1 Arriving traffic concept

As long as trajectory exchange is not a possibility, the method of issuing vectors or defining a limited number of fixed arrival trajectories are the two remaining options for controlling the lateral part of the trajectories of arriving aircraft. Fixed routes have the advantage that they can be designed as noise optimal routes, can be flown with high navigation precision and at the same time allow for more optimal CDA procedures. The downside of fixed routes is that the controller looses the path stretching possibility. This means that any required delay should be absorbed before aircraft start their assigned fixed arrival route, apart from the fine tuning that might be achieved using speed control.
The arrival management model used here can deal with multiple fixed routes to a single runway of an airport. The configuration presented in this paper is based on three routes towards one of the runways of Amsterdam Airport Schiphol. This is depicted in figure 1.

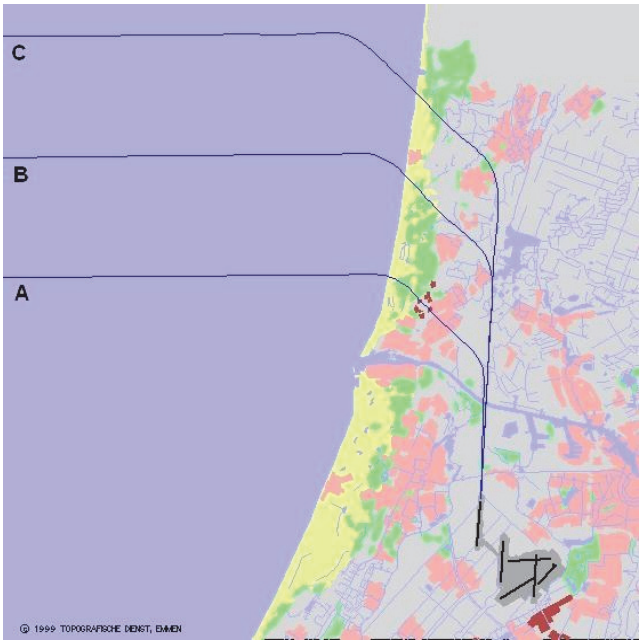


Fig. 1. The three arrival routes used

For this scenario, it is assumed that all traffic from the West is guided to this runway. Traffic from the East is not modelled, but is assumed to land at another runway independent of the modelled part. Of the three routes, the last part of the center route (labelled B) is very similar to the current night-time CDA for this runway. The two outer routes (A and C) are variants of the route in the sense that they cross the coastline either further away or closer to the runway. Please note that in reality, these routes do not exist and the use of fixed arrival routes is currently limited. For the model, all approaches are assumed to be CDA procedures, from the points where the fixed routes start. As a result, the three different routes cross the more densely populated areas close to the coastline at different altitudes, before turning towards the runway over a less populated area. This is expected to result in different noise exposure, except for the final part, where all three routes are equal.

Before a flight movement starts on one of the three depicted fixed approach routes, it is assumed that it crosses one of two available metering fixes, situated north-west and south-west of the starting points of the fixed approach routes, see figure 2. Since all three routes can be used from both fixes, this may lead to crossing traffic. The model does not regard separation before traffic is on the depicted routes. Therefore, possible conflicts in the area between the gates and the approach routes must be solved using altitude constraints.

The arrival model is provided with a traffic sample. The sample specifies aircraft type, metering fix and expected, (undelayed) time of crossing time of the metering fix. Furthermore, the model is aware of the (undelayed) transit times from both fixes to the runway threshold via the three different routes. These transit times have been determined using the NLR ATC simulator (NARSIM) for different aircraft types.

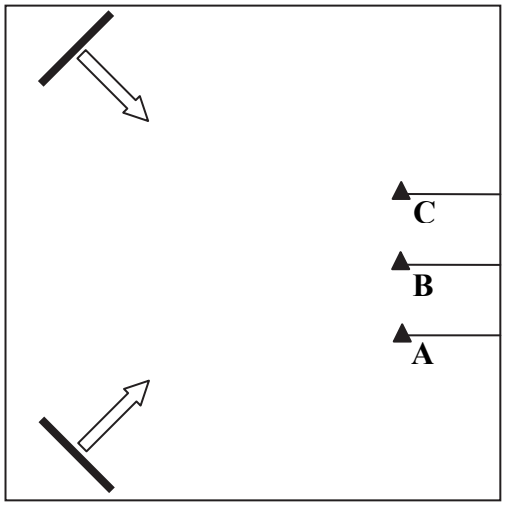


Fig. 2. Metering fixes and starting points for the CDA approaches

4.2 Noise metrics and indicators

In order to allow the model to perform routing selection based on noise criteria, it requires information on the noise exposure resulting from the selection of a route. The single event

noise exposure of flying the three different trajectories can be expressed using several metrics or indicators based on those metrics. Noise levels are typically computed at a large number of points within the area of interest. Although this type of result is suitable to compare different results graphically, it is not fit for a numerical comparison. Indicators that are derived from the metrics however can be seen as an aggregate of the result, often expressing the result using a single number. This enables easy comparison between the results of the different routes and aircraft types, but as with all aggregated data, a part of the original data is lost. Checking for consistency among multiple indicators is therefore always a safe option. Four different single event indicators based on two different metrics will be presented here.

The World Health Organization (WHO) recommends to limit the number of maximum A-weighted indoor peak level (LA_{max}) events of 45 dB(A) or more during night-time in the bedrooms (Berglund, 2000). This corresponds to a 60 dB(A) level outdoors, when assuming a sound transmission loss of 15 dB, a modest value that allows for people to sleep with the windows open. Based on this recommendation, the number of people exposed to higher peak levels during an aircraft flyover is a suitable single-number indicator for undesirable night-time noise events. This number is used as the first indicator. For day-time noise, there is no similar recommended or often applied limit. However one could argue that, based on the 10 dB penalty that is often applied to night-time events for the cumulative metrics, a daytime 70 dB(A) LA_{max} limit is equivalent to a 60 dB(A) night time limit. Therefore, the number of people exposed to peak levels higher than 70 dB(A) is used as the second indicator.

Another option is to use dose-response relations to estimate the effects of noise exposure. Therefore, the third indicator is based on the relationship as proposed by the Federal Interagency Committee on Aviation Noise (FICAN) in 1997 (Federal Interagency Committee on Aviation Noise, 1997). It represents an upper bound on the percentage of people likely to awake due to a flyover, where the percentage of awakenings is a function of the indoor Sound Exposure Level (SEL, L_{AE}), see figure 3. For this function, a sound transmission loss value of 20.5 dB is used, as an average value for a typical home.

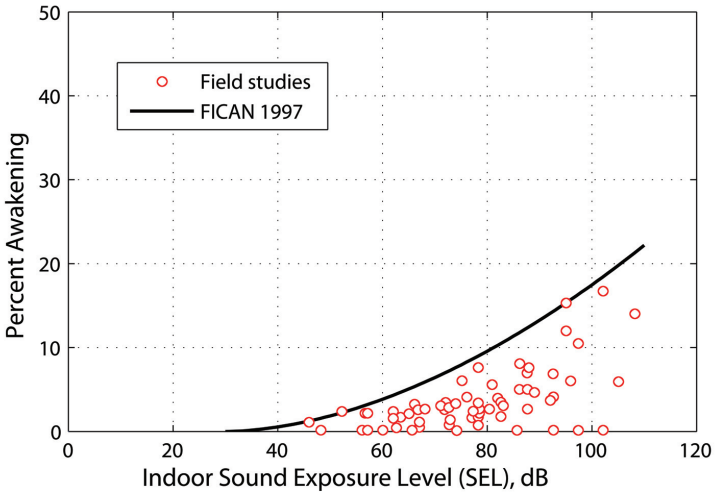


Fig. 3. FICAN proposed sleep disturbance dose-response relationship (Federal Interagency Committee on Aviation Noise, 1997)

As a fourth and final indicator, an estimation of the number of complaints will be used. This relation is based on Dutch research that relates the number of complaints concerning a flyover to its (computed) maximum noise level (Lieshout, 2006). It turned out that the different communities around the airport show different complaint rates. However, because the study was performed for a limited number of communities, here the complaint rates for the most sensitive community are used for the complete study area. Therefore, the number should be interpreted as a worst bound on the expected number of complaints due to an aircraft noise event. The resulting dose-response relationship is given in table 1.

LA_{max} dB(A)	Complaint rate (per 1000 inhabitants)
$x < 50$	None
$50 \leq x < 60$	0.130
$60 \leq x < 70$	0.437
$x \geq 70$	1.269

Table 1. Exposure-response relationship for expected complaints

Using the Dutch noise computation model, the results for different aircraft types and the three different routes have been computed for the four single event indicators. The results are presented in table 2 for a limited number of aircraft types. Please note that aircraft types are categorized in the Dutch noise model based on their noise performance. This may lead to the same results for different types, like for the Airbus A320 and the Boeing 737-800.

Indicator	Route	Aircraft type				
		CRJ700 Dash 8-400	A320-200 B737-800	A330-300 B767-400	B777-300	B747-400
1 60 dB(A) LA_{max}	A	228	545	30350	24030	50335
	B	178	1045	7108	4845	24655
	C	178	1045	8448	6945	25480
2 70 dB(A) LA_{max}	A	35	73	298	348	700
	B	35	78	308	345	1095
	C	35	78	308	345	1093
3 FICAN awakenings	A	445	773	1860	1672	2756
	B	287	470	1896	1560	2929
	C	321	501	2483	1946	3909
4 Expected complaints	A	3.22	5.83	18.82	16.33	26.86
	B	0.97	2.66	11.84	9.55	21.72
	C	1.29	2.78	13.18	10.42	26.78

Table 2. Single event noise indicators for the three approach routes

When analyzing the different results, the first observation is that the 'inhabitants within the 70 dB(A) contour' metric can hardly make a distinction between the three different approach routes, except maybe for the Boeing 747. This means that it is not fit for our purpose and is

further discarded. Not very surprising, the remaining indicators show that exposure increases with aircraft size, with the exception of the B777-300 category performing slightly better than the smaller A330-300 / B767-400 category. Concerning choice of routing, the B-route generally performs best, with a second place for the C-route. Again some exceptions can be observed, especially for the B747-400.
Instead of selecting one of the single event metrics for the arrival management model, it is also possible to use a composite function of the three remaining indicators, by summing them up. However, because of the very large differences in absolute number, some sort scaling is required to prevent one indicator from dominating the other ones.

4.3 Optimization model description

The scheduling problem is stated as a Mixed Integer Linear Programming (MILP) problem. This approach allows the problem to be formulated in generic form using (algebraic) constraint equations instead of designing a dedicated algorithm. Apart from the clarity of using equations, it also allows for easy changes or additions to the model, such as changing the goal function.
The sequencing is based on the existing principle of Constrained Position Shift (CPS), where an aircraft is allowed a difference of n positions between First Come First Serve (FCFS) order and the actual landing order (Balakrishnan & Chandran, 2006). When using FCFS, all aircraft land in order of their scheduled arrival times at the runway. When using CPS, an aircraft that is for example fourth in the FCFS sequence, is allowed to take the landing positions 3, 4 and 5 when n = 1. For a sequence of four aircraft, this leads to a decision tree as depicted in figure 4. Aircraft 1 and 6 do not join the sequencing process. Aircraft 1 can be thought of as the last aircraft that already has a fixed or frozen landing position and time. It prevents the aircraft taking position 2 from landing earlier than possible, based on the landing time of aircraft 1 and the required separation. The last aircraft, number 6 in this example is not necessarily a real (future) aircraft, but is mainly used to prevent the scheduler to push heavy aircraft to the back of sequence. Without this additional aircraft, the scheduler might do so because it does not regard the required separation behind the last aircraft in the sequence. Adding the dummy aircraft automatically adds the required separation behind the last real aircraft

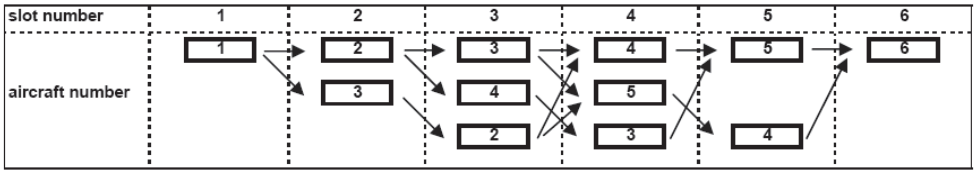


Fig. 4. Decision tree for the example problem

Next to the scheduling, there is also the route selection process. The scheduler is forced to choose exactly one of the three offered approach routes for each flight. The route selection determines the noise score for a specific flight as discussed in the previous section, as well as the earliest possibility to land. For example, a Boeing 737-800 approaching via the southern metering fix cannot land earlier than the time required to reach the fix plus 737 seconds when using route A, 807 seconds when using route B or 867 seconds when using route C.

Aircraft are prohibited to land earlier than their predecessor in the sequence and are also required to respect a minimum separation time based on the wake vortex category of the pair under consideration. The separation values used are 95 seconds minimum following a medium aircraft, 125 seconds minimum for a heavy aircraft following a heavy one and 155 seconds minimum for a medium following a heavy.
Due to the constraints described above, it is very likely that an aircraft is scheduled to land later than its earliest possible landing time. The difference - that is the delay - needs to be absorbed at some point. The model does not look into how the delay is absorbed; it only calculates the required amount. In practice, delay can probably best be accommodated before crossing the metering fix.
Finally, the objective function for this problem is defined as:

$$\text{Min}: \sum_{j=1}^{m} LT_j + k \sum_{j=1}^{m-1} NE_j \tag{1}$$

where LT_j is the landing time of the jth aircraft, k is the noise cost multiplier and NE_j is the noise exposure of the jth aircraft. The landing time is expressed in seconds from the instant the schedule is created, and is used as a proxy for delay. The noise exposure itself is formulated as:.

$$NE_j = a_j \cdot NE_{j,r=A} + b_j \cdot NE_{j,r=B} + c_j \cdot NE_{j,r=C} \tag{2}$$

$$\begin{aligned} a_j + b_j + c_j &= 1 \\ a_j, b_j, c_j &\in \{0,1\} \end{aligned} \tag{3}$$

where $NE_{j,r=A}$ is the noise exposure of aircraft j when using approach route A, etc. As can be seen, the noise cost for the last aircraft is excluded from objective function, since this is not a real aircraft to be scheduled. Its flying time is included on the other hand, because of the reason the aircraft was added in the first place. The noise cost and flying time for the first aircraft are taken into consideration, although the scheduler will not be able to optimize for these values, since they are already fixed. Noise cost multiplier k determines the importance of the noise related performance relative to the delay related performance. When k equals zero, noise exposure is not regarded at all, turning the optimizer into a traditional, delay driven only tool. When k is very large, the optimizer will still generate an optimal landing schedule, but the routing process is completely dominated by noise considerations.
The problem itself is generated by a script that reads the input variables, and writes the mathematical formulation for the problem. This can than be solved by a solver such as ILOG CPLEX (commercial) or LP_SOLVE (open source). Finally, the solution as returned by the solver is post-processed for ease of interpretation. In the post processing, the solution is also converted to a traffic file for NARSIM. Using this file, NARSIM can be instructed to 'playback' the solution on a radar screen, making it very easy to visualize, check and interpret the results.

4.4 Scheduler results

Scheduler results and the trade-off between average delay and noise exposure are shown in figures 5, 6 and 7 for 20 arrivals in a mix of 30% heavy and 70% medium aircraft. Figure 5 is

based on an arrival rate of 45 aircraft per hour, which is higher than the runway capacity. Figure 6 is based on the same traffic sample of 20 aircraft, but arriving at a rate of 36 aircraft per hour and figure 7 is based on an arrival rate of 30 aircraft per hour. All figures show the average delay per operation and the resulting (combined) noise exposure index (NEI), both against the noise cost multiplier k. This multiplier is varied between 0 and 200. The average delay can also be compared to the average delay that is achieved when using FCFS. The FCFS solution is based on time optimal routing only, so it does not regard noise exposure at all.

From the results, it can be concluded that adding noise considerations to the model does indeed reduce the noise exposure indicator. Furthermore, increasing the importance of the noise objectives relative to the efficiency objectives, the noise indicator value can be reduced further, at the cost of increased delay, possibly leading to a solution that is worse than the reference FCFS solution. Of course, in such a situation, the noise exposure of the optimized solution is lower than that of the FCFS solution. Interesting to note is that a (small) noise improvement can be achieved without an increase in delay. This can be seen in all three figures by looking at the differences in solutions resulting from k = 0 and k =1. Apparently, routing can sometimes be changed in favour of noise without affecting the sequence and the schedule.

When comparing the three figures, the effect of the arrival rate can be seen. In the situation where the arrival rate is below the runway capacity, it can be seen that reducing the noise exposure indicator easily leads to solutions that are worse than the FCFS solution in terms of efficiency. At arrival rates higher than the runway capacity, the situation is clearly different. In this situation, all aircraft need to be delayed. For efficiency, it does not matter whether the required amount of delay is absorbed before the metering fix or during the approach by using a different route. This allows the scheduler to assign longer, but noise optimal routes without effecting landing times and runway throughput.

Apart from the results shown here, which are based on the expected complaints indicator, additional results have been generated for the other optimization criteria. Similar results are obtained when using the other two indicators, as well as with an indicator based on a combination of the three. More results have also been generated using higher and lower arrival rates and different traffic samples, all showing different results of course, but similar trends.

4.5 Cumulative noise exposure

Instead of single event metrics or indicators, true community noise exposure is often based on cumulative exposure metric, such as the day-night level (L_{dn}) or the day-evening-night level (L_{den}). Both metrics describe weighted average noise levels, where both apply a 10 dB(A) penalty for night time events, but only L_{den} applies a 5 dB(A) penalty for evening noise as well. However, when assigning routes to aircraft based on the single event noise indicators only using a scheduling, sequencing and routing method as showed above, it is likely that one route is used exclusively, especially when the noise cost multiplier is high. When only using single event indicators, a route that is optimal for a certain flight, is still optimal 50 flights later. This can easily lead to an extremely high exposure for the area under that specific route, resulting in a situation that is considered unacceptable.

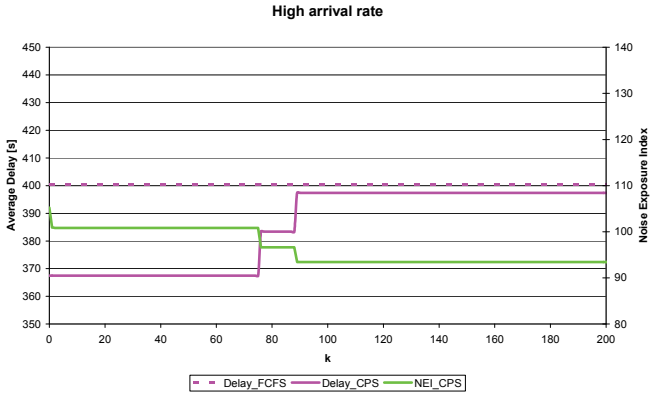


Fig. 5. Scheduler results for an arrival rate higher than the runway capacity

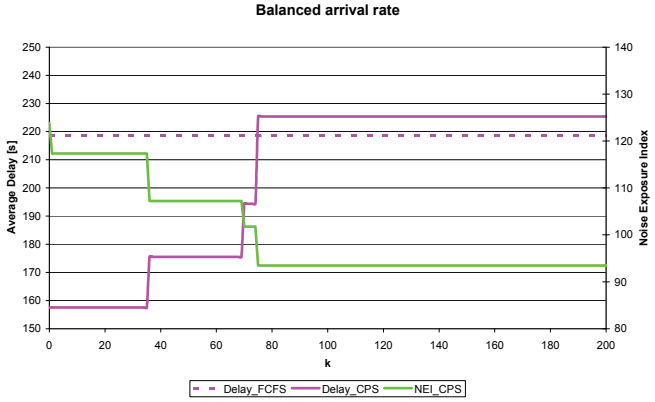


Fig. 6. Scheduler results for an arrival rate near the runway capacity

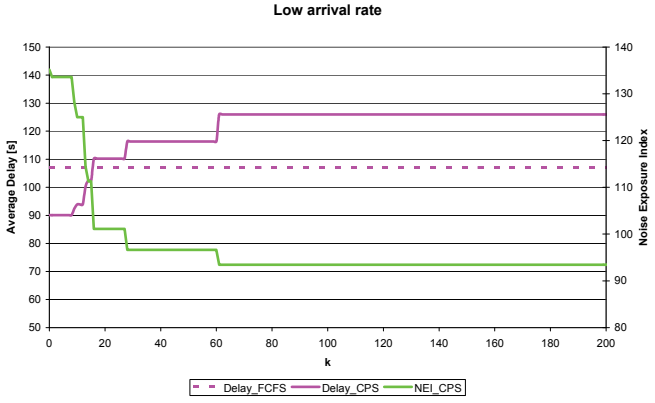


Fig. 7. Scheduler results for an arrival rate lower than the runway capacity

An obvious solution is to add a cumulative exposure indicator to the problem. A well-known indicator based on either L_{dn} or L_{den} is the number of people annoyed by aircraft noise, based on the relations as established by Miedema (Miedema & Oudshoorn, (2001). These dose-response relationships predict the long term average percentage of people annoyed by aircraft noise, based on the L_{dn} or L_{den} levels. Combining this information with population data, this leads to the single number annoyance indicator.
Whether using a single event or cumulative exposure metric, the scheduler will still need to make decisions on a per flight basis. A possible setup is to use the cumulative exposure of a past period of certain duration, and calculate the increase in exposure due to the flight currently under consideration. Based on the difference, the rise in total population annoyance can be computed, resulting from the marginal contribution of that movement. The first problem with this approach is that the dose-response relations have been established for long term average and stabilized exposure. As such, the additional annoyance calculated from a single flight, is certainly not guaranteed to be near the actual increased annoyance due to single flyover, if such increase could be quantified in the first place. However, when aware of this limitation, it can still be used to compare different alternatives.
A more fundamental problem lies in the behaviour of the dose-response relation itself. This is illustrated in figure 8. The annoyance percentage is a function of L_{dn} or L_{den} in dB(A) and is plotted against decibels on a linear scale, as in the left part of the figure. Plotted like this, the function appears to be convex. This would be desirable for our course, because when minimizing for annoyance, the increasing slope would result in traffic being directed away from the areas were exposure is already high. However, the function can also be plotted against a number of noise events, say the number of annual noise events of 90 dB(A) SEL each, as indicated in the right part of the figure. Here the function turns out to be concave, resulting in exactly the opposite behaviour: as soon as a certain area is experiencing high noise levels, annoyance is hardly increased by adding more flights. This cannot only be observed when examining this particular dose-response relation. When plotting other dose-response relations - like the one as established by Schultz (Federal Interagency Committee on Aviation Noise, 1997) - against the number of noise events, the same observation can be made. Apparently, total community annoyance can be minimized by maximizing exposure in the least sensitive area.

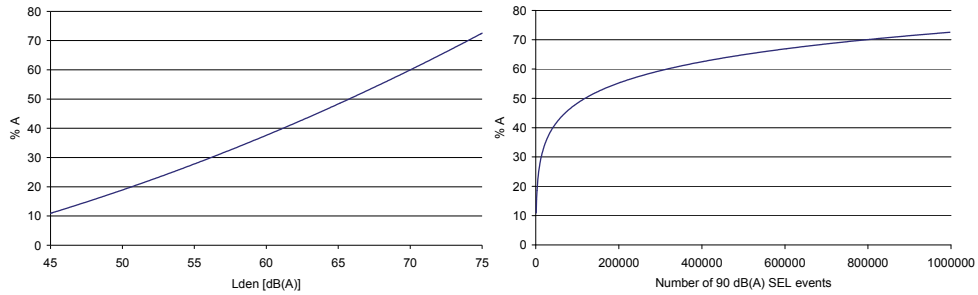


Fig. 8. The Miedema dose-response relation plotted against Lden in dB(A) as well as against a number of annual 90 dB(A) SEL events

Summarizing, when flight movement concentration results in unacceptable (or 'unfair') cumulative noise levels, adding indicators based on cumulative exposure annoyance does not solve the problem. An alternative is to set maximum allowable levels. When enforcing these maximum level limits by adding constraints to the scheduling problem, traffic will be redistributed in order not to break the limits. If cumulative level limits are defined as annual maxima, it can be desirable to derive a daily or hourly allowance, based on the year-to-date exposure. This can prevent very high exposure in one part of the year compared to a very low exposure in the remaining part, or vice versa.

5. Conclusions and future work

The foreseen transition towards TBO offers a unique possibility to integrate environmental management into the actual ATC process. When using such a form of integrated environmental management, it can be ensured that all actions taken to minimize the nuisance caused are consistent. At the same time, this concept allows all aircraft to fly trajectories that are optimized with respect to several objectives, including airport and airline efficiency and environmental ones.

Since the envisioned concept is years or even decades away from realization, adding noise considerations to the objective function of current or near future arrival managers appears to be an attractive interim solution. Based on the results of a model using the concept of fixed arrival routes in combination with CDA procedures, a small improvement in noise exposure can be achieved even without sacrificing efficiency. The noise indicators can be reduced further, but only when allowing increased delay. For low traffic situations, this can easily lead to situations that are worse than the FCFS solution in term of delay, but for heavy traffic situations, the trade-off is more advantageous.

Based on these results, it appears advisable to incorporate noise information in the arriving traffic scheduling process, even when sacrificing efficiency is deemed unacceptable. Further research will be conducted to look into the possibility of adding noise information to the performance indicator of a real arrival manager, as well as the effect of incorporating annual cumulative noise exposure limits.

