UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA IDRAULICA, MARITTIMA, AMBIENTALE E GEOTECNICA

SCUOLA DI DOTTORATO DI RICERCA

IN SCIENZE DELL'INGEGNERIA CIVILE ED AMBIENTALE

INDIRIZZO: IDRODINAMICA E MODELLISTICA AMBIENTALE

CICLO XX

INVESTIGATIONS ON VITREOUS HUMOUR DYNAMICS

Direttore della Scuola : Ch.mo Prof. Stefano Lanzoni

Relatori : Ch.mo Prof. Marco Colombini

Ch.mo Prof. Alessandro Stocchino

Controrelatore :Ch.mo Prof. Michael Dumbser

Dottoranda :Chiara Cafferata

UNIVERSITA' DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Sede Consorziata: Università degli studi Genova

Dipartimento di Ingegneria Idraulica, Marittima, Ambientale e Geotecnica

SCUOLA DI DOTTORATO DI RICERCA

IN SCIENZE DELL'INGEGNERIA CIVILE ED AMBIENTALE

INDIRIZZO: IDRODINAMICA E MODELLISTICA AMBIENTALE

CICLO XX

INVESTIGATIONS ON VITREOUS HUMOUR DYNAMICS

Direttore della Scuola : Ch.mo Prof. Stefano Lanzoni

Supervisore : Ch.mo Prof. Marco Colombini

Ch.mo Prof. Alessandro Stocchino

Controrelatore : Ch.mo Prof. Michael Dumbser

Dottoranda :Chiara Cafferata

DATA CONSEGNA TESI
30 Novembre 2009

Contents

List of Figures

Abstract

The vitreous humour is the substance that fills the posterior chamber. The vitreous body ensures the adherence of the retina to the underlying layers and acts as a barrier for heat and drug transport between the anterior and the posterior segments of the eye. In the medical literature many authors have postulated the existence of a connection between the fluid dynamics within the vitreous chamber during eye movements (especially when the substance filling the chamber is a liquid) and the occurrence of the retinal detachment.

In the present PhD Thesis the vitreous humour dynamics is studied by means of both an experimental and a numerical point of view.

The experiments consist in the measurement of the velocity field on the equatorial plane of a magnified model of the vitreous chamber using PIV techniques. The vitreous chamber has been modeled by means of a cavity carved in a rigid Perspex container. Two different geometries were used: a spherical geometry and a deformed geometry, which closely resembles that of the real eye. The model is filled with Glycerol, a highly viscous Newtonian fluid, and it is mounted on the shaft of a computer controlled motor, which rotates according to a generic time law. Visualizations of the fully three-dimensional flow show that the primary flow occurs on planes perpendicular to the axis of rotation. Secondary flows can be detected, the magnitude of which is however three or four orders smaller than the primary flow. In the case of spherical geometry, theoretical results, based on a simplified solution, are shown to be in very good agreement with the experimental findings. The maximum value of the shear stress at the wall does not significantly depend on the amplitude of saccadic movements but is strongly influenced by the vitreous viscosity. Velocity fields are found to be strongly influenced by the deformed geometry of the domain. When using the deformed model the formation of a vortex in the vicinity of the lens is invariably observed. The path described by this vortex during a period of oscillation is found to depend on the Womersley number of the flow.

The second part of the Thesis is devoted to the formulation of a numerical model which provides a direct numerical simulation of the flow of a viscous fluid inside an oscillating sphere. The problem is formulated in the primi-

tive velocity-pressure form. Replacing the continuity equation by the Poisson equation for the pressure supported by an independent condition of integral character an uncoupled formulation of the problem is derived. The equations are discretized in time by means of a non-fractional-step scheme. A spectral technique is used to transform the original problem in a sequence of ordinary differential problems for the coefficients of a spherical harmonics expansion of the variables. Spectral techniques have been tested on a simplified problem providing encouraging results.

Sommario

L'umor vitreo è la sostanza che riempie la camera posteriore dell'occhio. Esso assicura l'aderenza della retina agli strati sottostanti e costituisce una barriera per il trasporto di calore o sostanze tra la camera anteriore e la camera posteriore dell'occhio. Molti autori nella letteratura medica hanno ipotizzato una connessione tra la dinamica del vitreo durante i movimenti oculari (specialmente nei casi in cui la sostanza che riempie la camera è un liquido) e l'insorgenza del distacco retinico.

Nella presente Tesi di Dottorato lo studio della dinamica del vitreo oculare è stato