6. References

Davison Reynolds, H.J.; Reynolds T.G. & Hansman R.J. (2006). Human factors implication of continuous descent approach procedures for noise abatement. *Air Traffic Control Quarterly*, Vol. 14, No. 1, (January 2006), pp. 25-45, ISSN 1064-3818

Kershaw, A.D.; Rhodes, D.P. & Smith, N.A. (2000). The influence of ATC in approach noise abatement, In: *Proceedings of the 3rd USA/Europe Air Traffic Management R&D Seminar,* 06.01.2010, Available from http://atmseminar.eurocontrol.fr/past-seminars/3rd-seminar-napoli-italy-june-2000/papers/paper_107/view

Weitz, L.A.; Hurtado, J.E. & Bussink, F.J.L. (2005). Increasing Runway Capacity for Continuous Descent Approaches Through Airborne Precision Spacing. *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, ISBN-10: 1-56347-737-8, San Francisco, California, USA, August 15-18, 2005

Joint Planning and Development Office (2007). Concepts of Operations for the next generation air transportation system, version 2.0, 06.01.2010, Available from http://www.jpdo.gov/library/NextGen_v2.0.pdf

Clairbois, J-P.; Debroux, Ph. ; Meganck, D. & van der Poorten, D. (2005). Liege/Bierset airport, Brussels Bouth/Charleroi airport: an integrated approach of the noise impact of airports. *Proceedings of the Internoise Congress and Exposition on Noise Control Engineering*, Rio de Janeiro, Brazil, August 7-10, 2005

Hebly, S.J. & Wijnen, R.A.A. (2005). Development of a runway allocation optimisation model. *Proceedings of the Eleventh Australian International Aerospace Congress,* ISBN 0 85825-732-7, Melbourne, Australia, March 13-17, 2005

Galis, S.P.; Brouwer, M.A. & Joustra, T.J. (2004). Optimization of yearly airport capacity within noise limits at Schiphol Airport. *Proceedings of the 33rd International Congress and Exposition on Noise Control Engineering*, Prague, Czech Republic, August 22-25, 2004

Vormer, F.J.; Mulder, M.; van Paassen, M.M. & Mulder, J.A. (2006). Optimization of Flexible Approach Trajectories Using a Genetic Algorithm. *Journal of Aircraft*, Vol. 43, No. 4, (July, 2006), pp. 941-952, ISSN 0021-8669

Jung, Y.C. & Isaacson, D.R. (2002). Design concept and development plan of the expedite departure path. *Proceedings of the AIAA Aircraft Technology, Integration and Operations (ATIO) Technical Forum*, ISBN-10: 1-56347-554-5, Los Angeles, California, USA, October 1-3, 2002

Spit, W.F.M. & Sluis., C.M. (2006). Verkenningsstudie verkeersmanagement voor luchtkwaliteit, (in Dutch), 06.01.2010, Available from http://www.ipl-airquality.nl/data/DT225-2.007-rap-eindrapport.128.pdf

Capozzi, B.; Augustine, S. ;Thompson, T.R. & Robinson III, J.E. (2002). An Initial Assessment of Benefits for Noise-Aware Decision-Support Tools. *Proceedings of the AIAA Aircraft Technology, Integration and Operations (ATIO) Technical Forum*, ISBN-10: 1-56347-554-5, Los Angeles, California, USA, October 1-3, 2002

Capozzi, B.; Augustine, S. & Thompson, T.R. (2003). Dynamic Noise Avoidance Routing Through Terminal Airspace. *Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit*, ISBN 9040714061, Austin, Texas, USA, August 11-14, 2003.

Hebly, S.J.; Hanenburg, V.A.; Wijnen, R.A.A. & Visser, H.G. (2007). Development of a Noise Allocation Planning Tool. *Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit*, ISBN-10: 1-56347-889-7, Belfast, Northern Ireland, September 18-20, 2007

Berglund, B.; Lindvall, T. & Schwela, D.H. (2000). Guidelines for Community Noise. 06.01.2010, Available from http://whqlibdoc.who.int/hq/1999/a68672.pdf

Federal Interagency Committee on Aviation Noise. (1997). Effects of Aviation Noise on Awakenings from Sleep, 06.01.2010, Available from http://www.fican.org/pdf/Effects_AviationNoise_Sleep.pdf

Lieshout, R.B.T.; Veldhuis, J. & Balke, P. (2006). Klachtenanalyse 2005, (in Dutch), SEO Economic Research, SEO report nr. 915, Amsterdam, The Netherlands, 2006

Balakrishnan, H. & Chandran, B. (2006). Scheduling Aircraft Landing under Constrained Position Shifting. *Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit,* ISBN-10: 1-56347-802-1, Keystone, Colorado, USA, 2006

Miedema, H.M.E. & Oudshoorn, C.G.M. (2001). Annoyance from Transportation Noise: Relationships with Exposure Metrics DNL and DENL and Their Confidence Intervals. *Environmental Health Perspectives*. Vol., 109, No. 4, (April 2001), pp. 409-416, ISSN 0091-6765

21

A Conceptual Framework and a Review of Conflict Sensing, Detection, Awareness and Escape Maneuvering Methods for UAVs

B. M. Albaker and N. A. Rahim
UMPEDAC Research Centre, Faculty of Engineering,
University of Malaya, Kuala Lumpur,
Malaysia

1. Introduction

Because of the key characteristics of Unmanned Aerial Vehicles (UAVs), removal of pilot, UAVs will be highly suited for repetitive, dirty, and dangerous operations. A wide range of civil and military applications are being explored in the community (Clapper et al., 2007). As a result, UAVs are given serious considerations in worldwide, making them the next step in evolution of aviation. Whatever missions are chosen for UAVs, their number and use will significantly increase in future.

Currently, UAVs do not have convenient access to civil and military operation theatres due to their inability to provide an equivalent level-of-safety comparable to see-and-avoid requirements for manned aircraft. The current procedure requires a certificate of authorization be applied for every mission. Obtaining such an authorization may take more than a month. This lengthy process is not in line with increasing number of UAVs development. Therefore, an autonomous collision sensing, detection, awareness and avoidance system will be a key enabler for the integration of unmanned with manned aircraft in a shared airspace. The main objective of the Collision Avoidance System (CAS) is to allow UAVs to operate safely within non segregated civil and military airspace on a routinely basis. For this purpose, the UAV must be able to identify and be identified by the surrounding traffic.

The diversity of UAVs and their missions involve a wide-range of system operating concept. Current unmanned aircraft range in size from small hand launch vehicles to large fixed-wing UAV with a wing span similar to Boeing 737. In addition, some UAV autonomously, semiautonomous or completely guided by ground pilot. Furthermore, unmanned vehicles cruise speed, climb/dive rate, turn rate and operating altitudes are similarly varied. Therefore, many CAS methods were proposed to account for that variation and to ensure that the unmanned aircraft efficiently avoids other cooperative traffic while also avoids fixed and moving obstructions such as terrain, obstacles and no flying zones.

Numerous technologies are being explored in the community addressing CAS systems. Much of the research in collision avoidance methods for UAVs had been imparted from the air traffic management, maritime and mobile ground robot research communities. However, aircraft complicates the avoidance problem by added dynamic constraints that must be fulfilled for

adequate separation. Although large efforts have been done to address collision detection and avoidance problem to manned and unmanned aircraft, however there had been little survey and comparative discussion of the techniques and methods deployed to resolve conflicts.
Some efforts towards describing and understanding the differences among proposed approaches have been introduced in the literature. The majority of conflict detection and resolution methods review tried to highlight the differences among different methods. Warren (Warren, October 1997) conducted an evaluation among three conflict detection methods. Zeghal (Zeghal, August 1998) provides a review of the differences among force field collision avoidance methods. In the last decade, Krozel et al. (Krozel et al., 1997) and Kuchar and Yang (Kuchar & Yang, 2000) presented a comprehensive survey of conflict detection and resolution methods for manned aircraft. Current technology advances allow for innovative CAS systems to be more effective in reducing the number of collisions and utilizing airspace more efficiently. Those new systems need to be addressed and compared.
Recently, Utt et al. (Utt et al., 2005) addressed some of the lessons learned in development of a sense and avoid system for UAVs. Karhoff et al. (Karhoff et al., 2006) identified see and avoid requirements necessarily to avoid collisions and defined criteria specific to the warrior UAVs consistent with Federal Aviation Administration (FAA) guidelines. That is to obtain routine access to airspace. Lacher et al. (Lacher et al., 2007) investigated the challenges associated with UAV collision avoidance from a civil aviation perspective and presented results from MITRE's research addressing collision avoidance technologies and systems performance analysis. Albaker and Rahim (Albaker & Rahim, 2010b) proposed a generic collision avoidance system and presented a survey of some methods in the air traffic domain.
A little explanation is given in the literature addressing the problem of how complete collision sensing and avoidance system is functioning to solve conflicts. Towards addressing these problems, this chapter has three main goals: (1) To explore the fundamental concept of operation and presents up-to-date literatures review of the collision sensing, detection, awareness and avoidance methods those deployed for aircraft, especially for unmanned aircraft. (2) To introduce a conceptual framework to assist in the design of context-aware application in collision avoidance domain. This is done by providing a better understanding of what context is and how it can be used in the conflict resolution domain. (3) To categorize methods into what type it is designed as well as point out its advantages and disadvantages. Furthermore, this work identifies common issues that should be considered in avoidance systems design process.
The following sections are organized as follows: Firstly, the main functions carried by the collision avoidance system with an introduction on how to get knowledge of incoming threats are presented. Secondly, each function in CAS system is discussed in details, pointing out the significant researches done in each function. A context-awareness engine is explored as one of the functions in CAS system design. Next, the major design factors of collision avoidance systems are addressed. Finally, the developed trajectory escape maneuvering methods are classified.

2. Generic process functional model of collision avoidance system

A general collision avoidance system must detect and predict traffic conflicts. A conflict is defined as the event in which the Euclidean distance between two aircraft is less than the minimum desired separation distance. Collision avoidance system must be able to detect conflicting traffic in sufficient time to perform an avoidance maneuver and then propose a

course of action and maneuver so as not to create collision. Depending on the level of autonomy inherent in the UAV, these functions could fall into wide range from simple conflict detection and warning to full autonomous conflict detection and avoidance.
The basic idea of CAS is composed of two main phases. These phases are collision sensing and collision avoidance. Sensing operation involves monitoring the environment for any encounter including cooperative aircraft as well as stationary and/or moving obstacles in a shared airspace. As an example, a UAV performing an operation within a shared airspace segment that includes both manned and unmanned aircraft together with no flying zone, as depicted in Figure 1. When a UAV gets too close to any moving or stationary obstacles, less than a predefined protection zone, a potential collision will occur.

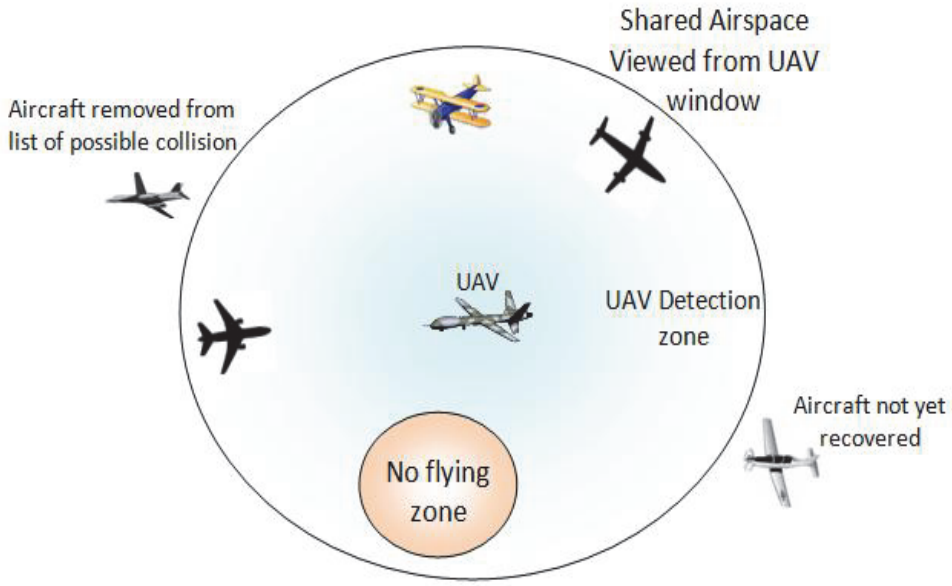


Fig. 1. A scenario illustrating shared Airspace as viewed from UAV

A fully autonomous CAS system is composed of six key functions. These functions include monitoring the environment, broad conflict detection, awareness, escape trajectory selection, maneuver realization and interception. Figure 2 illustrates the generic functional architecture of the CAS system for autonomous UAV.

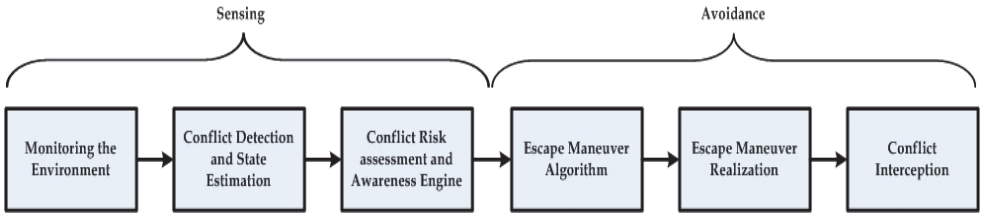


Fig. 2. Functional units implemented in the autonomous Sense & Avoid System

The sensing function refers to the ability of the system to monitor the environment and collect appropriate current state information for encounters, e.g. aircraft position, velocity and heading, about the environment surrounding UAV. This is done through the utilization of active and/or passive sensors and communication equipments.
The detection function is the ability of the system to acquire the sensed data, process it to extract useful information and discover collision risks to the UAV. Whereas, awareness function is used to dynamically projects the states into the future to check whether a potential conflict will occur in the near future or not. It also extracts the collision parameters in case of a potential conflict detected. In addition, it handles the process of when action should be taken.
The main role of avoidance function is to evade from a possible collision. This function will be invoked after detection of a near future collision. It determines how and what action should be performed. The maneuvering of the UAV will be performed based on the scheduled flight plan along with the level of responsibility assigned by the ground controller, which is further depends on the level of UAV autonomy (Asmat et al., 2006). Conflict interception handles the process of returning back to original UAV's course path after the conflicting object is resolved by the avoidance algorithm.
A fully automated collision avoidance system must address these six key functions, as stated earlier. For each function there were one or more design factor(s) that should be taken into account when consider selecting a suitable method for conflict resolution. The key functions together with its design factors will be discussed in the following sections. These factors represent principal categories by which approaches differ. Figure 3 shows the main and subdivisions of these factors

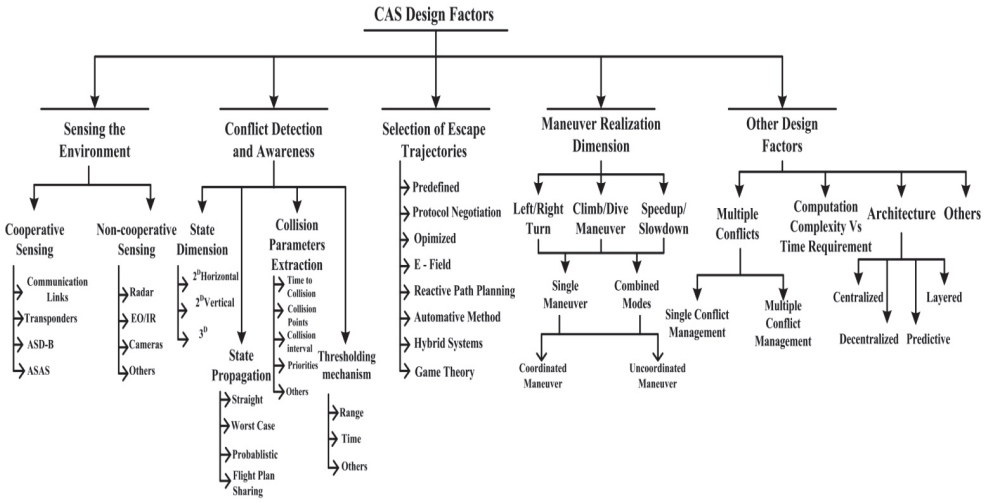


Fig. 3. Main CAS design factors.

Many methods have been proposed by various researchers to address collision avoidance problem. These methods have been developed not only for aerospace, but also for ground vehicles, robotics, and maritime applications. That is because the fundamental collision avoidance issues are similar across different transportation systems.

Some of the existing operational systems in use or which have been evaluated in the field are: Airborne Information for Lateral Spacing (AILS)(Waller & Scanlon, 1996), County Technical Assistance Service (CTAS)(Isaacson & Erzberger, 1997), Ground Proximity Warning System (GPWS)(RTCA, 1976), and its enhanced version (EGPWS) (Bateman, 1999), Precision Runway Monitor (PRM)(Federal_Aviation_Administration, 1991), Traffic alert and Collision avoidance system (TCAS) (RTCA, 1983; Ford, 1986; Ford & Powell, 1990; Committee147, 1997), Traffic and Collision Alert Device (TCAD)(Ryan & Brodegard, 1997), User Request Evaluation Tool (URET)(Brudnicki et al., 1997), and a prototype conflict detection system for Cargo Airline Association(Kelly, 1999). The other approaches range from abstract concepts to prototype conflict detection and resolution systems being evaluated and used in laboratories. Some approaches were developed for robotics, automobile or naval applications (Coenen et al., 1989; Iijima et al., 1991; Taylor, 1990), but are still not applicable to aviation (Chakravarthy & Ghose, 1998; Kuchar & Yang, 2000).

3. Monitoring the environment

The monitoring function refers to the ability of the system to provide traffic information about surrounding environment around unmanned aircraft. Determining the type of sensor that is appropriate for the UAV and environment is a challenging multidimensional problem. The fundamental information that a sensor or group of sensors need to acquire is the range, azimuth and elevation of all targets of interest (Lacher et al., 2007).There are a wide variety of sensors those were deployed for aircraft, which is mainly divided into two main categories: cooperative and non-cooperative traffic sensors.

3.1 Cooperative monitoring

Collision avoidance systems employed among UAVs usually assumes cooperative behavior in which inter-agent communication of position, heading, waypoints, and proposed trajectory is allowed. This is a common trait in collision avoidance methods for cooperative UAV systems, cooperative mobile ground robots and air traffic management systems.
Cooperative traffic Sensors includes all communication equipments those enable exchange information between the cooperative agents like position, heading, speed and waypoints. These devices like transponders mode S or emerging technologies like Airborne Separation Assistance Systems (ASAS) and Automatic Dependant Surveillance Broadcast (ADS-B) (Gazit, 1996; RTCA, 1997; Koencke et al., 1997; Holdsworth, 2003). As an example, ADS-B transfers the information: location, speed, UAV identification from UAV to other agents and Air Traffic Controllers. The data update rate received from other aircraft is important, so that the aircraft are working on timely data.

3.2 Non-cooperative monitoring

UAVs not fitted with such communication equipment may use non-cooperative traffic sensors to get knowledge about surrounding environment. In this case, the solution needs new sensors to replace communication links. Another case when a UAV use this kind of sensors to detect non-cooperative conflicts which include moving and/or stationary obstacles. Sensing the environment can be done in a variety of ways from the available technologies for non cooperative traffic including laser range finders, optical flow sensors Electro-Optical/Infra-Red (EO/IR), radar systems, acoustic or stereo camera pairs or moving single camera. Laser range finders are commonly used as an active sensor to detect

obstacles. The use of a single fixed laser range finder is of limited capabilities to detect conflicts in the environment. Moreover, this type of sensors is considered costly.
Alternatively, the utilization of radar system for active sensor detection is used to detect any moving/stationary obstacles whether they are cooperative or not. However, it is not used in small scale UAV due to its weight and size. Some of the research conducted in this field can be found in (Kumar & Ghose, 2001; Hyeon-Cheol & In-Kyu, 2004; Ariyur et al., 2005; Tatkeu et al., 2006; Kwag & Chung, 2007; Kemkemian et al., 2009). An efforts toward extracting radar parameters for continuous and interrupt driven data acquisition techniques were covered by Albaker and Rahim (Albaker & Rahim, 2009a; Albaker & Rahim, 2011b). Acoustic sensors can also be used for perceiving the target by passively listening.
As the advances of new powerful processing units, cameras can be used as a passive sensor to detect the obstacles around UAV. Many efforts are already being conducted to use camera in CAS systems such as the research found in (Matthies et al., 1998; Oh, 2004; Boon Kiat et al., 2004; Muratet et al., 2005; Mehra et al., 2005; De Wagter & Mulder, 2005; Zhihai et al., 2006; Ortiz & Neogi, 2006; Prazenica et al., 2006; Frew et al., 2006; Subong et al., 2008; Moore et al., 2009; Zufferey et al., 2010). Video cameras are light and inexpensive and thereby fit to the UAV requirements especially the small one. Video camera can be configured for obstacle detection as a stereo pair, or a moving single camera. However, video cameras provide information in a way that requires significant data processing for autonomous unmanned aircraft CAS system implementation.
Accurate monitoring and tracking of conflicting aircraft is an essential step in CAS systems. Erroneously identified conflicts would reduce the overall effectiveness of a CAS system. Simply, the accuracy of the information feed to the CAS system specifies how they are good. The accuracy of data available from sensors is limited and depends on the type of sensor used. The accuracy required also depends on aircraft packing density. The safety distance between aircraft can be increased to account for sensor inaccuracy.

4. Conflict detection and context awareness implementation

The detection function is the ability of the system to acquire the sensed data, process it to extract useful information and conflicts reporting to the UAV. The output of this function is to provide primary course collision detection in case of any intruder enters its detection zone around the UAV. If an encounter is detected, the detection function will acquire its state information and pass it to awareness function.
The needs behind context awareness engine comes from its importance for CAS algorithm where the UAV's surrounding environment context is changing rapidly. The goal behind introducing this function is to make interacting handle and manage conflict easier. In order to facilitate the building of context aware function, it needs to understand fully what constitutes a context aware application and what context is. The increase in mobility creates situations where CAS's context such as location of aircraft and cooperative/non-cooperative objects around it is more dynamic.
The detected conflicts, in the detection function, will be refined by reading out the reported threat position upon threat presence and project the states into the near future. In order to implement that requirement, the computing aircraft either tries to subscribe the conflicting aircraft in case of cooperative conflict detection or acquiring threats data registered by the sensing and detection functions in case for any moving and/or stationary obstacle. Depending on the type of conflict, different types of avoidance will be activated. The escape

trajectory generation and maneuver realization functions will be invoked after detection of future potential collision is detected. After resolving the conflict, the aircraft can return to its original course. Figure 4 illustrates the subunits that the collision detection and awareness should address.

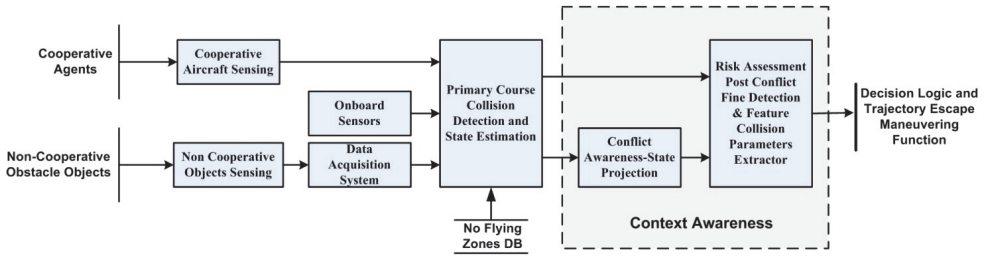


Fig. 4. A block diagram illustrating the tasks executed in the Conflict detection and awareness functions.

4.1 Context definition and context-awareness

Realizing the need for context is only the first step forward using it effectively in order to efficiently use context. A better understanding of what context is should be handled first. According to Webster's dictionary, context is the whole situation, backround or environment relevant to some happening or personality. This definition is too general to be useful in context-aware computing. Schilit et al. (Schilit et al., 1994; Schilit & Theimer, 1994) who first defined the term context-aware; refered to context as location, identities of nearby people and objects and changes to those objects. This definition is difficult to apply. When considering a potential new type of context information, it is not clear how the definition can help in deciding whether to classify the information as context or not (Dey et al., 2001).

Dey and Abowd (Dey & Abowd, 2000) defined context awareness as 'any information that can be used to characterize the situation of an entity, where an entity can be a person, place, or physical or computational object'. They went on to define context-awareness or context-aware computing as 'the use of context to provide task-relevant information and/or services to a user, wherever they may be'. Morse et al. (Morse et al., 2000) and Dey et al. (Dey et al., 2001) defined the context as any information that can be used to categorize the situation of entities whether a person, place or object that are considered relevant to the interaction between a user and an application themselves. Context is typically the location, identity and state of people, groups and computational and physical objects.

Based on these prior attempts to define context and following on from that, the context in conflict avoidance domain is defined as any information that can be used to characterize the situation around the UAV. Context-aware looks at who's, when's, where's and what's of entities and use this information to determine why the situation is occurring. The activity answers a fundamental question of what is occurring in the situation. Basically, the general model of context-awareness is divided into three main units, which are: generation, in which the contextual information is obtained from sensors and cooperative communication; processing, which process raw data acquired by the sensors and communication systems to obtain meaningful information and finally the usage, which use of context to activate the reaction as output and handles the process of when action should be taken.

The characteristics of context given by Dey et al. (Dey et al., 2001) is closest in spirit to the operational context characteristics that we seek. Four essential characteristics of context information are identified to get a more extensive assessment of a situation. These characteristics are: identity, location, status and time stamp. Identity refers to the ability to assign a unique identifier to a UAV. The identifier has to be unique in the namespace that is used by the CAS system. Location is more than just position information in three dimensional space. It is expanded to include the three degree orientation of an object, as well as all information that can be used to deduce spatial relationships between UAVs. Status identifies current negotiation situation of the UAV with other cooperative conflicting objects in the shared airspace. It also shows whether the UAV that is involved in resolving cooperative conflict is busy or ready for negotiation. Finally, time stamp helps to characterize a situation, used in conjunction with other pieces of context. It enables to leverage off the richness and value of historical information for the purpose of projecting states of the encounter into the future to check for collision risk.

4.2 Awareness engine in collision avoidance systems

The implementation of the context-awareness function in the collision avoidance system is utilized for monitoring surrounding situations and overtaking aircraft management in critical conditions and return control when flight conditions become normal. The context awareness engine involves the sub-functions: estimation of the current traffic situation, refining the reported encounters, computing the collision parameters in case of a potential collision is detected to occur in the near-future, and invoking avoidance function at a suitable time. The main objective of this function is to detect the protection zones violation raised between conflicting aircraft and measure the conflicting parameters in case of potential collision event is detected. A conflict between aircraft and encounter occurs when the protective zone of an aircraft overlaps with protective zone of an encounter.

Information from aircraft in the vicinity is used to create track files and projections of intention on the three dimensional map. This map provides a situational awareness of all neighboring aircraft, surrounding obstacles and no flying zones. The profiles in three dimensional space establish future intersection points that will results in a collision.

Once a potential collision is predicted, the Time To collision and Maneuver (TTC and TTM) for the avoidance phase is calculated. When TTM reaches zero, an automatic escape maneuver is realized by the awareness function and continues until the aircraft is no longer in danger.

The awareness function is based on the estimation of future UAV position and the application of predefined metrics such as time, distance, cost …etc., on the conflict situation to decide whether or not a potential conflict is exist. Although many studies were conducted focusing on the required detection matrices, collision risk assessment technique, maneuver execution time and so on, however the work in this chapter introduces the new concept, awareness engine, that handles these factors to simplify the problem specially in case of dense environment. The design factors associated with the collision detection and awareness are explored in the following subsections.

4.2.1 Conflicting aircraft's state extractor

The first important factor for encounter's state acquisition is the encounter sensing dimension in which the UAV will get knowledge about encounter's state vector. This dimension demonstrates whether the monitoring of the environment used in a given

approach is in two dimensional horizontal plane (2D-H), two dimensional Vertical Plane (2D-V) or three dimensional state information (3D). The majority of the developed CAS approaches cover either 3D or 2D-H. However GPWS focuses on the 2D-V.
The coverage of a certain dimension doesn't necessary mean complete description of the situation in that dimension is available. For example, TCAS uses range measurements and range rate estimates to determine if a conflict exists in the horizontal plane. A better prediction of the threat condition could be obtained if additional information were available such as bearing.

4.2.2 Conflicting aircraft's state projection

Another collision detection and awareness design factor is the prediction of the encounter's future state vector. That is because it specifies the way of dynamic projecting states of UAV and encounter into the near future and check for collision risk. Four fundamental prediction methods have been identified. These methods are, as illustrated in figure 5: straight projection, worst case, probabilistic and flight plan sharing.

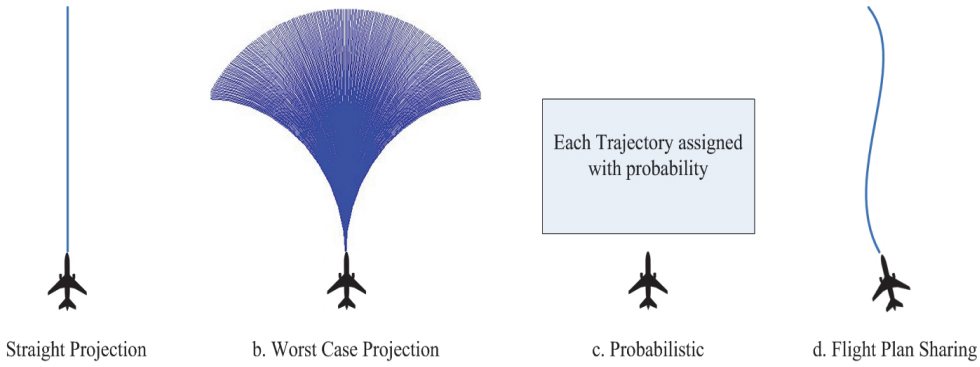


Fig. 5. Projection types of encounter's state vector

In the straight projection method, the states are projected into the future along a single straight trajectory, without direct consideration for uncertainties, as shown in Figure 5.a. This will simplify the problem but it is can be only used in situations in which aircraft trajectories is very predictable or used for short period of time. That is because the CAS approach that uses straight projection doesn't account for the possibility that an encounter can do any maneuvering in predicted time. Albaker and Rahim(Albaker & Rahim, 2009d) developed a new functional architecture for unmanned aircraft collision avoidance system with an avoidance algorithm utilized for deciding the collision criteria upon straight state projection in the near future.
The other extreme is the worst case projection illustrated in Figure 5.b, which assumes an aircraft will perform any range of maneuvers bounded by its physical limitation. If anyone of these trajectories could cause a conflict, then a conflict is predicted. It should be limited to a short period projection time to limit the computation requirement for risk assessment. Tomlim et al. (Tomlim et al., 2000) approached the collision avoidance problem from non-cooperative game theoretical angle. These approaches often solve for solutions that work in worst case scenarios. Although, these methods may provide an acceptable solution, they are far from optimal solution.

In the probabilistic method, the uncertainties are modeled to describe risk variation in the future trajectory of aircraft, as shown in Figure 5.c. This method is based on developing a complete set of possible future trajectories, each weighted by a probability of occurring, making a probability density function. The advantages of this method is that decisions can be made on the fundamental likelihood of conflict; safety and false alarm rate can be assessed and considered directly. However, the disadvantage is that the logic behind this method may be difficult to model the probabilities of future trajectories. Moreover, it requires heavy processing to cope calculations in case of large number of aircraft in a given shared airspace.

Most other methods of escape trajectory maneuvering rely on trajectory estimation filters based on previous intruder path history, position versus time. However, actual intended intruder path data, such as position; heading; and future waypoints, offers a much more reliable basis for path planning than trying to estimate where the intruder might go given its previous history.

The advancement of the technology allows for the forth method of encounters' states projection using path plan sharing, as depicted in Figure 5.d. It is a method of providing path trajectory (flight plan segment) and aircraft specific information (like position, heading and velocity) to all other aircraft in the vicinity. As an example, Albaker and Rahim (Albaker & Rahim, 2009c) and Sislak et al. (Sislak et al., 2008) use flight path sharing method for the assessment of collision risk then provide a solution for conflicting senario. Data from each aircraft will be sent to ground stations for monitoring and all neighboring aircraft as a broadcast. This will leads give all aircraft a 3D picture of neighboring aircraft movements, precise projection of encounters' states and exact collision parameters extraction. As an example ADS-B that is proposed to be fully deployed in aircraft by the year 2020 to support free flight capability (Asep et al., 1996). Other examples support free flight concept can be found in the references (K. Bilimoria, 1996; Holdsworth, 2003; J. Hill, July 2005; Christodoulou & Kodaxakis, 2006). However, the focus needs to be on removing the complexity of data exchanges and the quantity of data required to ensure safe maneuvers. Clearly, the more data needed to be exchanged in collision situation, the more complex and prone to error the system becomes.

4.2.3 Assessment of collision risk and collision parameters extraction

The design of any collision avoidance system should include some form of collision risk assessment. This is a complex issue that receives considerable attention in the literature. An example is given by Carlson and Lee (Carlson & Lee, 1997). Merz (Merz, 1991) describes a method of avoiding collision given the increased likelihood of collision as aircraft numbers and packing densities increase. The limit on packing density where these algorithms no longer work is examined by Bowers and smith et al. (Bowers, 1996; Smith et al., Mar 1998).

Approaches may use an extremely simple criterion like range information to determine when a conflict exists or may use a more complex threshold or set of logic. Some of them uses concept of a simple threat detection zone around each aircraft and determines a maneuver that ensures adequate separation even if one aircraft does not maneuver. This provides safe separation even if the link to one aircraft fails.

A determination of potential collision and request for trajectory maneuvering are done by utilizing relative position information, its rate of change and/or trajectory information between conflicting aircraft. When a potential conflict is reported, the context awareness

engine is then estimates the collision parameters. These parameters include the estimation of Time-To-Collision (TTC), Collision Interval (CI), time to activate the escape trajectory and collision angle (See Figure 6). Albaker and Rahim (Albaker & Rahim, 2009c) developed a method to extract collision parameters based on flight trajectory sharing.

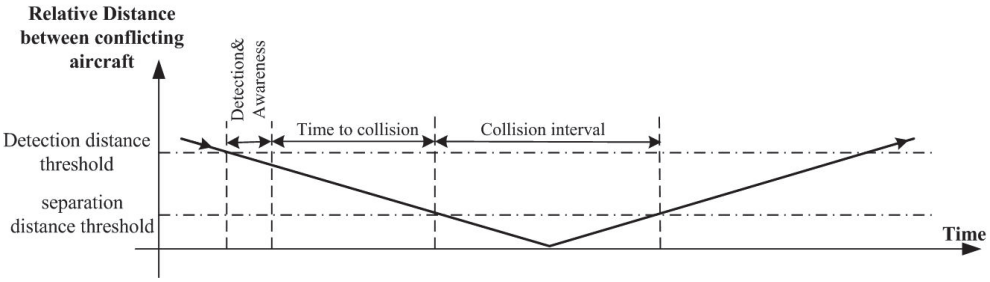


Fig. 6. Separation distance between two aircraft demonstrating some of the collision parameters in a future course collision scenario.

5. Escape maneuver algorithms

Various approaches have been proposed in the collision avoidance literature for choosing escape trajectories that generate solution to a conflict. Six main categories of the escape trajectory approaches are introduced in this paper, which are: predefined, negotiation protocol based, optimized, force-filed, game theory, automotive and hybrid systems. These approaches will be discussed in details in the next subsections.

To provide insight into different CAS algorithms, a literature review of previous research models and current developmental and operational systems is performed. Based on the collision avoidance system design factors as illustrated in Figure 3, the algorithms were catalogued according to their fundamental approaches to each phase of CAS function. The major collision avoidance algorithms for UAVs are categorized into four main methods. These methods are explained together with their advantages and disadvantages in the following subsystems.

5.1 Predefined trajectory escape

This type of collision avoidance is based on a fixed set of predefined rules without performing any additional computation to determine an escape trajectory. The advantage is on minimizing the response time to avoid the conflict. On the other hand the disadvantages will be on less effectiveness and less optimal than the maneuvers which are computed in online. That is because there is no way to alter the commanded maneuver, which is very essential to account for unexpected events. As an example, Ground Proximity Warning System (GPWS) issues a standard climb warning when a conflict with terrain exists (Bateman, 1999).

5.2 Optimized trajectory based algorithms

In this type, the collision avoidance problem is often formulated as an optimization problem. Algorithms using this kind of trajectory escape are generally combining a kinematic model with a set of constraints. An optimal resolution strategy is then computed

based on most desired optimization constraint. For example, the TCAS system does not seek to define an escape trajectory, instead requesting a climb or dive maneuver(RTCA, 1983; Committee147, 1997). It searches through a set of potential climb or descent maneuvers and selects the least-aggressive maneuver that provides adequate protection. The idea implies that somehow the system knows that the path planned towards the goal without taking account of intruders would be unsafe. An aircraft will head for its goal until a collision threat is detected and then find a trajectory that will avoid the collision. Path planning should be more elegant, that is finding a safe trajectory that still reaches the goal.

Tomilin et al. (Tomlim et al., 2000), Zhang and Sadtry (Zhang & Sastry, 2001) and Bayen et al. (Bayen et al., 2003) presented an optimization approach using game theorical technique for controller design that covers moving obstacles. In this technique, a pursuer's trajectories are examined based on all possible plans and the evader seek for collision free paths those are not intersecting with pursuer's trajectories. Although it is interesting, it does not appear practical at present. Archibald et al. (Archibald et al., 2008) described a multiagent solution to aircraft conflict resolution based on satisficing game theory. A key feature of the theory is that satisficing decision makers form their preferences by taking into consideration the preferences of others. The results in behavior is attractive both in terms of safety and performance. Mixed Integer Nonlinear Programming is also used as an optimization problem to solve traffic conflicts. However, this algorithm is hard to be extended to consider many maneuvering commands.

Another well-known safe navigation method originating from mobile ground robot research community is the dynamic window approach, presented by Fox et al. (Fox et al., 1997). This approach takes into account the dynamic model and kinematic constraints of aircraft to determine a safe control action. Nguyen (Nguyen, 2007), in his thesis, proposed collision avoidance system using horizon escape windows for UAVs. His proposal was based on proposing asymmetrical collision risk assessment metrics. Then an optimization is formulated to solve conflict based on possible trajectories for each UAV that can follow. Albaker and Rahim (Albaker & Rahim, 2011a) introduced a new collision avoidance algorithm based on geometrical intersection method for the estimation of collision risk. When a potential conflict along the trajectory exists, the collision avoidance is activated to take the action of filtering the possible trajectories to avoid the conflict and the best option to consider based on optimization problem. Van Dam et al. (Van Dam et. al, 2008) defined the workspace key functions required by the airborne separation assistance tool. A geometrical approach, supporting free flight concept, is proposed for conflict avoidance without the need to communicate among the conflicting aircraft. The authors based on implicit coordination among aircraft in a shared airspace. Speed and heading travel functions is utilized as resolution maneuvers to clear incoming threats.

Other optimized conflict resolution algorithms utilize techniques such as genetic algorithms, expert systems, or fuzzy control to the problem (Zengin, 2007; Tseng, 2008; Holdsworth, 2003). These techniques may be complex and therefore would require a large number of rules to completely cover all possible encounter scenarios. This leads to demanding high computational processing power. Resulting in difficult to certify that the system will always operate as intended.

Pre-mission path planning is often formulated as an optimization problem and many different optimization problems can be applied. Path planning for UAVs is difficult problem because it requires the ability to create paths in environments containing obstacles or no-flying zones. Additionally, UAVs are constrained by minimum turning radius, minimum

speed, and maximum climb rate constraints. Generally, CAS algorithms are used to sparsely search the space for solutions and then the best solution is chosen.

5.3 Negotiation protocol based maneuvers

This type offers a very elegant solution to conflict free navigation for a team of agents, each agent represent an aircraft. Inter-agent communication includes sharing position, velocities, waypoints and heading. Agents make decisions based on a common set of rules decided priori. This method is decentralized, highly scalable and guarantees safety. However, the trade off is that unnecessary long trajectories can be generated long mission completion times. (Albaker & Rahim, 2010a; Albaker & Rahim, 2009b; Wollkin et al., 2004; Wangermann & Stengel, 1999; Sislak et al., 2011; Sislak et al., 2010; Pechoucek & Sislak, Jan 2009) are examples use this kind of collision avoidance approach.

5.4 Force-field based collision avoidance

Many methods have been proposed for safe navigation in static obstacle strewn environment. Most popular obstacle avoidance methods are artificial potential field methods. Researchers have considered the force field to map the volume between aircraft in terms of a potential field. The methods treat each aircraft as a charged particle and the repulsive forces between aircraft are used to generate maneuvering trajectories. This type is considered as a path planning technique that estimates the trajectories by creating trajectory estimation filters based on the previous paths. Trajectories with low flex densities can be then selected as the preferred courses. The method shows some success through the sense that conflict avoidance is continuously available using simple electrostatic equations. However, the algorithms presented have limited relevance due to sharp discontinuities in the commanded maneuvers that may occurs. Furthermore, it requires a high level of flight guidance, leads to increase in complexity beyond issuing simple maneuvering commands.

Artificial potential field methods were first presented by Khatib (Khatib, 1985). Other methods utilize same escape method can be found in references (Miura et al., 1995; Jen-Hui, 1998; McQuade & McInnes, 1997; Veelaert & Bogaerts, 1999). Obstacle and other agents are modeled as repulsive forces and waypoints as attractive forces; the gradient of the summation of these forces yields the control command. These methods provide very simple and elegant solutions to general collision avoidance scenarios. However, the existence of local minima could trap an aircraft for infinite time (Krogh & SME., 1984). Potential field like methods that didn't have local minima were later demonstrated (Rimon & Koditschek, 1992; Kim & Khosla, 1992). The design for multi-agent systems presented in (Chang et al., 2003), in which the repulsion force from neighboring agents is replaced by a gyroscopic force from the nearest neighbor. This force will enable an agent to spin free in symmetrical conflict scenarios.

5.5 Other escape maneuvering algorithms

In addition to the above most famous approaches, there are several other CAS approaches to be considered. Like automotive collision avoidance, that offers some interesting analogies for aircraft but does not appear to have been considered for this purpose in the literature. It attempts to predict the vehicle trajectory using historical information or forward looking sensors (Min Young et al., 1996).

Another method uses hybrid CAS algorithm as presented by Tomlin et al. (Tomlin et al., 1998; Tomlim et al., 2000) and Pappas et al. (Pappas et al., 1996). This type of realization is

concerned with the modeling and control of systems combining continuous and discrete states. In this method, vehicle and its maneuver is modeled as a hybrid system and its reachable sets of states is filtered based on safety specifications to get a safe subset of the reach set. Then Hamilton-Jacobi equations are employed to calculate control commands that can guarantee UAV will remain in its safe set. Although this method is decentralized and guarantees safety, it scales poorly for large UAVs.

6. Trajectory maneuvering realization

A maneuver is the combination of actions by all conflicting aircraft in the vicinity. Initiating a resolution maneuver requires at least one aircraft to change its flight trajectory. Maneuver realization can implement all degree of freedom of aircraft control. Three maneuvers dimensions are identified for maneuver realization. These maneuvers include: horizontal plane, turn left/right; vertical maneuver, climb/dive; and/or speedup slowdown commands. The maneuvers depends on the CAS approach used, limited by the physical constraints of the aircraft as given by its flight dynamics. It may be issued separately (e.g. change of only one dimension) or combined maneuvers may be performed (e.g. speed and vertical and horizontal planes). Furthermore, the combined maneuvers can be performed simultaneously or in sequence.

Issues such as coordinated and uncoordinated maneuvers also need to be addressed. Coordinated maneuver refers to the choice of the direction when there is a choice of two alternative versions of maneuver. As an example in TCAS in which the preferred maneuver might be for aircraft A to climb while aircraft B descends. While the uncoordinated maneuver refers to the worst case scenario, in which the other aircraft does not respond and only the computing aircraft should do all the maneuvering commands.

7. Other collision avoidance factors

One of the other important CAS design factors is that, complex computation performed by an approach versus time requirement to resolve the conflict. The designed approaches should take into consideration finding the solution in real time. This means compromise between two factors must be done. That is the complexity of the calculation needs to be bounded, to provide an approach that is effective and robust but reasonably simple.

Collision avoidance systems are also differ by their system architecture those designed for. Basically, there are four type of CAS architecture, which are: Centralized, layered, predictive and decentralized. Centralized approaches, such as current air traffic management, are considered easy, one system controls all. Therefore, this type will improve the overall global performance. However, it is considered computationally expensive and the whole system fails in case if centralized controller failed. Furthermore, this type fails to prevent collisions among conflicting aircraft when their number increases. On the other hand, modular layered CAS architecture type is scalable but it adds design interfaces that may delay the response to solve the conflicts. Such a system can be found in (Casalino et al., 2009). Predictive control type is interesting as it handle time delay and packet loss. However, it design for uncertainties which much complicates the problem. Therefore it may fail to provide solution in a dense environments. Some of the research addressing this type can be found in (Lapp & Singh, 2004; Boivin et al., 2008). Most of the research done in this field based on decentralized CAS architecture (Borrelli et al., 2004; Lalish & Morgansen, 2008;

Keviczky et al., 2008; Roozbehani et al., 2009; Sislak et al., 2011). This is a critical requirement for autonomous UAVs. That is for implementing a self decision in which each UAV handles its own avoidance. A comparison of centralized and decentralized conflict resolution strategies were presented in (Bilimoria et al., 2000).

Another important design factor is the consideration of the CAS system for detection and accordingly resolving conflicts in multiple encounters scenarios. It describes how an approach handles traffic situations with multiple aircraft. It is divided into two types: Single conflict management approaches in which multiple sequential conflicts are avoided sequentially in pairs, and multiple conflict management approaches in which the entire situation is handled simultaneously. The general problem raises questions such as does this maneuver work on multi aircraft? Is there a maximum packing density where maneuvers no longer work and is it dependent on aircraft type or separation criteria?

Cooperative and non-cooperative collision avoidance algorithm can be also considered as one of the factors affacting CAS design. Technically, cooperative has equiped with ATC transponder or the recent ADS-B technology. In general, the case for which aircraft can communicate together resolves the conflicts more efficiently in term of their flight paths as compared with non-cooperative collision avoidance algorithms.

8. Conclusion

In this chapter, the fundamental concept and the key functions of the unmanned aircraft collision avoidance system are carried out. Special attention is given to the context-aware implementation in the collision avoidance domain. The intent of this chapter is to introduce a new conceptual framework for CAS system to handle conflicts more efficiently. Accordingly, providing an up-to-date review of collision avoidance algorithms based on the main CAS design factors those which also handled in details.

Building collision avoidance capability into flight controller requires detailed knowledge of the aircraft dynamics and deployment. In theory, CAS algorithms for UAVs like to assume that their models and methods are working efficiently. However in reality, due to system weaknesses and sensor error compound over time their systems may fail to prevent collisions. Due to a wide variety of UAVs types, their operating conditions, environment and missions, leads to the need for different degrees and types of collision avoidance algorithms. Therefore, no one sensing method and one CAS algorithm should be expected to cover all types and conditions. However, two or more fused sensors may be required to provide a complete picture of the surounding environment. Thereby resolve conflicts more efficiently. Availability of new techniques and sensors will lead to new and exciting CAS algorithms to be continually developed.

9. References

Albaker, B. & Rahim, N. (2009a). Signal Acquisition and Parameter Estimation of Radio Frequency Pulse Radar Using Novel Method. IETE Journal of Research, Vol.55, No.3, pp. 128-134.

Albaker, B. M. & Rahim, N. A. (2009b). Collision Detection and Resolution for Cooperative UAVs Based onto Peer-to-Peer Communication and Predefined Maneuvering Rules. proceedings of the Defense & Security Technology Conference. Kuala Lumpur, Malaysia.

Albaker, B. M. & Rahim, N. A. (2009c). Intelligent conflict detection and awareness for UAVs. Innovative Technologies in Intelligent Systems and Industrial Applications, 2009. CITISIA 2009, pp. 261-264.

Albaker, B. M. & Rahim, N. A. (2009d). Straight Projection Conflict Detection and Cooperative Avoidance for Autonomous Unmanned Aircraft Systems. proceedings of the 4th IEEE conference on Industrial Electronics and Applications, (ICIEA 2009), Xian P. R. China, pp. 1956-1960.

Albaker, B. M. & Rahim, N. A. (2010a). Unmanned Aircraft Collision Avoidance System Using Cooperative Agent-Based Negotiation Approach. International Journal of Simulation, System, Science and Technology, Vol.11, No.4, pp. 1-8.

Albaker, B. M. & Rahim, N. A. (2010b). Unmanned aircraft collision detection and resolution: Concept and survey. Industrial Electronics and Applications (ICIEA), 2010 the 5th IEEE Conference on, pp. 248-253.

Albaker, B. M. & Rahim, N. A. (2011a). Autonomous Unmanned Aircraft Collision Avoidance System Based on Geometric Intersection. International Journal of Physical Sciences, Vol.6, No.3, pp. 391-401.

Albaker, B. M. & Rahim, N. A. (2011b). Detection and Parameters Interception of a Radar Pulse Signal Based on Interrupt Driven Algorithm. Journal of Scientific Research and Essays, Vol.6, No.6, pp. 1380-1387.

Archibald, J. K.; Hill, J. C.; Jepsen, N. A.; Stirling, W. C. & Frost, R. L. (2008). A Satisficing Approach to Aircraft Conflict Resolution. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol.38, No.4, pp. 510-521.

Ariyur, K. B.; Lommel, P. & Enns, D. F. (2005). Reactive inflight obstacle avoidance via radar feedback. Proceedings of the American Control Conference, pp. 2978-2982.

Asep, R.; Achaibou, A. & Mora-Camino, F. (1996). Automatic collision avoidance based on supervised predictive controllers. Control Engineering Practice, Vol.4, No.8, pp. 1169-1175.

Asmat, J.; Rhodes, B.; Umansky, J.; Villavicencio, C.; Yunas, A.; Donohue, G. & Lacher, A. (2006). UAS Safety: Unmanned Aerial Collision Avoidance System (UCAS). IEEE Systems and Information Engineering Design Symposium, pp. 43-49.

Bateman, C. (1999). The introduction of Enhanced Ground-Proximity Warning Systems(EGPWS) into civil aviation operations around the world. Flight safety: Management, measurement and margins, pp. 259-273.

Bayen, A.; Santhanam, S.; Mitchell, I. & Tomlin, C. (2003). Differential game formulation of alert levels in ETMS data for high altitude traffic. AIAA Guidance, Navigation and Control conference.

Bilimoria, K. D.; Lee, H. Q.; Mao, Z. & Feron, E. (2000). Comparison of Centralized and Decentralized Conflict Resolution Strategies for Multiple-Aircraft Problems. proceedings of the AIAA Guidance, Navigation, and Control Conference.

Boivin, E.; Desbiens, A. & Gagnon, E. (2008). UAV collision avoidance using cooperative predictive control. Control and Automation, 2008 16th Mediterranean Conference on, pp. 682-688.

Boon Kiat, Q.; Ibanez-Guzman, J. & Khiang Wee, L. (2004). Feature detection for stereo-vision-based unmanned navigation. IEEE Conference on Cybernetics and Intelligent Systems, pp. 141-146.

Borrelli, F.; Keviczky, T. & Balas, G. J. (2004). Collision-free UAV formation flight using decentralized optimization and invariant sets. Decision and Control, 2004. CDC. 43rd IEEE Conference on, pp. 1099-1104 Vol.1.

Bowers, K. L. (1996). Determining the feasibility of a flight profile in a Free Flight environment. 15th AIAA/IEEE Digital Avionics Systems Conference, pp. 81-86.

Brudnicki, D.; Lindsay, K. & Mcfarland, A. (1997). Assessment of Field Trials, Algorithmic Performance, and Benefits of the User Request Evaluation Tool (URET) Conflict Probe. in Proc. 16th Digital Avionics Systems Conf., Irvine, CA, pp. 9.3-35 - 9.3-44.

Carlson, R. & Lee, J. (1997). Detecting Near Collisions for Satellites. IEEE Transactions on Aerospace and Electronic Systems, Vol.33, No.3, pp. 921-929.

Casalino, G.; Turetta, A. & Simetti, E. (2009). A three-layered architecture for real time path planning and obstacle avoidance for surveillance USVs operating in harbour fields. OCEANS 2009 - EUROPE, pp. 1-8.

Chakravarthy, A. & Ghose, D. (1998). Obstacle Avoidance in a Dynamic Environment – A Collision Cone Approach. IEEE Transactions on Systems, Man, and Cybernetics, Vol.28 No.5, pp. 562-574.

Chang, D. E.; Shadden, S. C.; Marsden, J. E. & Olfati-Saber, R. (2003). Collision avoidance for multiple agent systems. Proceedings of the 42nd IEEE Conference on Decision and Control, 2003, Vol.1, pp. 539-543.

Christodoulou, M. A. & Kodaxakis, S. G. (2006). Automatic commercial aircraft-collision avoidance in free flight: the three-dimensional problem. IEEE Transactions on Intelligent Transportation Systems, Vol.7, No.2, pp. 242-249.

Clapper, J.; Young, J.; Cartwright, J. & Grimes, J. (2007). Unmanned Systems Roadmap 2007-2032. Dept. of Defense.

Coenen, F. P.; Smeaton, G. P. & Bole, A. G. (1989). Knowledge-Based Collision Avoidance. Journal of Navigation, Vol.42, No.1.

Committee147, R. S. (1997). Minimum Aviation System Performance Standards for Traffic Alert and Collision Avoidance System II (TCAS II) Airborne Equipment.

De Wagter, C. & Mulder, J. (2005). Towards vision-based uav situation awareness. AIAA Guidance, Navigation and Control Conference, San Francisco, California, AIAA-2005-5872, pp. 15–18.

Dey, A. K. & Abowd, G. D. (2000). Towards a better understanding of context and context-awareness. Georgia Institute of Technology: GVU Technical Report GITGVU-99-22, College of Computing.

Dey, A. K.; Abowd, G. D. & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. Human-Computer Interaction, Vol.16, No.2, pp. 97-166.

Federal_Aviation_Administration (1991). Precision Runway Monitor Demonstration Report.

Ford, R. L. (1986). The Protected Volume of Airspace Generated by an Airbourne Collision Avoidance System. The Journal of Navigation, Vol.. 39, No.2, pp. 182-187.

Ford, R. L. & Powell, D. L. (1990). A New Threat Detection Criterion for Airbourne Collision Avoidance Systems. The Journal of Navigation, Vol.43, No.3, pp. 391-403.

Fox, D.; Burgard, W. & Thrun, S. (1997). The Dynamic Window Approach to Collision Avoidance. IEEE Robotics and Automation Magazine, Vol.4, No.1.

Frew, E. W.; Langelaan, J. & Sungmoon, J. (2006). Adaptive receding horizon control for vision-based navigation of small unmanned aircraft. American Control Conference, pp. 6-11.

Gazit, R. Y. (1996). Aircraft surveillance and collision avoidance using gps. PhD thesis, Stanford University.

Holdsworth, R. (2003). Autonomous in-flight path planning to replace pure collision avoidance for free flight aircraft using automatic depedent surveillance broadcast. PhD thesis, Swinburne University of Technology.

Hyeon-Cheol, L. & In-Kyu, K. (2004). Antenna design of collision avoidance radar system for the smart UAV. The 23rd Digital Avionics Systems Conference pp. 12.D.3-12.1-8.

Iijima, Y.; Hagiwara, H. & Kasai, H. (1991). Results of Collision Avoidance Maneuver Experiments Using a Knowledge-Based Autonomous Piloting System. Journal of Navigation, Vol.44, No.2.

Isaacson, D. & Erzberger, H. (1997). Design of a Conflict Detection Algorithm for the Centre/TRACON Automation System. in Proceeding 16th Digital Avionics Systems Conference, Irvine, CA, pp. 9.3-1 - 9.3-9.

J. Hill, F. J., J. Archibald, R. Frost and W. Stirling. (July 2005). A cooperative multi-agent approach to free flight. Proceedings of the 4th international joint conference on Autonomous agents and multi agent systems, Netherlands, pp. 1083-1090.

Jen-Hui, C. (1998). Potential-based modeling of three-dimensional workspace for obstacle avoidance. IEEE Transactions on Robotics and Automation, Vol.14, No.5, pp. 778-785.

K. Bilimoria, B. S., and G. Chatterji. (1996). Effects of Conflict Resolution Maneuvers and Traffic Density of Free Flight. in Proceeding 1996 AIAA Guidance, Navigation, and Control Conference, San Diego, CA.

Karhoff, B. C.; Limb, J. I.; Oravsky, S. W. & Shephard, A. D. (2006). Eyes in the Domestic Sky: An Assessment of Sense and Avoid Technology for the Army's "Warrior" Unmanned Aerial Vehicle. IEEE Systems and Information Engineering Design Symposium, Sieds.

Kelly, W. E. (1999). Conflict Detection and Alerting for Separation Assurance Systems. in Proc.18th Digital Avionics Systems Conf., St. Louis, MO.

Kemkemian, S.; Nouvel-Fiani, M.; Cornic, P.; Le Bihan, P. & Garrec, P. (2009). Radar systems for Sense and Avoid on UAV. International Radar Conference - Surveillance for a Safer World, 2009, pp. 1-6.

Keviczky, T.; Borrelli, F.; Fregene, K.; Godbole, D. & Balas, G. J. (2008). Decentralized Receding Horizon Control and Coordination of Autonomous Vehicle Formations. IEEE Transactions on Control Systems Technology, Vol.16, No.1, pp. 19-33.

Khatib, O. (1985). Real-time obstacle avoidance for manipulators and mobile robots. IEEE International Conference on Robotics and Automation, pp. 500-505.

Kim, J. O. & Khosla, P. K. (1992). Real-time obstacle avoidance using harmonic potential functions. IEEE Transactions on Robotics and Automation, Vol.8, No.3, pp. 338-349.

Koencke, E. J.; Geyer, M.; Lay, R. & Vilcans, J. (1997). Beacon multilateration to facilitate ADS-B. Proceedings of the 53rd Institute of Navigation Annual Meeting: Alexandria, VA: Institute of Navigation, pp. 327-336.

Krogh, B. H. & Sme., R. I. O. (1984). A generalized potential field approach to obstacle avoidance control, RI/SME.

Krozel, J.; Peters, M. & Hunter, G. (1997). Conflict Detection and Resolution for Future Air Transportation Management.

Kuchar, J. K. & Yang, L. C. (2000). A Review of Conflict Detection and Resolution Modeling Methods. IEEE Transactions on Intelligent Transportation Systems, Vol.1, No.4.

Kumar, B. A. & Ghose, D. (2001). Radar-assisted collision avoidance/guidance strategy for planar flight. IEEE Transactions on Aerospace and Electronic Systems, Vol.37, No.1, pp. 77-90.

Kwag, Y. K. & Chung, C. H. (2007). UAV based collision avoidance radar sensor. Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, pp. 639-642.

Lacher, A.; Maroney, D. & Zeitlin, A. (2007). Unmanned Aircraft Collision Avoidance-Technology Assessment and Evaluation Methods. in The 7th Air Traffic Managemtent Research & Development Seminar. Barcelona, Spain.

Lalish, E. & Morgansen, K. A. (2008). Decentralized reactive collision avoidance for multivehicle systems. Decision and Control, 2008. CDC 2008. 47th IEEE Conference on, pp. 1218-1224.

Lapp, T. & Singh, L. (2004). Model predictive control based trajectory optimization for nap-of-the-earth (NOE) flight including obstacle avoidance. Proceedings of the 2004 American Control Conference, pp. 891-896.

Matthies, L.; Litwin, T.; Owens, K.; Rankin, A.; Murphy, K.; Coombs, D.; Gilsinn, J.; Tsai, H.; Legowik, S.; Nashman, M. & Yoshimi, B. (1998). Performance evaluation of UGV obstacle detection with CCD/FLIR stereo vision and LADAR. Intelligent Control (ISIC), 1998. Held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA), Intelligent Systems and Semiotics (ISAS), Proceedings, pp. 658-670.

Mcquade, F. & Mcinnes, C. (1997). Autonomous control for on-orbit assembly using potential function methods. Aeronautical Journal, Vol.101, No.1006, pp. 255-262.

Mehra, R.; Byrne, J. & Boškovic, J. (2005). Intelligent Autonomy and Vision Based Obstacle Avoidance for Unmanned Air Vehicles. in Proceedings of the 13th Yale Workshop on Adaptive and Learning Systems, New Haven. Citeseer.

Merz, A. W. (1991). Maximum-Miss Aircraft Collision Avoidance. Dynamics and Control, Vol.1, No.1, pp. 25-34.

Min Young, C.; Lichtenberg, A. J. & Lieberman, M. A. (1996). Minimum stopping distance for linear control of an automatic car-following system. IEEE Transactions on Vehicular Technology, Vol.45, No.2, pp. 383-390.

Miura, A.; Morikawa, H. & Mizumachi, M. (1995). Aircraft Collision Avoidance with Potential Gradient Ground Based Avoidance for Horizontal Meneuvers. Electronics and Communications in Japan, Part 3, Vol.78, No.10, pp. 104-113.

Moore, R. J. D.; Thurrowgood, S.; Bland, D.; Soccol, D. & Srinivasan, M. V. (2009). A stereo vision system for UAV guidance. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3386-3391.

Morse, D. R.; Armstrong, S. & Dey, A. K. (2000). The what, who, where, when, why and how of context-awareness. Citeseer, pp. 371-371.

Muratet, L.; Doncieux, S.; Briere, Y. & Meyer, J. A. (2005). A contribution to vision-based autonomous helicopter flight in urban environments. Robotics and Autonomous Systems, Vol.50, No.4, pp. 195-209.

Nguyen, D. L. (2007). A Collision Avoidance System using Horizon Escape Windows for Unmanned Aerial Vehicles. University of California.

Oh, P. Y. (2004). Flying insect inspired vision for micro-air-vehicle navigation. Autonomous Unmanned Vehicles System International Symposium, Anaheim, USA: Citeseer.

Ortiz, A. E. & Neogi, N. (2006). Color Optic Flow: A Computer Vision Approach for Object Detection on UAVs. 25th Digital Avionics Systems Conference, 2006 IEEE/AIAA, pp. 1-12.

Pappas, G. J.; Tomlin, C. & Sastry, S. S. (1996). Conflict resolution for multi-agent hybrid systems. Proceedings of the 35th IEEE on Decision and Control, pp. 1184-1189.

Pechoucek, M. & Sislak, D. (Jan 2009). Agent-Based Approach to Free-Flight Planning, Control, and Simulation. IEEE Intelligent Systems Journal, No.1, pp. 14–17.

Prazenica, R.; Kurdila, A.; Sharpley, R. & Evers, J. (2006). Vision-based geometry estimation and receding horizon path planning for UAVs operating in urban environments. American Control Conference, 2006, pp. 6-11.

Rimon, E. & Koditschek, D. E. (1992). Exact robot navigation using artificial potential functions. IEEE Transactions on Robotics and Automation, Vol.8, No.5, pp. 501-518.

Roozbehani, H.; Rudaz, S. & Gillet, D. (2009). On decentralized navigation schemes for coordination of multi-agent dynamical systems. IEEE International Conference on Systems, Man and Cybernetics, pp. 4807-4812.

Rtca, R. T. C. O. A. (1976). Minimum Performance Standards -Airborne Ground Proximity Warning Equipment. Washington.

Rtca, R. T. C. O. A. (1983). Minimum Performance Specifications for TCAS Airborne Equipment. Washington.

Rtca, S. C. (1997). Minimum Aviation System Performance Standards for Automatic Dependent Surveillance Broadcast (ADS-B). In: RTCA (ed.).

Ryan, P. & Brodegard, W. (1997). New Collision Avoidance Device is Based on Simple and Passive Design to Keep the Cost Low. ICAO Journal, Vol.52, No.4, pp.

Schilit, B.; Adams, N. & Want, R. (1994). Context-aware computing applications. Workshop on Mobile Computing Systems and Applications Proceedings, pp. 85-90.

Schilit, B. N. & Theimer, M. M. (1994). Disseminating active map information to mobile hosts. IEEE Network, Vol.8, No.5, pp. 22-32.

Sislak, D.; Pechoucek, M.; Volf, P.; Pavlicek, D.; Samek, J.; Marik, V. & Losiewicz, P. 2008. (AGENTFLY: Towards Multi-Agent Technology in Free Flight Air Traffic Control). Defense Industry Applications of Autonomous Agents and Multi-Agent Systems. . Birkhauser Verlag.

Sislak, D.; Volf, P. & Pechoucek, M. (2010). Large-scale Agent-based Simulation of Air-traffic. In Proceedings of the Twentieth European Meeting on Cybernetics and Systems Research.

Sislak, D.; Volf, P. & Pechoucek, M. (2011). Agent-Based Cooperative Decentralized Airplane-Collision Avoidance. IEEE Transactions on Intelligent Transportation Systems, Vol.12, No.1, pp. 36-46.

Smith, K.; Scallen, S. F.; Knecht, W. & Hancock, P. A. (Mar 1998). An Index of Dynamic Density. Human Factors, Vol.40, No.1, pp. 69-78.

Subong, S.; Lee, B.; Jihoon, K. & Changdon, K. (2008). Vision-based real-time target localization for single-antenna GPS-guided UAV. IEEE Transactions on Aerospace and Electronic Systems, Vol.44, No.4, pp. 1391-1401.

Tatkeu, C.; Deloof, P.; Elhillali, Y.; Rivenq, A. & Rouvaen, J. M. (2006). A cooperative radar system for collision avoidance and communications between vehicles. Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE, pp. 1012-1016.

Taylor, D. H. (1990). Uncertainty in Collision Avoidance Maneuvering. Journal of Navigation, Vol. 43, No.2.

Tomlim, C. J.; Lygeros, J. & Sastry, S. S. (2000). A Game Theoretic approach to Controller Design for Hybrid Systems. Proceedings of the IEEE, Vol.88, No.7, pp. 949-970.

Tomlin, C.; Pappas, G. & Sastry, S. (1998). Conflict resolution for air traffic management: a study in multi agent hybrid systems. IEEE Transactions on automatic control, Vol.43, No.4, pp. 509-521.

Tseng, W. (2008). Using Fuzzy Logic Control in Simulation of Autonomous Navigation and Formation Flight of Unmanned Aerial Vehicles. MSc thesis.

Utt, J.; Mccalmont, J. & Deschenes, M. (2005). development of a sense and avoid system. In the proceedings of the American Institute of Aeronautics and Astronautics (AIAA) Infotech Aerospace conference, Virginia.

Van Dam S. B. J.; Mulder, M.& Van Paassen, M. M. (2008). Ecological Interface Design of a Tactical Airborne Separation Assistance Tool. IEEE Transactions on Systems, Man & Cybernetics, Part A, Vol. 38, No. 6, pp. 1221–1233, 2008.

Veelaert, P. & Bogaerts, W. (1999). Ultrasonic potential field sensor for obstacle avoidance. IEEE Transactions on Robotics and Automation, Vol.15, No.4, pp. 774-779.

Waller, M. & Scanlon, C. (1996). Proceedings of the NASA Workshop on Flight Deck Centred Parallel Runway Approaches in Instrument Meteorological Conditions. NASA Conference Publication 10191, Hampton, VA.

Wangermann, J. & Stengel, R. (1999). Optimization and coordination of multi agent systems using principled negotiation. Journal of guidance control and dynamics, Vol.22, No.1, pp. 43-50.

Warren, A. (October 1997). Medium Term Conflict Detection for Free Routing: Operational Concepts and Requirements Analysis. 16th Digital Avionics Systems Conference, Irvine, CA, pp. 9.3-27 - 9.3-34.

Wollkin, S.; Valasek, J. & Ioerger, T. (2004). Automated conflict resolution for air traffic management using cooperative multiagent negotiation. American institute of aeronautics and astronautics, pp. 2004-4992.

Zeghal, K. (August 1998). A Review of Different Approaches Based on Force Fields for Airborne Conflict Resolution. AIAA-98-4240, AIAA Guidance, Navigation, and Control Conference, Boston, MA, pp. 818-827.

Zengin, U. (2007). Autonomous and cooperative multi-UAV guidance in adversarial environment. PhD thesis, University of Texas at Arlington.

Zhang, J. & Sastry, S. (2001). Aircraft conflict resolution: Lie-Poisson reduction for game on SE(2). Proceedings of the 40th IEEE Conference on Decision and Control, pp. 1663-1668.

Zhihai, H.; Iyer, R. V. & Chandler, P. R. (2006). Vision-based UAV flight control and obstacle avoidance. American Control Conference.

Zufferey, J. C.; Beyeler, A. & Floreano, D. (2010). Autonomous flight at low altitude with vision-based collision avoidance and GPS-based path following. Robotics and Automation (ICRA), 2010 IEEE International Conference on, pp. 3329-3334.

22

Collision Probabilities, Aircraft Separation and Airways Safety

Luís Campos[1,3] and Joaquim Marques[2,3]
[1]*Área Científica de Mecânica Aplicada e Aeroespacial (ACMAA), Instituto Superior Técnico (IST), Lisboa,*
[2]*Faculdade de Ciências Aeronáuticas (FCA), Universidade Lusófona de Humanidades e Tecnologias (ULHT), Lisboa,*
[3]*Centro de Ciências e Tecnologias Aeronáuticas e Espaciais (CCTAE), Lisboa,*
Portugal

1. Introduction

The steady growth of air traffic at a rate of 3-7% per year over several decades has placed increasing demands on capacity that must be met with undiminished safety (Vismari & Júnior, 2011). The trend is in fact to improve safety, while meeting more stringent requirements for environment impact, efficiency and cost. The traditional method of safety assurance in Air Traffic Management (ATM) is the setting of separation rules (Houck & Powell, 2001). The separation distances are determined by: (i) wake vortex effects on approach to land and take-off queues at runways at airports (FAA, 2011; International Civil Aviation Organization [ICAO], 2007; Rossow, 1999); (ii) collision probabilities for the in-flight phases of aircraft operations (Campos & Marques, 2002; Reich, 1966; Yuling & Songchen, 2010). Only the latter aspect is considered in the present chapter.

A key aspect of ATM in the future (Eurocontrol, 1998) is to determine (i) the technical requirements to (ii) ensure safety with (iii) increased capacity. The concepts of 'capacity', 'safety' and 'technology' can be given a precise meaning (Eurocontrol, 2000) in the case of airways with aircraft flying on parallel paths with fixed lateral/vertical (Figure 1), or longitudinal (Figure 2) separation: (i) the 'capacity' increases for smaller separation L; (ii) navigation and flight 'technology' should provide a reduced r.m.s. position error σ; (iii) the combination of L and σ should be such that the probability of collision (ICAO, 2006) does not exceed ICAO Target Level of Safety (TLS) of 5×10^{-9} per hour (ICAO, 2005). Thus the key issue is to determine the relation between aircraft separation L and position accuracy σ, which ensures that the ICAO TLS is met. Then the technically achievable position accuracy σ specifies L, viz. the safe separation distance (SSD). Conversely, if an increase in capacity is sought, the separation L must be reduced; then the ICAO TLS leads to a position accuracy σ which must be met by the 'technology'. The position accuracy σ includes all causes, e.g. navigation system (Anderson, 1966) error, atmospheric disturbances (Campos, 1984, 1986; Etkin, 1981), inaccuracy of pilot inputs (Campos, 1997; Etkin & Reid, 1996; Etkin & Etkin, 1990), etc.

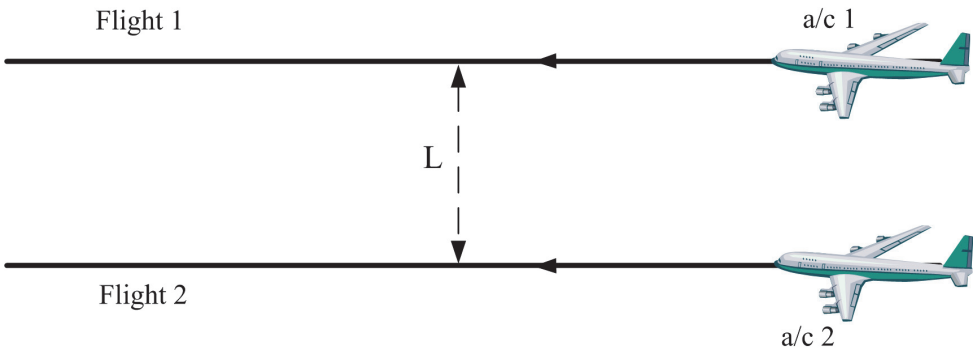


Fig. 1. Aircraft flying always at minimum lateral/vertical separation distance *L*.

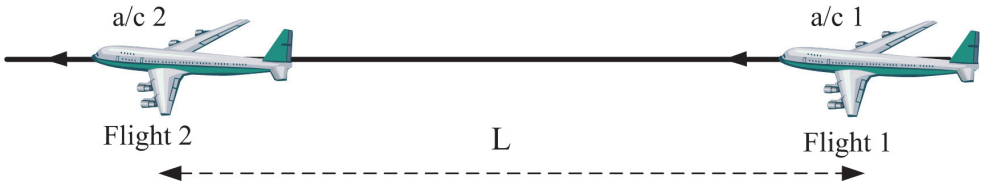


Fig. 2. Aircraft flying always at minimum longitudinal separation distance *L*.

The two main ATM flight scenarios are: (i) parallel paths (Figure 1) with fixed separations in flight corridors typical of transoceanic flight (Bousson, 2008); (ii) crossing (Figure 3) and climbing/descending (Figure 4) flight paths typical of terminal flight operations (Shortle at al., 2010; Zhang & Shortle, 2010). Since aircraft collisions are rare, two-aircraft events are more likely and this the case considered in the present chapter.

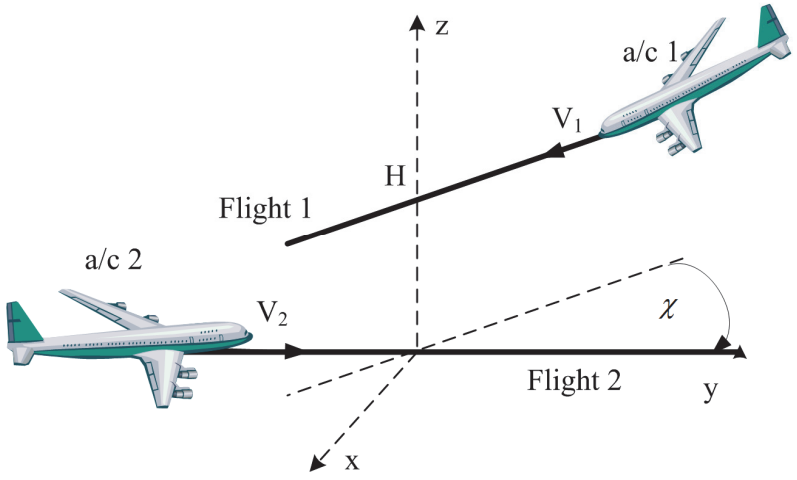


Fig. 3. Geometry of crossing aircraft.

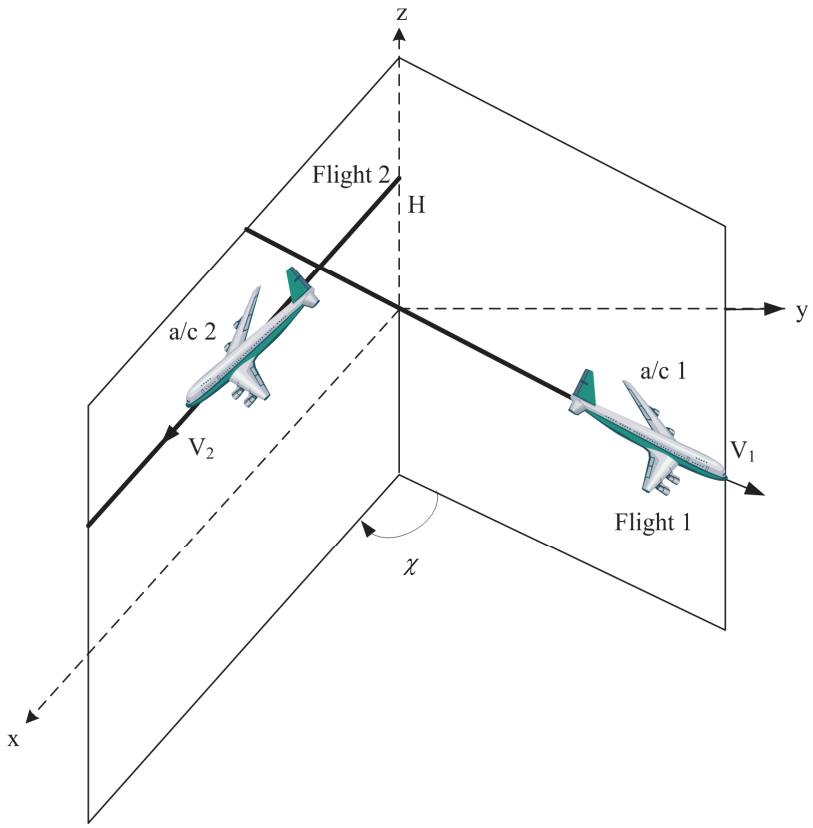


Fig. 4. Geometry of climbing/descending aircraft.

The methods to calculate collision probabilities (Reich, 1966) have been applied to Reduced Vertical Separation Minima (RSVM), to lateral separation (Campos, 2001; Campos & Marques, 2002), to crossing aircraft (Campos & Marques, 2007, 2011), to free flight (Barnett, 2000) and to flight in terminal areas (Shortle et al., 2004). The fundamental input to the models of collision probabilities, is the probability distribution (Johnson & Balakrisshann, 1995; Mises, 1960) of flight path deviations; since it is known that the Gaussian distribution underestimates collision probabilities, and the Laplace distribution though better (Reich, 1966) is not too accurate, the generalized error distribution (Campos & Marques, 2002; Eurocontrol, 1988), and extensions or combinations have been proposed (Campos & Marques, 2004a). It can be shown (Campos & Marques, 2002) that for aircraft on parallel flight corridors (Figure 1) an upper bound to the probability of collision is the probability of coincidence (PC). Its integration along the line joining the two aircraft leads to the cumulative probability of coincidence (CPC); the latter has the dimensions of inverse length, and multiplied by the airspeed, gains the dimensions of inverse time, i.e., can be compared to the ICAO TLS. Alternatively the ICAO TLS can be converted to collision per unit distance, which is directly comparable to the CPC. Since most commercial aircraft fly no faster than $V_0 = 625\,k\,t$, the ICAO TLS of $P_0 \le 5 \times 10^{-9}$ /h, is met by

$Q_0 = P_0 / V_0 \leq 8 \times 10^{-12}$ /nm. The latter can thus be used as an Alternate Target Level of Safety (ATLS).

In the present chapter the CPC is calculated (Section 2) for comparison with the ICAO ATLS of 8×10^{-12} probability of collision per nautical mile; three probability distributions are compared (Section 2.1) and discussed in detail: the Gaussian (Section 2.2); the Laplace (Section 2.3); a generalized error distribution (Section 2.4), which is less simple but more accurate, viz. it has been shown to fit aircraft flight path deviations measured from radar tracks (Campos & Marques, 2002, 2004a; Eurocontrol, 1988). The comparison of the CPC with the ATLS, is made (Section 3) for four typical cruise flight conditions: (i/ii) lateral separation $L_a = 50nm$ in uncontrolled (e.g. oceanic) airspace (Section 3.1) and $L_b = 5nm$ in controlled airspace (Section 3.2); (iii/iv) standard altitude separation $L_c = 2000\ ft$ used worldwire (Section 3.3) and RVSM L_d = 1000 ft introduced (figure 5) by Eurocontrol (1988) to increase capacity at higher flight levels (FL290 to FL410). Longitudinal separation along the same flight path could be considered to the limit of wake vortex effects (Campos & Marques, 2004b; Spalart, 1998). In each of the four cases: (i) the CPC is calculated for several position accuracies σ, to determine the minimum which meets the safety (ATLS) standard; (ii) the Gauss, Laplace and generalized distributions are compared for the collision probabilities of two aircraft with similar position errors σ; (iii) the case of aircraft with dissimilar position errors σ_1 and σ_2 is considered from the beginning, and analysed in detail for the most accurate probability distribution. The discussion (Section 4) summarizes the conclusions concerning airways capacity versus position accuracy, for an undiminished safety.

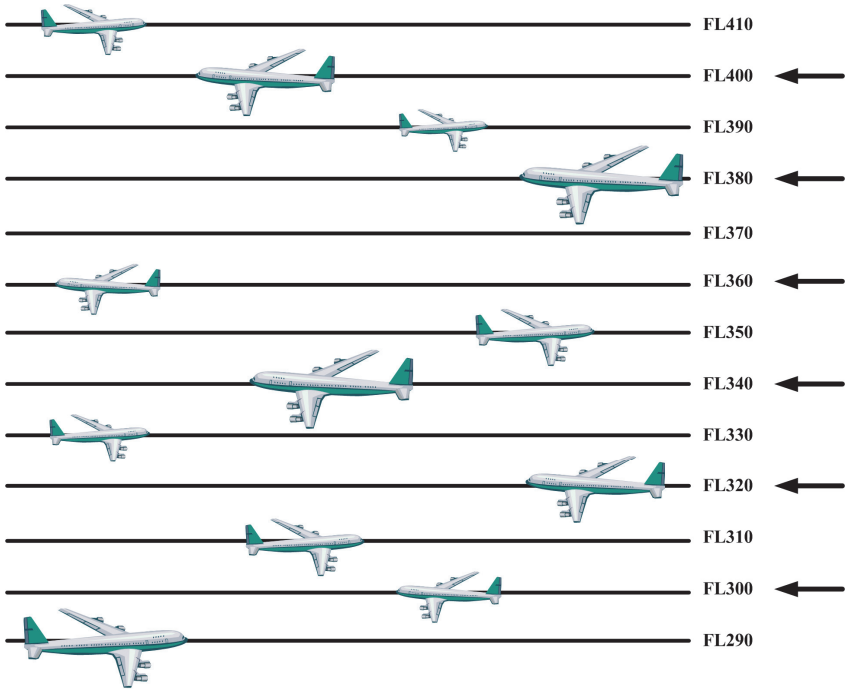


Fig. 5. RVSM between flight levels (FL) 290 and 410 inclusive.

2. Comparison of probability distributions for aircraft flight path

An upper bound for the probability of collision of aircraft on parallel flight tracks (Section 2.1) is calculated using Laplace (Section 2.2), Gaussian (Section 2.3) and generalized (Section 2.4) probability distributions, for aircraft with generally dissimilar r.m.s. position errors.

2.1 Comparison of three probability distributions for flight path deviations

Consider two aircraft flying at: (i) either constant lateral or altitude separation L in parallel flight paths (Figure 1), (ii) or at constant longitudinal separation L on the same flight path (Figure 2). In the case of vertical separation there may be an asymmetry in the probability distributions, which has been treated elsewhere (Campos & Marques, 2007); in the case of longitudinal separation wake effects need to be considered as well (Campos & Marques, 2004b; Spalart, 1998). Apart from these effects, a class of probability distributions (Johnson & Balakshishnan, 1995; Mises, 1960) relevant to large aircraft flight deviations (Campos & Marques, 2002; Eurocontrol, 1998), which are rare events (Reiss & Thomas, 2001; Nassar et al., 2011), is the generalized error distribution (Campos & Marques, 2004a), viz.:

$$F_k(x;\sigma) = A \exp\left(-a|x|^k\right), \tag{1}$$

where k is the weight. The constant a is determined by the condition of unit total probability:

$$A = k\, a^{1/k} \,/\left[2\Gamma\left(1/k\right)\right], \tag{2a}$$

where $\Gamma(\alpha)$ is the Gamma function of argument α. The constant a can be related by:

$$a^{2/k} = \sigma^{-2}\left[\Gamma\left(3/k\right)/\Gamma\left(1/k\right)\right], \tag{2b}$$

to the r.m.s. position error σ or variance σ^2. The case of weight unity in (2a,b), viz.:

$$k = 1: \qquad a = \sqrt{2}/\sigma, \qquad A = 1/\left(\sigma\sqrt{2}\right), \tag{3a,b}$$

corresponds by (1) to the Laplace probability distribution:

$$F_1(x;\sigma) = \left[1/\left(\sigma\sqrt{2}\right)\right] \exp\left(-\sqrt{2}\,|x|/\sigma\right); \tag{4}$$

the case of weight two in (2a,b), viz.:

$$k = 2: \qquad a = 1/\left(2\sigma^2\right), \qquad A = 2/\left(\sigma\sqrt{2\pi}\right), \tag{5a,b}$$

leads by (1) to the Gaussian probability distribution:

$$F_2(x;\sigma) = \left[1/\left(\sigma\sqrt{2\pi}\right)\right] \exp\left[-x^2/\left(2\sigma^2\right)\right]; \tag{6}$$

the best approximation to large aircraft flight path deviations (Campos & Marques, 2002, 2007; Campos, 2001) corresponds approximately to weight one-half, so that (2a,b):

$$k = 1/2: \qquad a^4 = 120/\sigma^2, \qquad A = \sqrt{15/2}\,/\,\sigma, \tag{7a,b}$$

substituted in (1) leads to:

$$F_{1/2}(x;\sigma) = \left(\sqrt{15/2}\,/\,\sigma\right) \exp\left(-\sqrt[4]{120}\,|x/\sigma|^{1/2}\right), \tag{8}$$

which may be designated for brevity the 'generalized' distribution. For any probability distribution, it can be shown (Campos & Marques, 2002) that an upper bound for the probability of collision is the probability of coincidence, which (Figure 6): implies (i) a deviation for the first aircraft, with r.m.s. position error σ_1; (ii) a deviation *L-x* for the second aircraft error σ_2.

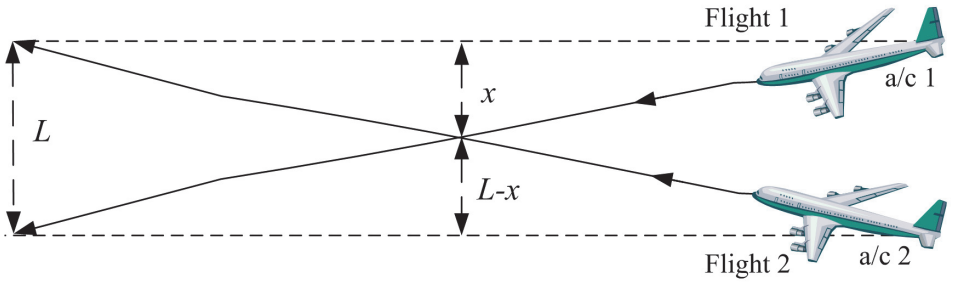


Fig. 6. Aircraft flying on parallel paths: a coincidence will occur if position errors are x (aircraft 1) and L-x (aircraft 2).

For statistically independent aircraft deviations, the probability of coincidence at position *x* the product:

$$P_k(x;L,\sigma_1,\sigma_2) = F_k(x;\sigma_1)\; F_k(L-x;\sigma_2). \tag{9}$$

Its integral over all positions along the line joining the two aircraft is the CPC, viz.:

$$Q_k(L;\sigma_1,\sigma_2) = \int_{-\infty}^{+\infty} P_k(x;L,\sigma_1,\sigma_2)\,dx = \int_{-\infty}^{+\infty} F_k(x;\sigma_1)\; F_k(L-x;\,\sigma_2)\,dx, \tag{10}$$

and, in particular, for aircraft with equal r.m.s. position errors:

$$\sigma \equiv \sigma_1 = \sigma_2: \qquad Q_k(L;\sigma) \equiv Q_k(L;\sigma,\sigma) = \int_{-\infty}^{+\infty} F_k(x;\sigma)\; F_k(L-x;\sigma)\,dx. \tag{11}$$

The CPC has the dimensions of inverse length. The ICAO TLS of 5×10^{-9} /h (12a) can be converted for a maximum airspeed $V_0 = 625\,kt$ in (12b) to a ATLS given:

$$\bar{Q}_0 = 5\times10^{-9} \quad h^{-1}, \qquad V_0 \le 625\,kt, \qquad Q \le Q_0 = \bar{Q}_0 / V_0 \le 8\times10^{-12}\ nm^{-1}, \tag{12a,b,c}$$

which is an upper bound for the CPC. The safety criterion (12c) is applied next to the Laplace (Section 2.2), Gaussian (Section 2.3) and generalized (Section 2.4) probability density functions.

2.2 Laplace distributions for the dissimilar aircraft

The ATLS (12c) is the upper bound for the CPC (10) calculated for aircraft whose position errors follow the Laplace probability distribution (4), with dissimilar r.m.s. position errors for the two aircraft:

$$Q_0 \geq Q_1(L;\sigma_1,\sigma_2) = \left[1/(2\sigma_1\sigma_2)\right] \int_{-\infty}^{+\infty} \exp\left[-\sqrt{2}\left(|x|/\sigma_1 + |L-x|/\sigma_2\right)\right] dx. \tag{13}$$

The appearance of modulus in the argument of the exponential in (13), requires that the range of integration $-\infty,+\infty$ be split in three parts. The first part corresponds to coincidence at $0 \leq x \leq L$ between the flight paths of the two aircraft:

$$\begin{aligned} 2\sigma_1\sigma_2\, Q_{11} &= \int_0^L \exp\left[-\sqrt{2}\left(x/\sigma_1 + (L-x)/\sigma_2\right)\right] dx = \\ &= \exp\left(-\sqrt{2}L/\sigma_2\right) \int_0^L \exp\left[-\sqrt{2}\,x\,(1/\sigma_1 - 1/\sigma_2)\right] dx, \end{aligned} \tag{14}$$

and involves an elementary integration:

$$2\sigma_1\sigma_2\, Q_{11} = \exp\left(-\sqrt{2}\,L/\sigma_2\right)\left\{1-\exp\left[-\sqrt{2}\,L(1/\sigma_1 - 1/\sigma_2)\right]\right\}\left[\sqrt{2}\,(1/\sigma_1 - 1/\sigma_2)\right]^{-1}, \tag{15}$$

and simplifies to:

$$Q_{11} = \left[2\sqrt{2}\,(\sigma_2 - \sigma_1)\right]^{-1}\left[\exp\left(-\sqrt{2}\,L/\sigma_2\right) - \exp\left(-\sqrt{2}\,L/\sigma_1\right)\right], \tag{16}$$

and should be the main contribution (i) to (13). To evaluate (13) exactly, the remaining contributions, besides (i), are also considered: (ii) the coincidence at a point $x \geq L$ outside the path of second aircraft:

$$2\sigma_1\sigma_2 Q_{12} = \int_L^\infty \exp\left\{-\sqrt{2}\left[x/\sigma_1 + (x-L)/\sigma_2\right]\right\} dx, \tag{17}$$

leads to an elementary integral:

$$\begin{aligned} 2\sigma_1\sigma_2\; Q_{12} &= \exp\left(\sqrt{2}\,L/\sigma_2\right) \int_L^\infty \exp\left\{-\sqrt{2}\,x\,(1/\sigma_1 + 1/\sigma_2)\right\} dx \\ &= \exp\left(\sqrt{2}\,L/\sigma_2\right)\left[\sqrt{2}\,(1/\sigma_1 + 1/\sigma_2)\right]^{-1} \exp\left[-\sqrt{2}L(1/\sigma_1 + 1/\sigma_2)\;\right], \end{aligned} \tag{18}$$

which simplifies to:

$$Q_{12} = \left[2\sqrt{2}\,(\sigma_1 + \sigma_2)\right]^{-1} \exp\left(-\sqrt{2}\,L/\sigma_1\right); \tag{19}$$

(iii) the coincidence $-\infty < x < 0$ outside the flight path of the first aircraft:

$$\begin{aligned} 2\sigma_1\sigma_2\, Q_{13} &= \int_{-\infty}^0 \exp\left\{\sqrt{2}\left[x/\sigma_1 - (L-x)/\sigma_2\right]\right\} dx \\ &= \exp\left(-\sqrt{2}\,L/\sigma_2\right) \int_0^\infty \exp\left\{-\sqrt{2}\,x\,(1/\sigma_1 + 1/\sigma_2)\right\} dx, \end{aligned} \tag{20}$$

is again an elementary integral:

$$Q_{13} = \left[2\sqrt{2}\left(\sigma_2 + \sigma_1\right)\right]^{-1} \exp\left(-\sqrt{2}\,L/\sigma_2\right). \tag{21}$$

The sum of (21), (19) and (16) specifies the CPC where:

$$\begin{aligned} Q_1\left(L;\sigma_1,\sigma_2\right) &= \left[2\sqrt{2}\left(\sigma_2 - \sigma_1\right)\right]^{-1}\left[\exp\left(-\sqrt{2}L/\sigma_2\right) - \exp\left(-\sqrt{2}L/\sigma_1\right)\right] \\ &+ \left[2\sqrt{2}\left(\sigma_2 + \sigma_1\right)\right]^{-1}\left[\exp\left(-\sqrt{2}\,L/\sigma_1\right) + \exp\left(-\sqrt{2}\,L/\sigma_2\right)\right], \end{aligned} \tag{22a}$$

for the Laplace distribution:

$$Q_1\left(L;\sigma_1,\sigma_2\right) = Q_{11} + Q_{12} + Q_{13} \le Q_0 = 8\times 10^{-12}\ n\,m^{-1}, \tag{22b}$$

and hence (12c) the safety criterion. Of the preceding expressions, only (16) breaks down for $\sigma_2 - \sigma_1 = 0$, i.e., aircraft with the same r.m.s. position error $\sigma_1 = \sigma_2 \equiv \sigma$. In this case the probability of coincidence is given: (i) between the flight paths of the two aircraft, instead of (14-16) by:

$$\sigma_1 = \sigma_2 \equiv \sigma: \qquad \bar{Q}_{11} = \left(2\sigma^2\right)^{-1}\int_0^L \exp\left(-\sqrt{2}\,L/\sigma\right)dx = \left[L/\left(2\sigma^2\right)\right]\exp\left(-\sqrt{2}\,L/\sigma\right); \tag{23}$$

(ii) outside the flight path of the second aircraft (17-19) is replaced by:

$$\begin{aligned} \sigma_1 = \sigma_2 \equiv \sigma: \qquad \bar{Q}_{12} &= \left(2\sigma^2\right)^{-1}\exp\left(\sqrt{2}\,L/\sigma\right)\int_L^\infty \exp\left(-2\sqrt{2}\,x/\sigma\right)dx \\ &= \left(4\sqrt{2}\,\sigma\right)^{-1}\exp\left(-\sqrt{2}\,L/\sigma\right); \end{aligned} \tag{24}$$

(iii) outside the flight path of the second aircraft (20-22) is replaced by:

$$\begin{aligned} \sigma_1 = \sigma_2 \equiv \sigma: \qquad \bar{Q}_{13} &= \left(2\sigma^2\right)^{-1}\exp\left(-\sqrt{2}L/\sigma\right)\int_{-\infty}^0 \exp\left(2\sqrt{2}\,x/\sigma\right)dx \\ &= \left(4\sqrt{2}\sigma\right)^{-1}\exp\left(-\sqrt{2}\,L/\sigma\right). \end{aligned} \tag{25}$$

The sum of (23), (24) and (25) specifies:

$$Q_1\left(L;\sigma\right) = \exp\left(-\sqrt{2}\,L/\sigma\right)\left(2\sigma\right)^{-1}\left(L/\sigma + 1/\sqrt{2}\right), \tag{26a}$$

as the safety criterion:

$$\sigma_1 = \sigma_2 \equiv \sigma: \qquad Q_1\left(L;\sigma\right) = \bar{Q}_{11} + \bar{Q}_{12} + \bar{Q}_{13} \le Q_0 = 8\times 10^{-12}\ nm^{-1}, \tag{26b}$$

for Laplace probabilities with equal r.m.s. position errors for both aircraft.

2.3 Gaussian distribution with distinct variances

The ATLS (12c) is the upper bound for the CPC (10) calculated next for aircraft whose flight path deviations satisfy the Gaussian probability distribution (6) for aircraft with dissimilar variances of position errors:

$$Q_0 \geq Q_2\left(L;\sigma_1,\sigma_2\right)=\left(2\pi\sigma_1\sigma_2\right)^{-1}\int_{-\infty}^{+\infty}\exp\left\{-\left[\left(x/\sigma_1\right)^2+\left(\left(L-x\right)/\sigma_2\right)^2\right]/2\right\}dx. \quad (27)$$

The integral in (27) does not need splitting to be evaluated, e.g. in the case of equal variances:

$$\begin{aligned}\sigma_1=\sigma_2\equiv\sigma: \qquad Q_0 \geq Q_2\left(L;\sigma\right)&=\left(2\pi\sigma^2\right)^{-1}\int_{-\infty}^{+\infty}\exp\left\{-\left[x^2+\left(L-x\right)^2\right]/\left(2\sigma^2\right)\right\}dx,\\ &=\left(2\pi\sigma^2\right)^{-1}\exp\left[-L^2/2\sigma^2\right]\int_{-\infty}^{+\infty}\exp\left\{-\left(x^2-xL\right)/\sigma^2\right\}dx,\end{aligned} \quad (28)$$

the change of variable (29a):

$$y=\left(x-L/2\right)/\sigma: \qquad \int_{-\infty}^{+\infty}\exp\left(-y^2\right)dy=\sqrt{\pi}, \quad (29a,b)$$

leads to a Gaussian integral (29b), viz.:

$$Q_2\left(L;\sigma\right)=\left(2\pi\sigma^2\right)^{-1}\exp\left[-L^2/\left(2\sigma^2\right)\right]\int_{-\infty}^{+\infty}\exp\left(-y^2+L^2/4\sigma^2\right)dy; \quad (30)$$

using (29b) in (30) leads to:

$$Q_2\left(L;\sigma\right)=\left(2\sqrt{\pi}\sigma^2\right)^{-1}\exp\left[-\left(L/2\sigma\right)^2\right]\leq Q_0=8\times10^{-12}\ nm^{-1}, \quad (31)$$

as the safety criterion.
In the more general case (27) of aircraft with dissimilar r.m.s. position errors:

$$Q_2=\left(2\pi\sigma_1\sigma_2\right)^{-1}\exp\left(-L^2\sigma_2^{-2}/2\right)\int_{-\infty}^{+\infty}\exp\left\{-\left[\left(x^2/2\right)\left(\sigma_1^{-2}+\sigma_2^{-2}\right)-xL\sigma_2^{-2}\right]\right\}dx, \quad (32)$$

the change of variable:

$$y=\left[x\sqrt{\sigma_1^{-2}+\sigma_2^{-2}}-L\sigma_2^{-2}/\sqrt{\sigma_1^{-2}+\sigma_2^{-2}}\right]/\sqrt{2}, \quad (33)$$

leads again to a Gaussian integral (29b), viz.:

$$Q_2=\left(2\pi\sigma_1\sigma_2\right)^{-1}\exp\left(-L^2\sigma_2^{-2}/2\right)\exp\left\{L^2\sigma_2^{-4}/\left[2\left(\sigma_1^{-2}+\sigma_2^{-2}\right)\right]\right\}\int_{-\infty}^{+\infty}\exp\left(-y^2\right)dy, \quad (34)$$

which simplifies the safety condition to:

$$Q_0 \geq Q_2\left(L;\sigma_1,\sigma_2\right)=\left(2\sqrt{\pi}\sigma_1\sigma_2\right)^{-1}\exp\left\{-\left(L^2/2\right)/\left[\left(\sigma_1\right)^2+\left(\sigma_2\right)^2\right]\right\}. \quad (35)$$

This reduces to (31) in the case of equal r.m.s. position errors.

2.4 Generalized error or Gaussian distribution

The safety condition (12c) for (10) the more accurate (8) generalized probability distribution:

$$c \equiv \sqrt[4]{120}:$$

$$Q_0 \geq Q_3\left(L;\sigma_1,\sigma_2\right)=\left[15/\left(2\sigma_1\sigma_2\right)\right]\int_{-\infty}^{+\infty}\exp\left\{-c\left[\left|x/\sigma_1\right|^{1/2}+\left|\left(L-x\right)/\sigma_2\right|^{1/2}\right]\right\}dx, \tag{36}$$

requires again a split in the region of integration as for the Laplace distribution (Section 2.2), with the difference that the evaluation of integrals is not elementary. The contribution to the cumulative probability of coincidence of the position between the flight paths of the two aircraft is:

$$Q_{31}=\left[\left(15/2\sigma_1\sigma_2\right)\right]\int_0^L\exp\left\{-c\left[\sqrt{x/\sigma_1}+\sqrt{\left(L-x\right)/\sigma_2}\right]\right\}dx, \tag{37a}$$

$$=\frac{15}{2\sigma_1\sigma_2}\sum_{n=0}^{\infty}\frac{\left(-\right)^n c^n}{n!}\int_0^L\left\{\sqrt{x/\sigma_1}+\sqrt{\left(L-x\right)/\sigma_2}\right\}^n dx, \tag{37b}$$

where the exponential was expanded in power series, and binomial theorem:

$$\left[\sqrt{x/\sigma_1}+\sqrt{\left(L-x\right)/\sigma_2}\right]^n=\sum_{m=0}^{n}\left\{n!/\left[m!\left(n-m\right)!\right]\right\}\left(x/\sigma_1\right)^{m/2}\left[\left(L-x\right)/\sigma_2\right]^{(n-m)/2}, \tag{38}$$

can also be used:

$$Q_{31}=\frac{15}{2\sigma_1\sigma_2}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{\left(-\right)^n c^n}{m!\left(n-m\right)!}\sigma_1^{-m/2}\sigma_2^{-(n-m)/2}\,\mathrm{I}_{n,m}, \tag{39a}$$

and $\mathrm{I}_{n,m}$ denotes the integral:

$$\mathrm{I}_{n,m}\equiv\int_0^L x^{m/2}\left(L-x\right)^{(n-m)/2}dx, \tag{39b}$$

which can be reduced to an Euler's Beta function. The Beta function (40a) is defined (Whittaker & Watson, 1927) by:

$$B\left(\alpha,\beta\right)\equiv\int_0^1 y^{\alpha-1}\left(1-y\right)^{\beta-1}dy=\Gamma\left(\alpha\right)\Gamma\left(\beta\right)/\Gamma\left(\alpha+\beta\right), \tag{40a,b}$$

and can be evaluated (40b) in terms of Gamma functions (Goursat, 1950). The integrals (39b) are evaluated in terms of the Beta function via a change of variable.

$$y\equiv x/L:$$

$$L^{-1-n/2}\,\mathrm{I}_{nm}=\int_0^1 y^{m/2}\left(1-y\right)^{(n-m)/2}dy=B\left(1+m/2,1+\left(n-m\right)/2\right) \tag{41a,b,c}$$

$$=\Gamma\left(1+m/2\right)\Gamma\left(1+\left(n-m\right)/2\right)/\Gamma\left(2+n/2\right).$$

Substitution of (41c) in (39a) yields:

$$Q_{31} = \frac{15L}{2\sigma_1\sigma_2} \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{(-)^n c^n}{m!(n-m)!} \left(\frac{L}{\sigma_1}\right)^{m/2} \left(\frac{L}{\sigma_2}\right)^{(n-m)/2} \frac{\Gamma(1+m/2)\,\Gamma(1+n/2-m/2)}{\Gamma(2+n/2)}, \quad (42)$$

as the first contribution to (36).
Since (42) may be expected to be the main contribution to (36), we seek upper bounds for the two remaining contributions. The second contribution to (36) concerns coincidence outside the path of the second aircraft:

$$Q_{32} = \left[15/(2\sigma_1\sigma_2)\right] \int_L^{\infty} \exp\left\{-c\left[\sqrt{x/\sigma_1} + \sqrt{(x-L)/\sigma_2}\right]\right\} dx; \quad (43a)$$

an upper bound is obtained by replacing $x \geq L$ by L in the first exponential:

$$Q_{32} \leq \left[15/(2\sigma_1\sigma_2)\right] \exp\left(-c\sqrt{L/\sigma_1}\right) \int_L^{\infty} \exp\left[-c\sqrt{(x-L)/\sigma_2}\right] dx, \quad (43b)$$

the change of variable (44a) leads:

$$y = c\sqrt{(x-L)/\sigma_2}, \qquad Q_{32} \leq \frac{15}{\sigma_1 c^2} \exp\left(-c\sqrt{L/\sigma_1}\right) \int_0^{\infty} e^{-y} y\, dy, \quad (44a,b)$$

to an integral (44b) which is evaluated in terms (Whittaker & Watson, 1927; Goursat, 1950) of the Gamma function:

$$\int_0^{\infty} e^{-y} y^n dy = \Gamma(n+1) \equiv n!; \quad (45a)$$

using (45a) in (44b) leads to the upper bound for the second contribution to (36), viz.:

$$Q_{32} \leq \left[15/\left(\sigma_1 c^2\right)\right] \exp\left(-c\sqrt{L/\sigma_1}\right). \quad (45b)$$

The third contribution to (36) corresponds to coincidence outside the flight path of the first aircraft:

$$Q_{33} = \left[15/(2\sigma_1\sigma_2)\right] \int_{-\infty}^{0} \exp\left\{-c\left[\sqrt{-x/\sigma_1} + \sqrt{(L-x)/\sigma_2}\right]\right\} dx, \quad (46a)$$

$$= \left[15/(2\sigma_1\sigma_2)\right] \int_0^{\infty} \exp\left(-c\sqrt{x/\sigma_1}\right) \exp\left(-c\sqrt{(L+x)}/\sigma_2\right) dx; \quad (46b)$$

an upper bound is obtained by replacing in the second exponential $L + x \geq L$ by L:

$$Q_{33} \leq \left[15/(2\sigma_1\sigma_2)\right] \exp\left(-c\sqrt{L/\sigma_2}\right) \int_0^{\infty} \exp\left(-c\sqrt{x/\sigma_1}\right) dx. \quad (46c)$$

The last integral is evaluated via a change of variable:

$$y = c\sqrt{x/\sigma_1}: \qquad Q_{33} \leq \left[15/\left(\sigma_2 c^2\right)\right] \exp\left(-c\sqrt{L/\sigma_2}\right) \int_0^\infty e^{-y} y\, dy, \tag{47a}$$

leading by (45a) to:

$$Q_{33} \leq \left[15/\left(\sigma_2 c^2\right)\right] \exp\left(-c\sqrt{L/\sigma_2}\right). \tag{47b}$$

If the upper bounds (45b) and (47b) are small relative to the first contribution (42) to (36), viz.:

$$Q_{31} >> \left(15/c^2\right)\left[\sigma_1^{-1} \exp\left(-c\sqrt{L/\sigma_1}\right) + \sigma_2^{-1} \exp\left(-c\sqrt{L/\sigma_2}\right)\right] \geq Q_{32} + Q_{33}, \tag{48a}$$

then (46) alone can be used in the safety criterions (12c), viz.:

$$8 \times 10^{-12}\, nm^{-1} = Q_0 \geq Q_{31}, \tag{48b}$$

with an error whose upper bound is specified by the ratio of the r.h.s. to l.h.s. of (48a). If the latter error is not acceptable, then (43a) and (46b) must be evaluated exactly. Concerning the second contribution (43a) to (36), the change of variable (49a):

$$x = L \cosh^2 \alpha, \quad x - L = L \sinh^2 \alpha, \tag{49a,b}$$

implies (49b), and transforms (43a) to:

$$Q_{32} = \left[15\, L/\left(\sigma_1 \sigma_2\right)\right] \int_0^\infty d\alpha \cosh\alpha \sinh\alpha \exp\left\{-c\sqrt{L}\left(\sigma_1^{-1/2} \cosh\alpha + \sigma_2^{-1/2} \sinh\alpha\right)\right. \tag{49c}$$

Concerning the third contribution (46b) to (36) the change or variable (50a):

$$x = L \sinh^2 \alpha, \qquad x + L = L \cosh^2 \alpha, \tag{50a,b}$$

implies (50b), and leads to:

$$Q_{33} = \left[15L/\left(\sigma_1 \sigma_2\right)\right] \int_0^\infty d\alpha \sinh\alpha \cosh\alpha \exp\left\{-c\sqrt{L}\left(\sigma_1^{-1/2} \sinh\alpha + \sigma_2^{-1/2} \cosh\alpha\right)\right\}, \tag{50c}$$

which is similar to (49c) interchanging σ_1 with σ_2. The integrals (49c) and (50c) can be evaluated numerically, and coincide in the case of equal r.m.s. position errors:

$$\sigma_1 = \sigma_2 \equiv \sigma: \qquad Q_{32} = Q_{33} = \frac{15L}{4\sigma^2} \int_0^\infty \exp\left(-c\sqrt{L/\sigma}\, e^{\alpha}\right)\left(e^{2\alpha} - e^{-2\alpha}\right) d\alpha. \tag{51a}$$

A further change of variable (51b) yields:

$$\begin{aligned} & y = c\sqrt{L/\sigma}\, e^{\alpha}: \\ & Q_{32} + Q_{33} = \frac{15L}{2\sigma^2} \int_{c\sqrt{L/\sigma}}^\infty e^{-y} \left\{\left[\sigma/\left(c^2 L\right)\right] y - \left(c^2 L/\sigma\right) y^{-3}\right] d\, y. \end{aligned} \tag{51b}$$

The exponential integral of order n is defined (Abramowitz & Stegun, 1965) by:

$$E_n(z) = \int_z^\infty y^n e^{-y} dy\,, \tag{52a}$$

and allows evaluation of (51b), viz.:

$$Q_{32} + Q_{33} = \left[15L/\left(2\sigma^2\right)\right]\left\{\left[\sigma/\left(c^2L\right)\right]E_1\left(c\sqrt{L/\sigma}\right) - \left[\left(c^2L\right)/\sigma\right]E_{-3}\left(c\sqrt{L/\sigma}\right)\right\}. \tag{52b}$$

The sum of the three contributions (42) plus (49c) and (50c) or (52b), specifies:

$$8\times10^{-12}\,nm^{-1} = Q_0 \geq Q_3\left(L;\sigma_1,\sigma_2\right) = Q_{31} + Q_{32} + Q_{33}\,, \tag{52c}$$

as the safety condition.

3. Application to four ATM scenarios

The preceding safety-separation criteria are applied to the four major airway scenarios, viz. lateral separation in uncontrolled (Section 3.1) and controlled (Section 3.2) airspace and standard (Section 3.3) and reduced (Section 3.4) vertical separation.

Probability distribution		Laplace	Gauss	Generalized
quantity	σ_a	Q_{1a}	Q_{2a}	Q_{3a}
Unit	nm	-	-	-
10	nm	2,42E-04	5,45E-06	3,80E-04
5	nm	7,72E-07	1,57E-13	3,58E-05
4	nm	3,47E-08	1,91E-19	1,28E-05
3	nm	1,68E-10	2,17E-32	2,75E-06
2	nm	2,84E-15	9,77E-70	1,92E-07
1	nm	4,95E-30	1,04E-272	3,88E-10
0,5	nm	3,84E-60	0,00E-00	4,70E-14

Table 1. Lateral a CPC for the Laplace, Gaussian and generalized probabilities.

3.1 Lateral separation in oceanic airspace

The lateral separation in oceanic airspace is (53a):

$$L_a = 50\,nm, \qquad \sigma_a = 0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 10,\ nm, \tag{53a,b}$$

and the r.m.s. position error is given the values (53b) in Table 1, where the CPC are indicated for the Laplace, Gaussian and generalized probabilities. Taking as reference the generalized probability distribution, that is the most accurate representation of large flight

path deviation considerably underestimates the risk of collision, and the Laplace distribution although underestimating less is still not safe. For example the ICAO ATLS of 8×10^{-12}/nm is met according to the generalized probability distribution for a r.m.s. position deviation $\sigma_a \le 1$ nm; the Laplace distribution would give $\sigma_a \le 3$ nm and the Gaussian $\sigma_a \le 5$ nm. The latter are both unsafe, because for $\sigma_a = 3$ nm the generalized distribution gives a collision probability 2.75×10^{-6}/nm and for $\sigma_a = 5$ nm it gives 3.58×10^{-5}/nm and both significant exceed the ICAO ATLS.

3.2 Lateral separation in controlled airspace

In controlled airspace the lateral separation (53a) is reduced to (54a):

$$L_b = 5\,nm, \qquad \sigma_b = 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1.0\,nm, \tag{54a,b}$$

and the r.m.s. position errors considered (54b) are correspondingly smaller than (53b). Again the generalized distribution meets the ICAO ATLS for a r.m.s. deviation $\sigma_b \le 0.05$ nm smaller than predicted by the Laplace ($\sigma_b \le 0.2$ nm) and Gaussian ($\sigma_b \le 0.5$ nm) distributions. For the safe r.m.s deviation $\sigma_b = 0.05$ nm the Gaussian probability of collision is negligible.

Probability distribution		Laplace	Gauss	Generalized
quantity	σ_b	Q_{1b}	Q_{2b}	Q_{3b}
Unit	nm	-	-	-
1,0	nm	2,42E-03	5,45E-04	3,80E-03
0,5	nm	7,72E-06	1,57E-11	3,58E-04
0,4	nm	3,47E-07	1,91E-17	1,28E-04
0,3	nm	1,68E-09	2,17E-30	2,75E-05
0,2	nm	2,84E-14	9,77E-68	1,92E-06
0,1	nm	4,95E-29	1,04E-270	3,88E-09
0,05	nm	3,84E-59	0,00E-00	4,70E-13

Table 2. Lateral *b* CPC for the Laplace, Gaussian and generalized probabilities.

3.3 Vertical separation in oceanic airspace

The probabilities of vertical separation can be less upward than downward, due to gravity, proximity to the service ceiling, etc.; apart from this correction (Campos & Marques, 2007, 2011), the preceding theory can be used with the standard vertical separation (55a):

$$L_c = 2000\ ft, \qquad \sigma_c = 40, 50, 100, 200, 300\ ft, \tag{55a,b}$$

and r.m.s. deviations (55b). The r.m.s. height deviation that meets the ICAO ATLS is about 40 ft according to the generalized distribution, with larger and unsafe predictions for the Laplace (100 ft) and Gaussian (200 ft) distributions.

Probability distribution		Laplace	Gauss	Generalized
Quantity	σ_c	Q_{1c}	Q_{2c}	Q_{3c}
Unit	ft	-	-	-
300	ft	9,88E-07	4,68E-11	4,03E-06
200	ft	1,93E-08	9,79E-17	8,76E-07
100	ft	5,39E-14	1,05E-48	2,11E-08
50	ft	1,10E-25	2,16E-178	8,12E-11
40	ft	1,24E-31	6,49E-276	8,21E-12

Table 3. Vertical *a* CPC for the Laplace, Gaussian and generalized probabilities.

3.4 Reduced vertical separation

The RSVM (Figure 5) introduced by Eurocontrol in upper European air space halves the vertical separation (56a) to (58a):

$$L_d = 1000\,ft, \qquad \sigma_d = 15, 50, 100, 150\ ft, \tag{56a,b}$$

and the r.m.s. position errors are correspondingly reduced from (56b) to (58b) in Table 4.

Probability distribution		Laplace	Gauss	Generalized
quantity	σ_d	Q_{1d}	Q_{1d}	Q_{3d}
Unit	ft	-	-	-
150	ft	1,98E-06	1,87E-10	8,05E-06
100	ft	3,86E-08	3,92E-16	1,71E-06
50	ft	1,08E-13	4,20E-48	4,04E-08
15	ft	2,55E-41	0,00E-00	6,86E-13

Table 4. Vertical *b* CPC for the Laplace, Gaussian and generalized probabilities.

Taking as reference the generalized distribution to meet the ICAO ATLS: (i) the RVSM from 2000 ft (Table 3) to 1000 ft (Table 4) requires a reduction in r.m.s. altitude error from 40 ft to 15 ft; (ii) the reduction of lateral separation from 50 nm in transoceanic (Table 1) to 5 nm in controlled (Table 2) airspace required a reduction of r.m.s. deviation from 0.5 to 0.05 nm.

4. Discussion

The separation-position accuracy or technology-capacity trade-off was made for an air corridor ATM scenario with aircraft flying along the same flight path (Figure 2) or on parallel flight paths (Figure 1) with a constant separation. The generalized probability distribution leads to lower values of the r.m.s. deviation to meet the ICAO TLS, than the Laplace and Gaussian. Although the latter distributions are simpler, they underestimated

the collision risk, and do not yield safe predictions. Using simultaneously lateral and vertical separations leads to much lower collision probabilities, and allows reducing each separation for the same overall safety. In the case of aircraft flying on parallel tracks, it is desirable to use alternate directions of flight (Figure 5), because: (i) adjacent flight paths correspond to aircraft flying in opposite directions, which spend less time close to each other, reducing the collision probability (Campos & Marques, 2002; Eurocontrol, 1988; Reich, 1966); (ii) the aircraft which spend more time 'close' by are on a parallel track at twice the separation 2L, thus allowing a larger r.m.s. position error σ for the same safety. If the aircraft have both altitude and lateral separation, and the two position errors are statistically independent, the ICAO ATLS is $\sqrt{8\times10^{-12}}\ /\ nm \approx 2.8\times10^{-6}\ /\ nm$ in each direction. For transoceanic flight this is met by a lateral r.m.s. deviation $\sigma_l \leq 3$ nm; for flight in controlled airspace with RVSM the ICAO ATLS wold be met with lateral $\sigma_l \leq 0.2$ nm and altitude $\sigma_h \leq 150$ ft r.m.s. deviations. Using also along track or longitudinal separation adds a third dimension, requiring a smaller ICAO ATLS $\sqrt[3]{8\times10^{-12}}\ /\ nm = 2\times10^{-4}\ /\ nm$ and allowing larger r.m.s. deviations in three directions.

5. References

Abramowitz, M. & I., Stegun, (1965), *Tables of Mathematical Functions*, Dover.

Anderson, E.W., (1966), *The Principles of Navigation*, Hollis & Carter.

Barnett, A. (2000), Free Flight and En-route Air Safety a First-order Analysis, *Operations Research*, Vol. 48, pp. (833-845).

Bousson, K., (2008), Model predictive control approach to global air collision avoidance, *Aircraft Engineering and Aerospace Technology*, Vol. 80, (6), pp. (605-612).

Campos, L.M.B.C., (1984), On the Influence of Atmospheric Disturbances on Aircraft Aerodynamics, *Aeronautical Journal,* June/July, pp. (257-264).

Campos, L.M.B.C., (1986), On the Aircraft Flight Performance in a Perturbed Atmosphere, *Aeronautical Journal*, Paper 1305, Oct., pp. (301-312).

Campos, L.M.B.C., (1997), On the Non-linear Longitudinal Stability of Symmetrical Aircraft, *Journal of Aircraft*, Vol. 36, pp. (360-369).

Campos, L.M.B.C., (2001), On the probability of Collision Between Aircraft with Dissimilar Position Errors, *Journal of Aircraft*, 38, pp. (593-599).

Campos, L.M.B.C. & Marques, J.M.G., (2002), On Safety Metrics Related to Aircraft Separation, *Journal of Navigation*, Vol. 55, pp. (39-63).

Campos, L.M.B.C. & Marques, J.M.G., (2004a), On the Combination of the Gamma and Generalized Error Distribution With Application to Aircraft Flight Path Deviation, *Communications in Statistics-Theory and Methods*, Vol. 33, Nº10, pp. (2307-2332).

Campos, L.M.B.C. & Marques, J.M.G., (2004b), On Wake Vortex Response for All Combinations of Five Classes of Aircraft, *Aeronautical Journal*, June, pp. (295-310).

Campos, L.M.B.C. & Marques, J.M.G. (2007) On the Probability of Collision Between Climbing and Descending Aircraft, *Journal of Aircraft*, Vol. 44 (2), pp. (550-557).

Campos, L.M.B.C. & Marques, J.M.G., (2011), On the Probability of Collision for Crossing, *Aircraft Eng. Aerospace Technology* (to appear).

Etkin, B., (1981), The Turbulent Wind and its Effects on Flight, *Journal of Aircraft*, Vol. 18, Nº5.

Etkin, B. & Reid, L.D., (1996), *Dynamics of Flight: Stability and Control*, Wiley.

Etkin, B. & Etkin, D.A., (1990), Critical Aspects of Trajectory Prediction Flight in a non-Uniform Wind, Agardograph AG-301, Vol.1.

Eurocontrol, (1988), European Studies of Vertical Separation Above FL290 - Summary Report, Brussels, Oct.

Eurocontrol, (1998), Air Traffic Management Strategy for 2000+, November.

Eurocontrol, (2000), Objective Measures of ATM System: Safety Metrics, C-Integra Report to Eurocontrol.

FAA, (2011), Aeronautical Information Manual (AIM), Official Guide to Basic Flight Information and ATC Procedures.

Goursat, E., (1950), *Course of Analysis*, Dover.

Houck, S. & Powell, J.D., (2001), Assessment of Probability of Mid-air Collision During an Ultra Closely Spaced PParallel Approach, AIAA Paper 2001-4205.

International Civil Aviation Organization, (2005), Air Traffic Services, Annex 11.

International Civil Aviation Organization, (2006), Rules of the Air, Annex 2.

International Civil Aviation Organization, (2007), Procedures for Air Traffic Management, Doc 4444.

Johnson, N.L. & Balakrishnan, N., (1995), *Continuous Univariate Probability Distributions*, Wiley 1995.

Mises, R.V., (1960), *Theory of Probability and Statistics*, Academic Press.

Nassar, M.M., Khamis, S.M. & Radwan, S.S., (2011), On Bayesian sample size determination, *Journal of Applied Statistics*, Vol. 38, (5), pp. (1045-1054).

Reich, P.G., (1966), Analysis of Long-range Air Traffic Systems: Separation Standards. *Journal of Navigation,* Nº19, pp. (88-98), pp. (169-186), pp. (331-347).

Reiss, R.D. & Thomas, M., (2001), *Statistical Analysis of Extreme Values.* Birkhauser Verlag.

Rossow, V.J., (1999), Lift-generated Vortex Wakes of Subsonic Transport Aircraft, *Progress in Aerospace Sciences*, Vol. 35, pp. (507-660).

Shortle, J.F., Xie, Y., Chen, C.H. & Dunohue, G.L., (2004), Simulating Collision Probabilities of Landing Airplanes at Non-towered Airports. *Simulation*, Nº80, pp. (21-31).

Shortle, J.F., Zhang, Y., Wang, J., (2010), Statistical Characteristics of Aircraft Arrival Tracks, *Journal of the Transportation Research Board*, Vol. 2177, pp. (98-104).

Spalart, P.R., (1998), Airplane Trailing Vortices. *Annual Review of Fluid Mechanics*, Nº30, pp. (107-138).

Vismari, L.F. & Júnior, J.B.C., (2011), A Safety Assessment Methodology Applied to CNS/ATM-Based Air Traffic Control System, *Reliability Engineering & System Safety*, Vol. 96, pp. (727-738) .

Whittaker, E.T. & Watson, G.N., (1927), *Course of Modern Analysis*, Cambridge University Press.

Yuling, Q. & Songchen, H., (2010), A Method to Calculate the Collision Risk on Air-Route, Management and Service Science (MASS), Proceedings of IEEE International Conference.

Zhang, Y. & Shortle, J., (2010), Comparison of Arrival Tracks at Different Airports, Proceedings of 4th International Conference on Research in Air Transportation - ICRAT 2010, Budapest Hungary, June 01 – 04.

23

Development of an Aircraft Routing System for an Air Taxi Operator

F.M. van der Zwan[1], K. Wils[1] and S.S.A. Ghijs[2]
[1]*Delft University of Technology*
[2]*Fly Aeolus*
[1]*The Netherlands*
[2]*Belgium*

1. Introduction

Due to increasing congestion of the road network and the major airports, the established modes of transport like car and scheduled air travel are having difficulty to fulfil the need for efficient business travel. Therefore, recent years have seen a growing demand for business aviation services like business aircraft charter and (fractional) aircraft ownership. These services enable improved time efficiencies, access to a wider range of airports, and the ability to control flight scheduling (Budd & Graham, 2009). However, business aviation is characterized by high prices and due to the economic downturn of 2008 corporate travel departments now seek less costly and more efficient business travel solutions (American Express, 2009; Gall & Hindhaugh, 2009). This has paved the way for the development of a new phenomenon in the air transport industry, namely the on-demand air taxi service. Air taxi operators offer travellers a low fare, on-demand travel service (Budd & Graham, 2009). As such, they provide the flexibility and time efficiency of existing on-demand business aviation, giving them the same time-efficiency advantage over car and scheduled air transport, but at a lower price (Bonnefoy, 2005).
One of the key enablers that allows air taxi operators to offer lower fares and still remain profitable is, apart from their smaller and thus cheaper aircraft, the optimization of resources, particularly during the employment of the aircraft during the aircraft routing phase. This chapter therefore sets out to develop an aircraft routing system which allows a full on-demand per-aircraft air taxi company to generate aircraft routing plans which adhere to its planning objectives, during the first operational years. Firstly, the air taxi business model is discussed. Then the aircraft routing problem and modelling approaches are described, followed by the design, testing and validation of the aircraft routing model, after which conclusions on the approach taken are drawn.

2. The air taxi business model

This section discusses the air taxi business model, the planning process for on-demand air transport (ODAT) and how this differs from scheduled air transport services.

2.1 Air taxi characteristics

Existing air taxi services can be differentiated with respect to their offer. Both per-seat and per-aircraft air taxi services exist. In the former the customer buys a single seat on a flight,

while in the latter the customer charters a whole aircraft. Also 'semi on-demand' and 'full on-demand' air taxi services exist. In the former the air taxi service only operates from a predetermined set of airports, while in the latter the customer can freely choose the departure and arrival airport. Another difference lies in the type of aircraft used; operators either fly very light jets (VLJs) or small piston aircraft. A VLJ typically seats between four and seven passengers, can be flown by a single pilot, has a maximum take-off weight of less than 10,000 lb (4536 kg), and an average range of around 2,000 km (Budd & Graham, 2009, p.289). Piston aircrafts have one or more piston-powered engines connected to the propeller(s). Piston aircraft used for business typically seats one to six passengers and fly relatively short missions of 300-400 miles, using very small general aviation airports that are often without air traffic control towers (NBAA, 2010). The characteristics of the air taxi business model are summarised in Table 1 below.

Focused outside hub airports
Faster than scheduled air carriers for most point-to-point travel
One to six passengers per aircraft (though some take more)
High utilization of aircraft (via optimization of resources)
Cheaper than a charter or a fractionally owned aircraft
On-demand or semi on-demand (no fixed or published schedules)
Sold by the plane or per-seat
Clear, all-in pricing, by distance, time or zone
Using small piston or VLJ aircraft

Table 1. Characteristics of the air taxi business model (after Dyson, 2006)

As mentioned at the beginning of this chapter, air taxi services offer the flexibility and time efficiency of existing on-demand business aviation but at a lower price. Prices for per-aircraft air taxi (PAAT) operators in Europe range from € 2,500 per flight hour (London Executive Aviation, 2010) to € 800 per flight hour (FlyAeolus, 2010). Bonnefoy (2005) indicates that air taxi operators are able to offer these lower fares due to their use of a small aircraft single type fleet and optimization of resources. The small aircraft used by air taxi operators have lower acquisition and operating costs than jets used by existing business aviation (Bonnefoy, 2005). Also, Budd and Graham (2009) and Espinoza et al. (2008a) note that the landing and take-off characteristics of these smaller aircraft enable access to an even greater range of airports, which possibly lie closer to where business travellers want to go. Apart from the smaller aircraft, optimization of resources is also key to the cost feasibility of the air taxi model (Dyson, 2006; Mane & Crossley, 2009). The resource optimization challenges faced by air taxi operators present themselves during the operational planning and management of the day-to-day operations.

2.2 Planning process for air taxi model

The airline planning process for scheduled air transport typically consists of five phases: flight scheduling, fleet assignment, aircraft routing, crew scheduling and crew rostering (Bazargan, 2004). The planning problem for on-demand air travel has a different nature than that of scheduled airlines (Hicks et al., 2005; Yao, 2007). The primary cause for this is the demand mechanism (Yao, 2007): scheduled airlines decide on their flight schedule months in advance, while in contrast, on-demand air transport operators like air taxis and fractional management companies sometimes know their flight requests only several hours in advance

since customers can book up to six or eight hours before departure (Hicks et al., 2005; Yang et al., 2010). In addition, the flight legs flown by on-demand operators are less predictable and differ from day to day and week to week (Ronen, 2000; Yao et al., 2008). However, one can still distinguish the different phases of the scheduled airline's planning process in the planning process of on-demand air travel operators.

The **flight scheduling** phase for on-demand air transportation is customer driven. The customer contacts the air taxi operator a couple of days or hours in advance with his flight request. This process leads to an unpredictable and non-repeatable flight schedule. The customer's flight request specifies a departure location and departure time and an arrival location. Air taxi companies have the option to reject a request. The accept/reject decision depends on a number of factors, for example whether it is possible to fly the request and also whether is worthwhile flying the flight, both financially and strategically. This assessment should be performed with respect to the flight request itself, with respect to expected demand (Fagerholt et al., 2009), and also future customer demand as a customer denied service may choose not to request future trips (Mane & Crossley, 2007a). During the flight scheduling phase it must also be decided when and where aircraft will undergo scheduled maintenance (Keysan et al., 2010). If the request is accepted, the flight leg(s) of the customer's flight request are added to the flight schedule.

The second phase of the scheduling process considers **fleet assignment**, where fleet types (not specific aircraft) are assigned to each flight leg in the schedule (Bazargan, 2004). For an air taxi company, this phase is virtually non-existent as they mostly operate a single type fleet (Bonnefoy, 2005; Dyson, 2006).

The next phase is **aircraft routing**. During this phase each individual aircraft is assigned a routing, which is a sequence of flight legs, so that each leg is covered exactly once while ensuring that the aircraft visits maintenance stations at regular intervals thereby fulfilling maintenance requirements (Barnhart et al., 2003). The aircraft routing phase for on-demand air transport operators consists of assigning specific aircraft to the customer requested flight legs (Yao, 2007). During aircraft routing the operator thus aims to create aircraft routes that cover all the flight requests while minimizing the operational cost and adhering to the operational constraints, related to the aircraft's maintenance limit, crew regulations and the availability of fuelling facilities at airports (Martin et al., 2003; Yao et al., 2008). As departure and arrival locations can be chosen freely by the customer it is unlikely that an aircraft is always directly available at the departure location. Therefore an ODAT operator will have to conduct repositioning flights, also called 'deadhead' or 'non revenue' flights, to reposition their aircraft to the departure locations. Yao et al. (2008) mentions that deadhead flights may represent over 35% of the total flying conducted by a fractional management company (FMC). Air taxis and FMCs mostly need to bear the cost of repositioning themselves since the customer only pays for his actual flight request (Yang et al., 2008). Therefore, as a major driver during the aircraft routing phase is minimizing operational cost, it is important to minimize repositioning flights, both in number and in length (Bonnefoy, 2005; Mane and Crossley, 2007b; Yao, 2007). It must be noted that per-seat on-demand operators face an additional challenge during the aircraft routing phase. Apart from planning so that all customer requests are served, customer requests should be interleaved where possible such that that different customers can be put on the same aircraft, thereby obtaining a load factor sufficiently high to remain profitable (Espinoza et al., 2008a). This greatly increases the complexity of the underlying aircraft routing problem.

The two last phases in the scheduled airline's planning process involve the planning of crew resources. The fourth phase is **crew scheduling** or crew pairing. This treats the process of identifying sequences of flight legs that start and end at the same crew base, i.e. trip pairings (Bazargan, 2004). While creating trip pairings one strives for a minimum-cost set of pairings so that every flight leg is covered and that every pairing satisfies the applicable crew work rules (Bazargan, 2004). Finally, in the **crew rostering** phase, crew pairings are combined into monthly schedules which are then assigned to individual crew members (Bazargan, 2004). The crew pairing and crew rostering phase for ODAT operators are different from those used by scheduled airlines. As the definitive flight schedule is not known beforehand, no accurate monthly crew schedules can be created. Therefore crew of ODAT operators work with on-duty periods, during which they must be ready to operate flight legs which may not yet be known at the start of the on-duty period, and off-duty periods (Hicks et al., 2005; Yao et al., 2008). Yao et al. (2008) mention that, as aircraft are not always stationed at crew bases, crews coming on duty may need to travel to the aircraft's location, which counts as on-duty time and (possibly) needs to be paid for by the ODAT operator. Needless to say, ODAT crew schedules need to meet crew regulations as stipulated by the air transport authorities.

With respect to the demand uncertainty of ODAT, Hicks et al. (2005) report, based on the operations of an FMC, that approximately 80% of the trips are requested 48 hours or more prior to departure and 20% with as little as 4 hours notice. In addition, 30% of trips are changed at least once within 48 hours of the requested departure time (Hicks et al., 2005). Historical data analysis on FMC operations shows that on average 5% of demand is unknown for the first day for which the schedule is created, and 10-20% and 25-40% for the second and the third day respectively (Yao et al., 2007, 2008). Despite the uncertainty in demand, ODAT operators need to create advance flight schedules, aircraft routings and crew schedules. There are a number of reasons for this: some airports require arrival information 24 hours in advance, the crews need to be given enough time to relocate to where they are required to come on duty, and maintenance locations need to be booked for scheduled maintenance (Yang et al., 2010). However, these advance schedules need to be adapted as more demand information gets known. For example, Yang et al. (2008) reports that around 25% of the requested trips of an FMC arrived after the original schedule was created, which causes the need to re-optimize the schedule.

Because of this dynamic process of creating and adapting the schedules it is desirable to both have a flexible and persistent schedule (Karaesmen et al., 2007; Yang et al., 2010): a flexible schedule to be able to cost-effectively cater extra demand, and a persistent schedule to avoid that the planning completely changes every time a change in demand occurs. Next to that, the schedule needs to be robust. Robustness denotes the ability of the schedule to cope with and cost-effectively recover from unforeseen changes in supply (e.g. aircraft break-down) or other disruptions (e.g. adverse weather) (Ball et al., 2007). As Bian et al. (2003) state, in a robust schedule enough slack is built in, such that a single disruption does not cause later flights to be delayed. However, a robust schedule comes at a price as schedule slack basically means that aircraft will be standing idle. Finally, the ODAT schedule needs to achieve a sufficient customer service level. This means on the one hand carefully looking at accepting or rejecting of flight requests, but also considering whether to schedule their customer's flights as requested or not (Fagerholt et al., 2009). Delaying the departure time by 30 minutes may enable a more cost-effective aircraft routing but this may

have a negative impact on the service level as perceived by the client. In summary, the ODAT operator has to balance minimising costs whilst keeping a high service level and developing an appropriate schedule (flexible, persistent, and robust).

3. The aircraft routing problem

In this section the aircraft routing problem as it applies to the air taxi operator is laid out. As explained before, the aircraft routing problem arises during the aircraft routing phase, in which the ODAT operator aims to create aircraft routes that cover all flight legs and satisfy operational restrictions while minimizing operational cost. Below, the different aspects of the aircraft routing problem are treated:

- **The scope:** planning horizon
- **The input:** flight legs (trips) and aircraft
- **The main objective:** to create an aircraft routing plan that covers all flight legs while minimizing the operating cost
- **The secondary objectives:** robustness, flexibility and persistence
- **The constraints:** operational requirements

Please note that this work has been undertaken in collaboration with Fly Aeolus, a Belgian full on-demand per-aircraft air taxi operator. This routing problem and subsequent model has been kept as generic as possible. However, in some instances the specifics of the Fly Aeolus business model have been taken as an input for the routing problem. In the text it is clearly stated how the routing problem might be (slightly) different for ODAT operators with a different business model.

An important parameter in the planning process is the **planning horizon** that is used. This determines the size, and thus complexity, of the planning problems that need to be solved during the different phases. If a 24 hour horizon is chosen the aircraft routing system will find a local optimum for that day, but the ending locations of the aircraft will not be optimized towards the start of the next day. Theoretically, the more days that are included in the planning horizon, the closer the solution will be to a global optimum. However, due to the computational complexity and the uncertainty of demand, there is a limit on the number of days that can be incorporated in the planning horizon. Ronen (2000) has tested the effect of the planning horizon by solving aircraft routing problems for 24, 36 and 48 hour horizons for a fractional management company (FMC). For his application a 24 to 36 hour planning horizon proved most effective, as planning decisions taking place beyond that horizon were always changed again later. On the other hand, Martin et al. (2003) state that the operations of the FMC they worked with allowed a two to three day planning horizon. Overall, in the ODAT literature it is generally accepted that planning beyond 72 hours is highly speculative. Therefore, the planning horizon is set to a maximum of 72 hours.

At each instance the operations control centre (OCC) wants to create an aircraft routing plan, the aircraft routing system takes the known flight schedule (output of the flight scheduling phase) and extracts the booked flight requests. Each flight request consists out of at least one **flight leg** (trip). Each of these flight legs has a departure airport, a requested time of departure, an arrival airport and a specified number of passengers for that flight. Customers also have the possibility to make a booking consisting of multiple flight legs. For these requests they are given the option to have the aircraft and crew remaining on standby at the airport in between the coupled flight legs, such that all their separate trips are served by the same aircraft. Apart from the flight requests, the aircraft routing system also retrieves

the current status of all the **aircraft** available within the available fleet. For each aircraft the home base and its lease cost per hour are known. In addition to this, it is known when and where the aircraft will become available, and the aircraft's fuel status and flight hours remaining till maintenance at that time. An aircraft can be unavailable at the start of the planning horizon because it is still in-flight at that time, or because maintenance is performed on the aircraft.

The objective of the aircraft routing system is now to assign to each aircraft a feasible sequence of flight legs in a manner that **minimizes the cost of operating all the flights** while meeting the operational requirements. Of the total operating costs, the variable direct operating costs (DOCs) are the only costs that are controlled by the aircraft routing system. This is because these are the costs that are directly related to the amount of flying that is undertaken (Doganis, 2002), which in turn depends on the aircraft routing plan is created. The following DOC items incurred by the air taxi operator are directly influenced by the aircraft routing plan: variable aircraft costs, passenger and non-passenger related airport charges, air traffic service charges, fuel costs, and variable flight crew costs. With respect to the fuel costs, these costs depend on the amount of flying that is planned. In addition, they also rely on where and how much the aircraft is planned to be fuelled. Because fuel prices vary significantly between airports (Flyer Forums, 2010), so-called fuel ferrying can in certain cases be a valid strategy to reduce fuel cost. When adopting this approach, excess fuel is carried on a flight leg to decrease the amount of fuel that needs to be bought at the next airport where fuel prices are high. The crew costs consist of the pilot wages which are directly related to how much flying the pilot carries out (as measured by the Hobbs meter of the aircraft, which registers the time that the engine is running and thus also includes taxiing and idling (e.g. when waiting for take-off clearance)). Therefore, this part of crew costs is also controlled during the aircraft routing phase.

As explained in the previous section, achieving **robustness, persistence and flexibility** while creating operational plans are additional objectives to consider during the planning process. These three secondary planning objectives are also to be considered during the design of the aircraft routing system. However, these planning objectives are interrelated and cannot be fully achieved together. For example, requiring a very robust schedule (thus a large amount of slack time), decreases the ability of the air taxi operator to accept future flight requests. During the development of the aircraft routing system it is therefore ensured that the system parameters that control robustness, persistence and flexibility can be altered by the OCC such that they can prioritize these objectives as they wish.

The aircraft routing plans that are created also have to fulfil a number of **operational requirements**. Firstly, aircraft should not exceed their maintenance flight hour limit when flying the assigned flight sequence. Secondly, flight sequences should be fuel feasible. Both because there are airports where it is not possible to fuel and because the number of passengers that is carried influences the amount of fuel that can be taken. For example, an SR22 carrying a pilot and two passengers can only carry 38% of its maximum usable fuel (Cirrus, 2007). In addition, aircraft routing plans must be crew feasible, i.e. they should not violate the crew work rules. For this purpose, it is customary in the ODAT industry to couple an aircraft with its crew for scheduling purposes (Martin et al., 2003; Yao et al., 2007; Yang et al., 2008). Hence, an aircraft is viewed as an operational unit which has to meet both its own restrictions (e.g. maintenance) and crew duty time limits. Because of this, the aircraft routing system will generate routing plans that do not necessarily require crew swaps during the day as a single crew can (legally) stay with a single aircraft throughout the whole day.

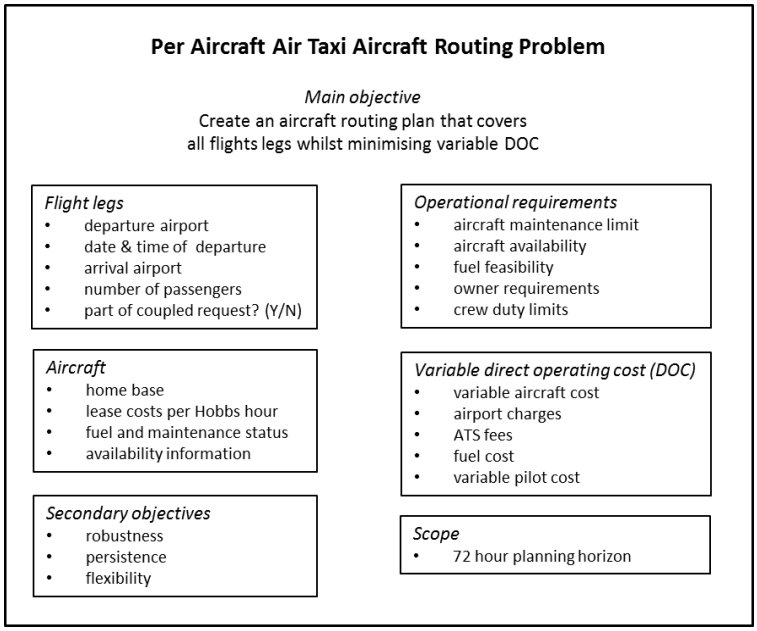


Fig. 1. Full on-demand per-aircraft air taxi routing problem

In summary, during the routing phase the air taxi operator's OCC needs to create aircraft routing plans covering all flight legs taking place during the scheduling horizon with the aim to minimize operating cost while adhering to the aircraft's maintenance flight hour limits, fuel feasibility, aircraft availability and crew duty rules, as shown in Figure 1 above.

4. Choice of modelling approach

In this section the modelling approach for the aircraft routing problem is chosen, after which the chosen methodology is further explained.

4.1 Choosing the appropriate modelling approach

To the authors' best knowledge there exists no literature on the aircraft routing problem faced by a per-aircraft air taxi (PAAT) operator. However, their aircraft routing problem is very similar to that of a FMC or charter. Both aim to minimize the operating cost while serving all the customer requests on time, and satisfying the applicable constraints. However, there are some differences between the aircraft routing problem of an FMC and that of a typical air taxi operator. First of all, since most air taxi operators operate a single type of aircraft, all its aircrafts are in principle compatible with every request. For an FMC, which operates different fleet types, there is a specific aircraft type that fits each request and this requirement must be taken into account during the aircraft routing. In addition, an FMC does not incur the flight hour related aircraft costs as a part of its variable operating costs because they either own or lease their aircraft for a continuous period of time. So their aircraft cost is irrespective of how much they use the aircraft. However, despite these differences the core of the aircraft routing problem for an FMC and a PAAT operator

remains the same. Therefore, this section focuses on studies related to FMCs to develop a system that solves the PAAT aircraft routing problem. These are listed in Table 2 below.

Modelling approach	Author(s)	Maintenance constraints	Crew constraints	Planning horizon [h]	Mathematical formulation*	Solution method#
M1	Keskinocak & Tayur (1998)	N	N	72	0 – 1 IP	CPLEX
	Keskinocak & Tayur (1998)	N	N	72	0 – 1 IP	Heuristic
M2	Ronen (2000)	Y	Y	48	SP	Heuristic
M3	Martin et al. (2003)	Y	Y	48 – 72	0 – 1 IP	CPLEX
M4	Hicks et al. (2005)	Y	Y	24 – 72	Mixed IP	CG, GENCOL
M5	Karaesmen et al. (2005)	Y	N	24	0 – 1 IP	CPLEX
M6	Karaesmen et al. (2005)	N	N	24	Mixed IP	CPLEX
	Yang et al. (2008)	N	N	24	Mixed IP	CPLEX
M7	Karaesmen et al. (2005)	Y	Y	24	0 – 1 IP	Heuristic, CPLEX
	Yang et al. (2008)	Y	Y	24	0 – 1 IP	Heuristic CPLEX
M8	Karaesmen et al. (2005)	N	Y	24	Mixed IP	CPLEX
	Yang et al. (2008)	N	Y	24	Mixed IP	CPLEX
M9	Karaesmen et al. (2005)	Y	Y	24	SP	BP, CPLEX
	Yang et al. (2008)	Y	Y	24	SP	BP, CPLEX
	Yang et al. (2008)	Y	Y	72	SP	BP, CPLEX
	Yang et al. (2008)	Y	Y	96	SP	BP, CPLEX
M10	Yao et al. (2005)	Y	Y	72	SP	CG
	Yao et al. (2008)	Y	Y	72	SP	CG
M11	Yao & Zhao (2006)	Y	Y	72	SP	na
	Yao et al. (2007)	Y	Y	72	SP	CG

* IP = integer programming, SP = set-partitioning # CG = column generation, BP = branch-and-price

Table 2. Modelling and solution approaches for the aircraft routing problem

To choose a modelling approach for a PAAT aircraft routing problem one must first specify the decision criteria that are used during the trade-off. These criteria are derived from the requirements that are posed upon the aircraft routing system and from the nature of the PAAT aircraft routing problem. The modelling approach should allow for both crew feasible and maintenance feasible routings. It should be able to accommodate a multiple day planning horizon and multiple fleet types. It should also have the flexibility to incorporate additional operational rules and finally, the flight leg cost should be dependent on the aircraft route. In Table 2 above the specifics of the different aircraft modelling approaches deployed in FMC literature are given. Since the criteria that schedules must be crew and maintenance feasible need to be fulfilled a number of modelling approaches can already be discarded. These are model 1 proposed by Keskinocak and Tayur (1998) and models 5, 6 and 8 described by Karaesmen et al. (2005) and Yang et al. (2008). Furthermore, model 7 can also be discarded since it is noted by Karaesmen et al. (2005) and Yang et al. (2008) that this modelling approach is not flexible enough to handle planning horizons longer than 24 hours. This leaves us with six models (2, 3, 4, 9, 10 and 11). All these models produce schedules that are crew and maintenance feasible and can be used in multiple day planning horizons.

However, model 10 and 11 do not incorporate the cost of flying customer trips in the operational cost of a schedule. The authors, Yao et al. (2005, 2008); Yao and Zhao (2006), state that the customer of an FMC pays for the fuel and the crew costs of their trip, thus they do not consider these trip costs as a part of operational cost. For the PAAT aircraft routing problem this is not a valid assumption. The cost for one specific flight leg namely differs depending on which aircraft routing the leg is in. Air taxi customers always pay a fixed price per flight hour, which is independent of what total cost the operator has to bear to operate this flight.
Models 2 and 9 utilize a set-partitioning formulation. In an overview of vehicle routing models, Bunte and Kliewer (2010) state that an advantage of the set-partitioning formulation, in contrast to other formulations, is that additional constraints or operational rules can be easily incorporated. Also, Ronen (2000) and Karaesmen et al. (2005) note that the set-partitioning approach is very appropriate to problems where costs are nonlinear and discrete, and complicated rules are imposed, as is the case in the PAAT aircraft routing problem. A direct comparison with a 0-1 and mixed IP formulation is made by Karaesmen et al. (2005) and Yang et al. (2008). They conclude that the set-partitioning formulation is the only one that has adequate flexibility to cover long planning horizons and to incorporate complex operational rules (Karaesmen et al., 2005). Therefore, in order to model the PAAT aircraft routing problem the set-partitioning formulation is chosen.

4.2 The set-partitioning model

For the set-partitioning formulation of the PAAT aircraft routing problem, which is adopted from Yang et al. (2008), the following notation is used:

$$
\begin{aligned}
&\Omega_i = \text{the set of feasible routes for aircraft } i, && i \in N\\
&p = \text{a feasible route}, && p \in \Omega_i\\
&c_{ip} = \text{the total cost of taking route } p \text{ by aircraft } i, && i \in N \text{ and } p \in \Omega_i\\
&a^i_{jp} = 1 \text{ if trip } j \text{ is on route } p; \quad 0 \text{ otherwise}, && i \in N,\ p \in \Omega_i \text{ and } j \in M\\
&\theta_{ip} = 1 \text{ if aircraft } i \text{ takes route } p; \quad 0 \text{ otherwise}, && p \in \Omega_i \text{ and } i \in N
\end{aligned}
$$

Where $N = \{1, 2;\ldots\ n + m\}$ is the set of all aircraft such that $\{1,\ldots, n\}$ represents the own fleet, and $\{n+1,\ldots, m\}$ represents any charter aircraft for subcontracting the trips (in case the own fleet proves insufficient to operate all the flights). For each $i \in (n + 1, \ldots, m)$, Ω_i consists of a single route that only takes trip j with a cost b_j (the cost that is incurred when subcontracting this trip). Furthermore, M denotes the set of all trips. Using this notation, the aircraft routing problem of a PAAT operator can be formulated as the following set-partitioning problem (SP):

$$
\begin{aligned}
(\text{SP}) \quad & \min \sum_{i\in N}\sum_{p\in\Omega_i} c_{ip}\theta_{ip} && && (1)\\
& \text{s.t.} \sum_{i\in N}\sum_{p\in\Omega_i} a^i_{jp}\theta_{ip} = 1, && j \in M && (2)\\
& \sum_{p\in\Omega_i} \theta_{ip} \le 1, && i \in N && (3)\\
& \theta_{ip} \in \{0,1\}, && i \in N \text{ and } p \in \Omega_i && (4)
\end{aligned}
$$

This way aircraft are assigned to feasible routes while minimizing the total cost of serving all trips (1). Constraints (2) ensure that each trip is covered exactly once in the solution. Constraints (3) make sure every aircraft is assigned at most one route. Also, either a route is part of the solution or it is not, therefore parameters θ_{ip} must be either 0 or 1 (4).
To solve the set-partitioning model of the aircraft routing problem it is possible to use either an exact solution method or a heuristic (Hillier and Lieberman, 2008). As Hillier and Lieberman (2008) and Silver (2004) note heuristics are often used when the time required to find an optimal solution for an accurate model of the problem would be very large. For the aircraft routing system under consideration computational time is indeed an important requirement. However, the FMC studies mentioned in Table 2 that use exact solution methods to solve a set-partitioning model already achieve solution times in the order of seconds. The second performance measure for a solution method, apart from computational speed, is the degree of optimality of the solution (Silver, 2004). In practice, and even for exact solution methods, there is a maximum optimality gap given which is the maximum percentage that the obtained solution differs from the optimal solution. For the exact methods computational time or memory limits can cause the model not to be solved to optimality. The maximum optimality gap for the exact solution methods for the set-partitioning models has been found to be smaller than half a percent. Also, in the considered literature on the aircraft routing problem for on-demand air transport no heuristics have been used to solve the set-partitioning model. In addition, when designing a heuristic or adapting an existing one, both computational performance and solution quality can only be assessed after implementation. So when opting for a heuristic solution method for this problem no performance guarantees can be given beforehand. Because exact solution methods have proven to yield an adequate computational performance and solution quality for solving the set-partitioning model of the aircraft routing problem for on-demand air transport, the authors have chosen to use an exact solution method to solve the set-partitioning model that is created by the model creator.
To implement an exact solution method for the solver of the aircraft routing system, there are two options. First of all, one can program and implement an exact solution method, like branch-and-price or column generation, themselves. However, as Feillet (2010) has recently noted, carrying out an accurate implementation of an exact solution algorithm is a long and difficult task due to the inherent complexity of these methods and the lack of simple and comprehensive descriptions of these methods. The second method is to use a solver software package that embeds one or more of the exact solution methods needed to solve the set-partitioning problem. There already exist a number of solver packages that are applicable for the problem at hand, as can be seen in Table 2. Therefore, in this project a solver software package is used to solve the set-partitioning formulation of the PAAT aircraft routing problem. The solver package that is used is the IBM ILOG CPLEX Optimizer, since it has already proven its worth in ODAT applications. To solve the set-partitioning problem with CPLEX use is made of the built-in "mipopt" optimizer module. This optimizer uses a branch-and-cut algorithm, which is a hybrid method of branch-and-bound and cutting plane methods (Smith and Taşkin, 2008).

5. Designing the aircraft routing system

To construct the set-partitioning model of the aircraft routing problem for a PAAT operator, the model creator receives the following input data via the I/O interface of the aircraft

routing system: flight leg (trip) data of customer and dummy flight requests, parameters of the real and dummy aircraft, airport parameters, a list of persistent aircraft-trip assignments, and operational parameters. This is shown in Figure 2 below.

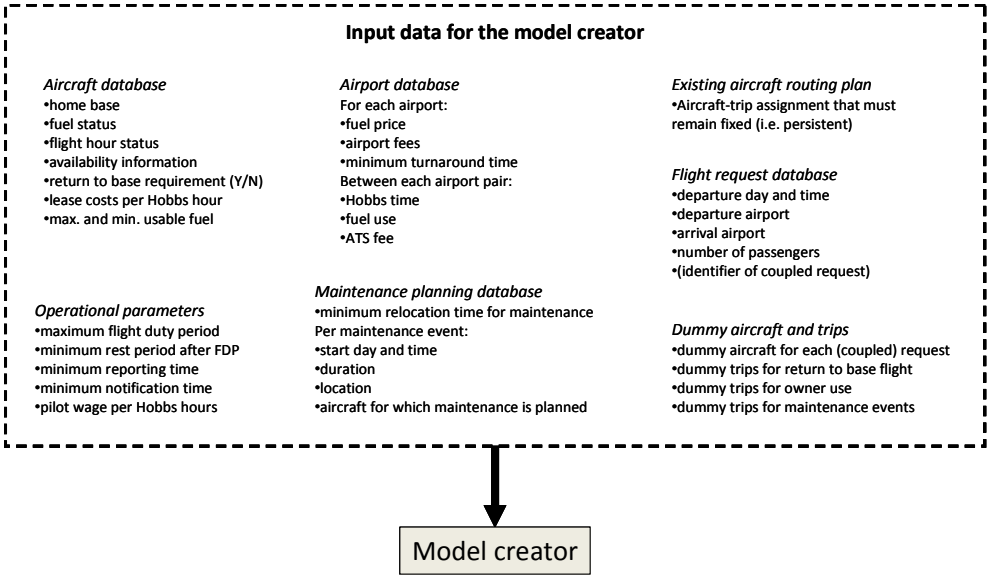


Fig. 2. Input data of the aircraft routing system.

Dummy requests are created for each return to base trip, for owner use and for maintenance events. Dummy aircraft on the other hand are added to represent the aircraft of another company when a trip is subcontracted. The model creator uses the above input data to generate the set of feasible routes for each aircraft (Ω_i), the total cost of operating each of these routes (c_{ip}) and information as to which trips are contained in each route (a^i_{jp}). With these parameters the set-partitioning formulation of PAAT aircraft routing problem is constructed and written to an LP file. This file is read by CPLEX which then solves the set-partitioning model. The obtained solution specifies which routes are part of the aircraft routing plan (i.e. which θ_{ip}'s are 1). This information is processed by the solver module which uses it to generate the aircraft routing plan that forms the solution of the aircraft routing problem. The model creator and the solver module of the aircraft routing system described in this chapter are programmed in MATLAB. The model creator uses the input data provided by the I/O interface to generate an LP file of the set-partitioning model that represents the aircraft routing problem. This LP file serves as input for the solver module. Figure 3 provides a detailed overview of this process.

First, the preprocessor module of the model creator uses the input data to create matrices in which it is specified whether aircraft i is initially compatible with trip j and whether aircraft i can fly trip j_2 after trip j_1. Together with the input data these two matrices are passed to the next module of the model creator, namely the route generator. The route generator module uses the information stored in the matrices to create the set of possible routes for each aircraft.

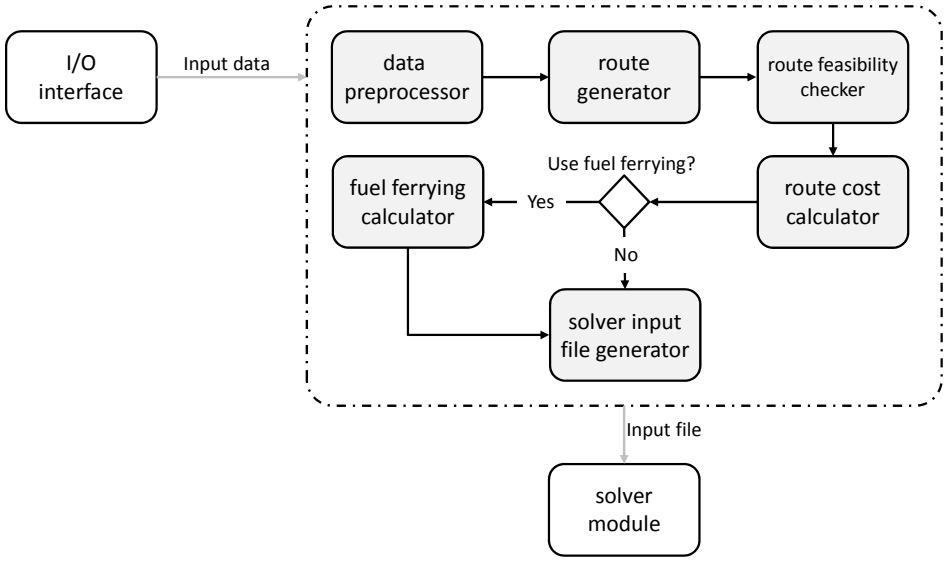


Fig. 3. The model creator module of the aircraft routing system.

Each of these routes consists of a sequence of flight legs. This set of routes for each aircraft is passed to the route feasibility checker. The route feasibility checker processes the route list for each aircraft generated by the route generator module and removes the routes that are infeasible. The list of feasible routes is then forwarded to the route cost calculator. The route cost calculator processes each of the feasible routes and calculates the total route cost for each of the feasible routes. The total route cost consists of the aforementioned variable aircraft cost, airport charges, ATS charges, fuel costs, and variable crew cost. If the fuel ferrying strategy is used, the fuel cost is calculated using the fuel ferrying calculator. This module uses the assumed percentage burn-off fuel ferrying calculation method. The final module of the model creator is the solver input file generator. This module writes an LP file of the set-partitioning model that is created using the previous modules. This LP file can be directly read and solved by CPLEX. The solver module of the aircraft routing system consists out of a MATLAB script that solves the created set-partitioning formulation, as written to the LP file, using CPLEX and retrieves the solution, i.e the aircraft routing plan.

6. Testing and validating the model

In this section the aircraft routing model is both tested and validated.

6.1 Testing the model

To test and validate the developed aircraft routing system operational data is needed as input. However, at the time of writing Fly Aeolus is not yet fully operational and as such the system can not be tested in the environment where it will eventually be used. Therefore a virtual test environment is created. This environment contains a virtual airport network in which a per-aircraft air taxi operates. The virtual environment is created by an input data generator which creates all the input data needed by the model creator. The input data generator utilizes operational parameters that reflect the (expected) PAAT operations of Fly Aeolus, see Table 3.

Parameter	Value	Unit
Parameters used to generate airport test data		
Farthest point-to-point distance in the network	800	km
Number of airports	225	-
Average cruise speed	310	km/h
Average fuel burn	1.12	litre per minute
ATS fee	0.00	€ per km
Passenger arrival fee	0 – 10	€ per pax
Passenger departure fee	0 – 30	€ per pax
Aircraft arrival fee	10 – 150	€
Aircraft departure fee	0 – 60	€
Turnaround time	45	minutes
Avgas price	0.90 – 2.70	€ per liter
Number of maintenance bases	6	-
Parameters used to generate trip test data		
Planning horizon	72	hours
Minimum Hobbs time of flight request	50	minutes
Average Hobbs time of flight request	90	minutes
Departure time distribution	-	xx
Average number of passengers	1.3	-
Percentage of coupled flight requests	10	% of flight requests
Number of trips in a coupled request	2	-
Waiting time between each trip of a coupled request	6 – 8	hours
Aircraft that have maintenance events	10	% of aircraft
Planned maintenance duration	6	hours
Minimum relocation for maintenance	120	minutes
Aircraft that have owner use	30	% of aircraft
Start time of owner use, between	5 – 20	hour
Aircraft that have return to base requirement	30	% of aircraft
Aircraft has to be back at base at	22	hour
Parameters used to generate aircraft test data		
Time when aircraft becomes available, between	0 – 17	hours of the 1st day
Initial fuel status, between	52 – 87	litres
Maximum useable fuel for 0, 1, and 2 pax	348; 272; 132	litres
Minimum reserve fuel	52	litres
Flight hours remaining till maintenance, between	6 – 50	flight hours
Lease costs per Hobbs hour, between	n/a	€
Charter cost per Hobbs hour, between	n/a	€
Parameters used to in the aircraft routing system		
Maximum flight duty period	13	hours
Minimum rest period after an FDP	12	hours
Time pilots need to report for duty	60	minutes
Pilot wage per Hobs hour	50	€ per hour
Surplus fuel burn (sfb_{hour})	4	% per hour
Allowance for unknown costs (a_u)	0.05	€ per litre
Number of closest routes to append (k)	10	-

Table 3. Parameters used to generate input test data

In addition, to test the use of persistent aircraft-trip assignments a demand uncertainty simulator is programmed. The input generator that is created to generate the test environment and corresponding data consists of two MATLAB scripts: an 'airport data generator' and a 'trip and aircraft data generator'. The airport data generator creates a virtual world containing a number of airports, i.e. the network in which a virtual per-aircraft air taxi service operates. The trip and aircraft data generator in turn creates flight requests that need to be served by that virtual operator and the data of the virtual operator's aircraft. Both the virtual world and the virtual PAAT operator are modelled on our Belgian air taxi company and its operations to ensure that the order of magnitude of the generated input data is the same as that of the environment where the system will eventually be employed.
The final parameter, *k*, determines the number of closest trips that are appended to a route during the route generation process. To choose the value for *k*, tests are conducted with the aircraft routing system with *k* ranging from 3 to 15. The results of these experiments are given in Tables 4 and 5. Note that each line of results that is reported in the results tables represents the average result of ten tests with independent random data sets. Furthermore the number of flight requests per day in the planning horizon is indicated in the results by *j* and the number of available aircraft by *i*. The data sets are generated using the virtual environment and virtual air taxi operator as described in this section.

1. k	2. (j, i)	3. Routes	4. ct_{MC}	5. ct_{CPLEX}	6. Solution cost
3	(7, 19)	1,729	5.3	0.01	18,649
5	(7, 19)	3,427	8.0	0.03	18,491
10	(7, 19)	8,090	15.6	0.08	18,401
15	(7, 19)	10,799	20.3	0.12	18,395
3	(15, 25)	6,086	17.7	0.13	38,469
5	(15, 25)	16,631	37.2	0.22	38,149
10	(15, 25)	72,260	177.9	1.21	37,849
15	(15, 25)	180,389	608.9	5.05	37,832

Table 4. The effect of *k* on the aircraft routing solution

The first column in Table 4 indicates the value of parameter *k* that is used for the experiments. Column 2 gives the number of flight requests per day that are generated and the number of aircraft in the fleet (excluding subcontractor aircraft). In the next column the total number of feasible routes present in the set-partitioning model is stated (thus for all the aircraft). Column 4 and 5 respectively denote the computational time in seconds of the model creator and CPLEX. Finally, in the last column the total cost of the aircraft routing plan is given. Table 5 below gives the results of the same experiments as shown in Table 4, but this time relative to the results for *k* = 15.

1. k	2. (j, i)	3. Routes	4. ct_{MC}	5. ct_{CPLEX}	6. Solution cost
3	(7, 19)	16%	27%	13%	101.4%
5	(7, 19)	33%	40%	33%	100.5%
10	(7, 19)	77%	78%	71%	100.0%
15	(7, 19)	100%	100%	100%	100.0%
3	(15, 25)	4%	3%	4%	101.8%
5	(15, 25)	10%	7%	5%	100.8%
10	(15, 25)	40%	30%	24%	100.0%
15	(15, 25)	100%	100%	100%	100.0%

Table 5. The effect of *k* on the aircraft routing solution (relative to *k* = 15)

It can be seen in column 3 of Table 4 that decreasing k causes the number of feasible routes that is generated by the model creator to decrease. This is because fewer of the closest trips are appended during the route generator loop. Due to the lower number of routes that are generated, both the model creator and CPLEX need less time to create and solve the set-partitioning model. This effect can be observed in column 4 and 5 of Table 5. While the number of created routes and the computational time for k = 3 can be as little as 4% of those for k = 15, the average increase in total cost of the aircraft routing solution is maximum 1.8%. When changing k from 15 to 10, the increase in solution cost is less than 0.1% while the computational time is decreased by at least 22% and in the (15, 25) case even by 70%. The latter can be explained by the fact that the number of routes that is created grows exponentially with the number of flight requests per day. A change in parameter k therefore has an increasingly larger effect on the change in the number of generated routes, and thus in computational time, when increasing the flight requests per day. In summary, varying parameter k from 15 to 3 has a large effect on the number of feasible routes that are created and, consequently, on the computational time needed by the aircraft routing system. However, the conducted experiments show that the effect on the solution cost is limited to an increase by maximum 1.8%. These findings are similar to the ones reported in the research of Ronen (2000) who varied k from 11 to 5 using 47 aircraft and 50 trips and reported a maximum cost increase of less than 0.1%. It can therefore be concluded that while having a high value of k causes the system to generate a larger number of feasible routes and thus possible solutions, the chosen aircraft routing solution mostly contains those routes that minimize the repositioning cost. In other words, when creating an aircraft routing it is a good rule of thumb to only consider those trips that are close to the arrival airport of the previous trip. As shown by the results in Table 5 only considering the five trips for which the departure airport is closest to the arrival airport of the previous trip, causes the routing solution to be at most 0.8% more costly compared to when considering the fifteen closest ones. Because changing k from 15 to 10 only increases the solution cost by less than 0.1% while reducing the solution time sometimes by as much as 70%, a k value of 10 is adopted as the standard value.

When creating the virtual airport network, demand and PAAT operator it is attempted to create input data for the aircraft routing system that has the same order of magnitude of the expected real input data. However, there are a number of differences between the virtual test environment and the air taxi operator expected operations. First of all, the virtual test environment uses a square airport grid in which airports are uniformly distributed instead of the actual locations of the airports in Western Europe. In addition, the airport charges the virtual environment uses are therefore not the ones encountered in reality. Also related to the airports is the fact that the turnaround time is taken to be 45 minutes for all airports. In reality the turnaround time depends on the airport.

The trip data generator randomly chooses a departure and arrival airport for all the trips. It therefore does not simulate demand hot spots, i.e. airports where a lot of demand originates or arrives. In reality, certain airports do attract significantly more business flight traffic than others. For example, Geneva Cointrin airport, which is the second busiest business airport in Europe, has almost 50% more business flight departures than the fifth busiest airport, Nice (Eurocontrol, 2010). In addition, the trip data generator only uses the hourly business flight departure pattern of France. In reality, however, the hourly departure pattern is dependent on the country the flight departs from and on the month and the day of the week (Eurocontrol, 2010). Also the trip data generator creates few day returns, i.e. customers that

fly from A to B in the morning and return from B to A in the evening. Furthermore, the data generator only earmarks aircraft unavailable for a maximum of a single day. However, it is possible that the aircraft is unavailable for a longer period of time. Yao et al. (2008) report that a FMC with 35 aircraft encountered 49 mid-day unscheduled maintenance events during a one-month period. It is therefore probable that unplanned maintenance will also occur during air taxi operator's operations. However, this virtual environment does not account for this.

It is apparent from the above that there exist differences between the virtual test environment that is created and the expected operational environment of the air taxi operator. However, as explained before, the virtual test environment is only used in this project to test and validate the workings of the aircraft routing system, and not for example to find out what the exact effect of the fleet size is on the air taxi operations. Therefore, it is of main importance that all the necessary input elements (like airport turnaround times, owner use trips and coupled requests) are specified in the virtual environment and that these have the same order of magnitude as in reality. As such it can be assessed whether the aircraft routing system is capable of handling all these inputs correctly. It is in this stage not important that the virtual test environment does exactly represent reality. The main aim is the development of the aircraft routing system and therefore the virtual test environment described in this section is adequate for its purpose, namely validating the aircraft routing system.

6.2 Validating the model

Validation is an important step in model development because through validation the level of confidence in and the credibility of the aircraft routing system is established. To carry out the validation in this project five types of validation are conducted, namely conceptual model validation, computerized model verification, data validation, experimental validation and operational validation (Landry et al., 1983; Sargent, 2007).

In the **conceptual validation** phase, in which it is determined whether the underlying theories and assumptions of the system are correct, face validation and desk checking are the used validation techniques. It is argued that the theory of the aircraft routing system is indeed correct. However, there are some assumptions of which the exact effect on the accuracy of the conceptual model could not be determined as no real operational data is available, namely the assumptions that a fixed percentage of ferried fuel is burned per hour, that the Hobbs time and fuel use between airport pairs is specified for the one passenger case, and that it is acceptable to impose crew constraints on the aircraft. It is therefore advised that these assumptions and their effect are checked as soon as real operational data is available. Next is the **computerized model verification**. This process ensures that the implementation of the conceptual model in programming code is correct. To make sure this holds true for the MATLAB implementation of the aircraft routing system the authors used bottom-up development and testing, debugging and desk checking. In addition, experiments show that the results produced by the routing system are consistent. Next, **data validation** is performed to ensure that the data used to build and evaluate the model is adequate and correct. For building the model the mental and written databases provided by the air taxi operator and found in ODAT literature are considered appropriate. In addition, the virtual test environment and corresponding numerical input data created to support the evaluation of the aircraft routing system is

deemed adequate for its purpose. **Experimental validation** aims to check whether or not the model contradicts qualitative, expert knowledge. To this end both the convergent validation and face validation techniques are utilized. For the first technique the results produced by the aircraft routing system are checked against six expert statements. In all cases the results of the routing system adhered to the expert knowledge. In addition face validation is carried out with Fly Aeolus' director. He is given an example aircraft routing problem to assess whether the solution produced by the routing system is correct, which he indeed found to be true. Finally, **operational validation** assesses the usefulness and timeliness of the solutions and cost of implementing the system. To carry out the routing plans produced by the aircraft routing system the air taxi operator must first decide on and then apply the waiting strategy for the routing plans. Furthermore, as mentioned in the conceptual validation, it must be assessed what the exact effect on the accuracy of the routing plans is of the fuel ferrying, and Hobbs time and fuel use assumptions before implementing the routing system. With respect to the timeliness it is found that the aircraft routing system is capable of solving routing problems with sizes that correspond to air taxi operator's first operational years in under five minutes. Noteworthy in this respect is that the routing system is capable of producing a solution to a routing problem in little over one second that takes a human dispatcher 30 minutes to solve.

7. Discussion & conclusions

In this section important limitations of the work are discussed and conclusions are drawn.

7.1 Discussion

As this aircraft routing system has been developed for an air taxi operator who – at time of writing – is not yet fully operational, no real operational input data can be used for system validation and testing and instead a virtual test environment that approximates the air taxi operator's operations is utilized. Therefore, this research has some limitations.

The air taxi operator needs to check some assumptions underlying the routing system before putting it to use. The exact effect of these assumptions on the routing plans could not be assessed in this project because no real operational data is available. First, it needs to be assessed whether the increase in routing costs caused by imposing crew constraints upon the aircraft is acceptable. In addition, the air taxi operator must check whether specifying Hobbs time and fuel use between airport pairs for the one passenger load case is accurate enough. Finally, the relation to reality of the assumption that a fixed percentage of ferried fuel is burned per hour must be checked. Once operational data becomes available, the air taxi operator is advised to check these assumptions. This is deemed a necessary step in the implementation of the aircraft routing system. However do note that with respect to the Hobbs time and fuel use assumption, the air taxi operator can also opt to remove this assumption from the system altogether by specifying these input values for each passenger load case.

The aircraft routing plans created with the aircraft routing system do not contain a waiting strategy for repositioning flights, only the earliest and latest start time for deadheads are indicated. If the air taxi operator wants to adopt a 'fly-first' or 'wait-first' waiting strategy, this can be directly applied to the created routing plans and as such this limitation is easily removed from the system. However, the authors suggest that the aircraft routing system is

augmented with a module which calculates the optimal waiting strategy per aircraft route instead of adopting the 'fly-first' or 'wait-first' approach as the latter two are suboptimal approaches. Once the waiting strategy is determined, also aircraft parking fees, crew waiting wages, and airport opening hours can be incorporated into the aircraft routing system, as these all depend on the strategy that is adopted.
With respect to the secondary planning objectives, the degree of robustness, flexibility and persistence of the aircraft routing plans are not measured. The developed system does allow the OCC to control these parameters, but only to a limited extent. Robustness is solely controlled by slack time, and no other means of enforcing robustness are currently provided. The same holds for persistence and flexibility, whose balance is only altered by using the persistent aircraft- trip assignments. But the aircraft routing system does not provide information as to exactly how robust, flexible or persistent the created routing plans are. Though the dispatcher knows he is increasing/decreasing robustness, flexibility or persistence when he is adapting the corresponding control parameters. The authors would like to further explore and implement other control parameters to enforce these three secondary planning objectives of an aircraft routing plan in an on-demand air transport (ODAT) context.

7.2 Conclusions

This paper has aimed to be the first study that provides a detailed description of the aircraft routing problem as faced by a per-aircraft air taxi (PAAT) operator and also the first study that treats the development of an aircraft routing system for a per-aircraft air taxi operator. The developed aircraft routing system, which consists out of the model creator, CPLEX and the solution module, is capable of solving the aircraft routing problem faced by the air taxi operator in under five minutes for the first operational years. This will allow the air taxi operator to achieve an estimated cost reduction of 12% on their routing plans with respect to using a human dispatcher. This amounts to estimated savings of 10 % on variable direct operating costs during the air taxi's first operational year. By altering the slack time parameter the OCC can control the robustness, i.e. the capability to cope with and recover from external disruptions, of the aircraft routing plans it creates. Furthermore, via the persistent aircraft-trip assignments input, the system allows the OCC to control the balance between persistence, the degree to which the aircraft routing plan deviates from the previous one when it is re-optimized, and flexibility, the ability of the schedule to cost-effectively deal with future demand, of the generated aircraft routing plans. In addition, by setting aircraft-trip assignments as input the OCC can specify 'hard' input decisions. The aircraft routing system also allows the air taxi operator to incorporate fuel ferrying in the aircraft routing phase as a strategy to reduce the operational cost of the aircraft routing plan. The next stage for our project is for our Belgian air taxi operator to become fully operational, further develop the routing system and eventually implement it at their OCC. Then the various databases in the aircraft routing system can be filled with real life operational data and help the air taxi operator to provide affordable and reliable full on-demand personal air transport.

8. References

Abdelghany, A. & Abdelghany, K. (2009). *Modeling Applications in the Airline Industry*, Ashgate, ISBN 978-0-7546-7874-8, Farnham, United Kingdom

American Express (2009). 2010 Global business travel forecast. In: *American Express*, 02.11.2011, Available at: http://corp.americanexpress.com/gcs/travel/news/ pressreleases.aspx

Ball, M., Barnhart, C., Nemhauser, G. & Odoni, A. (2007). Chapter 1 Air Transportation: Irregular operations and control, In: *Transportation. Vol. 14 of Handbooks in Operations Research and Management Science,* Barnhart, C. & Laporte, G. (Eds.), pp. 1 – 67, Elsevier, ISBN: 978-0-444-51346-5, Amsterdam, the Netherlands

Barnhart, C., Belobaba, P. & Odoni, A. (2003). Applications of operations research in the air transport industry, *Transportation Science*, Vol. 37, No. 4, pp. 368 – 391, ISSN: 0041-1655

Bazargan, M. (2004). *Airline Operations And Scheduling*, Ashgate Publishing, ISBN 978-0-7546-9772-5, Aldershot, United Kingdom

Bian, F., Burke, E., Jain, S., Kendall, G., Koole, G., Landa Silva, J., Mulder, J., Paelinck, M., Reeves, C., Rusdi, I., Suleman, M. (2003). Making airline schedules more robust. In: *Proceedings of the 1st Multidisciplinary International Conference on Scheduling: Theory and Applications (MISTA 2003)*, pp. 678–693

Bonnefoy, P. (2005). Simulating air taxi networks. *Proceedings of 37th Winter Simulation Conference*, pp. 1586 – 1595, ISBN: 0-7803-9519-0

Budd, L. & Graham, B. (2009). Unintended trajectories: liberalization and the geographies of private business flight. *Journal of Transport Geography*, Vol. 17, No. 4 (July 2009), pp. 285 – 292, ISSN: 0966-6923

Bunte, S. & Kliewer, N. (2010). An overview on vehicle scheduling models. *Public Transport*, Vol. 1, No. 4, (March 2010), pp. 299 – 317

Cirrus (2007). Cirrus SR22 weight and balance loading form, In: *Airplane information manual for the Cirrus Design SR Cirrus, 04.05.2011, Available on:* http://servicecenters.cirrusdesign.com/TechPubs/pdf/POH/sr22/pdf/20880-001InfoManual.pdf

Doganis, R. (2002). *Flying Off Course*, 3rd Edition, Routledge, ISBN: 0-415-21323-1, New York

Dyson, E. (2006). Visible demand: The new air-taxi market, In: *Fly Your Jet website*, 17.11.2010, Available at: http://www.flyyourjet.com/Portals/0/DysonRelease.pdf

Espinoza, D., Garcia, R., Goycoolea, M., Nemhauser, G. & Savelsbergh, M. (2008a). Per-seat, on-demand air transportation part i: problem description and an integer multicommodity ow model. *Transportation Science*, Vol. 42, No. 3, (August 2008), pp. 263 – 278, ISSN: 0041-1655

Espinoza, D., Garcia, R., Goycoolea, M., Nemhauser, G. & Savelsbergh, M. (2008b). Per-seat, on-demand air transportation part ii: Parallel local search. *Transportation Science*, Vol. 42, No. 3, (August 2008), pp. 279 – 291, ISSN: 0041-1655

Eurocontrol (2010). Business aviation in Europe 2009, In: *Eurocontrol website*, 04.05.2010, Available at: http://www.eurocontrol.int/statfor/public/subsitehomepage/ homepage.html

Fagerholt, K., Foss, B. A. & Horgen, O. J. (2009). A decision support model for establishing an air taxi service: A case study. *Journal of the Operational Research Society*, Vol. 60, No. 9, (September 2009), pp. 1173 – 1182, ISSN: 0160-5682

Feillet, D. (2010). A tutorial on column generation and branch-and-price for vehicle routing problems. *4OR*, A Quarterly Journal of Operations Research, Vol. 8, No. 4, (June 2010), pp. 407 – 424, ISSN: 1619-4500

Fly Aeolus (2010). Private Jet Price Private Plane Charter Prices, In: *Fly Aeolus website*, 03.05.2011, Available at http://www.flyaeolus.com/affordable-and-low-price/

Flyer Forums (2010). Avgas price table, In: *Flyer Forums website*, 25.10.2010, Available at: http://forums.yer.co.uk/viewtopic.php?t=51535

Gall, C. & Hindhaugh, J. (2009). Communicating change to sr. management economic & industry outlook 2010, In: *National Business Travel Association website*, 12.07.2010, Available at: http://www2.nbta.org/Lists/Resource%20Library/F103%20Communicating%20Change%20-%20Industry%20Outlook.pdf

Hicks, R., Madrid, R., Milligan, C., Pruneau, R., Kanaley, M., Dumas, Y., Lacroix, B., Desrosiers, J. & Soumis, F. (2005). Bombardier flexjet significantly improves its fractional aircraft ownership operations. *Interfaces*, Vol. 35, No. 1, (January – February 2005), pp. 49 – 60, ISSN: 0092-2102

Hillier, F. S. & Lieberman, G. (2008). *Introduction to Operations Research*. 8th Edition, McGraw Hill, ISBN: 007-123828-X, New York

Karaesmen, I., Keskinocak, P., Tayur, S. & Yang, W. (2005). Scheduling multiple types of time-shared aircraft: Models and methods for practice. In: *Proceedings of the 2nd Multidisciplinary International Conference on Scheduling: Theory and Applications*. Kendall, G., Lei, L. & Pinedo, M. (Eds.), New York (July 2005), Vol. 1, pp. 19 – 38

Karaesmen, I., Yang, W. & Keskinocak, P. (2007). Optimization vs. persistence in scheduling: heuristics to manage uncertainty in on-demand air travel. In: *Proceedings of the 3rd Multidisciplinary International Conference on Scheduling: Theory and Applications*, Paris (August 2007), Vol. 1., pp. 242 - 250

Keskinocak, P. & Tayur, S. (1998). Scheduling of time-shared jet aircraft, *Transportation Science*, Vol. 32, No. 3, pp. 277 – 294, ISSN: 1526-5447

Keysan, G., Nemhauser, G. & Savelsbergh, M. (2010). Tactical and operational planning of scheduled maintenance for per-seat, on-demand air transportation. *Transportation Science*, Vol. 44, No. 3, (August 2010), pp. 291 – 306, ISSN: 0041-1655

Landry, M. & Muhittin, O. (1993). In search of a valid view of model validation for operations research. *European Journal of Operational Research*, Vol. 66, No. 2, (April 1993), pp. 161 – 167, ISSN 0377-2217

London Executive Aviation (2010). In: *London Executive Aviation website*, 23.11.2010, Available at: http://www.flylea.com/

Mane, M. & Crossley, W. A. (2007a). An approach to predict impact of demand acceptance on air taxi operations. In: *7th AIAA Aviation Technology, Integration, and Operations Conference*. Vol. 1, Belfast, Northern Ireland, pp. 954 – 961

Mane, M., & Crossley, W. A. (2007b). Probabilistic approach for selection of maintenance facilities for on-demand operations. In: *7th AIAA Aviation Technology, Integration, and Operations Conference*. Vol. 1. Belfast, Northern Ireland, pp. 944 – 953

Mane, M. & Crossley, W. A. (2009). Importance of aircraft type and operational factors for air taxi cost feasibility. *Journal of Aircraft*, Vol. 46, No. 4, pp. 1222 – 1230, ISSN: 0021-8669

Martin, C., Jones, D. & Keskinocak, P. (2003). Optimizing on-demand aircraft schedules for fractional aircraft operators. *Interfaces*, Vol. 33, No. 5, (September – October 2003), pp. 22 – 35, ISSN: 0092-2102

NBAA (2010), Piston Engine Aircraft, In: *National Business Aviation Association website*, 03.05.2011, Available at: http://www.nbaa.org/businessaviation/aircraft/pistons/

Ronen, D. (2000). Scheduling charter aircraft. *The Journal of the Operational Research Society*, Vol. 51, No. 3, (March 2000), pp. 258 – 262, ISSN: 0160-5682

Sargent, R. (2007). Verification and validation of simulation models. *WSC '07 Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come*, IEEE Press Piscataway, NJ, pp. 124 – 137, ISBN: 1-4244-1306-0

Silver, E.A. (2004). An overview of heuristic solution methods. *The Journal of the Operational Research Society*, Vol. 55, No. 9, (May 2004), pp. 936 – 956, ISSN: 0160-5682

Smith, J. C. & Taşkin, Z. C. (2008). A tutorial guide to mixed-integer programming models and solution techniques. In: *Optimization in Medicine and Biology*. Lim, G. J., Lee, E. K. (Eds.), pp. 521 – 548, Taylor and Francis, Auerbach Publications, ISBN: 978-0-849-305-634. New York

Yang, W., Karaesmen, I., Keskinocak, P. & Tayur, S. (2008). Aircraft and crew scheduling for fractional ownership programs. *Annals of Operations Research,* Vol 159, (December 2007), pp. 415 – 431, ISSN: 1572-9338

Yang, W., Karaesmen, I. Z. & Keskinocak, P. (2010). Managing uncertainty in on-demand air travel. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 46, No. 6, (November 2010) p. 1169 – 1179, ISSN: 1366-5545

Yao, Y., Zhao, W., Ergun, O. & Johnson, E. (2005). Crew Pairing and Aircraft Routing for On-Demand Aviation with Time Window. In: *Social Sciences Research Network eLibrary*., 04.05.2011, Available from: http://ssrn.com/abstract=822265

Yao, Y. & Zhao, W. (2006). Optimization problem in fractional ownership operations. *Collection of Technical Papers - 6th AIAA Aviation Technology, Integration, and Operations Conference*, Vol. 1, pp. 319 – 323

Yao, Y. (2007). *Topics in fractional airlines*. Georgia Institute of Technology, School of Industrial and Systems Engineering, PhD thesis, May 2007

Yao, Y., Ergun, O., Johnson, E., (2007). Integrated model for the dynamic on-demand air transportation operations. In: *Dynamic Fleet Management*. Sharda, R., Vo, S., Zeimpekis, V., Tarantilis, C. D., Giaglis, G. M. &Minis, I. (Eds.), Vol. 38 of Operations Research/Computer Science Interfaces Series, pp. 95 – 111, Springer US, New York, ISBN 978-0-387-71721-0.

Yao, Y., Ergun, O., Johnson, E., Schultz, W. & Singleton, J. M. (2008). Strategic planning in fractional aircraft ownership programs. *European Journal of Operational Research*, Vol. 189, No 2, (September 2008), pp. 526 – 539, ISSN: 0377-2217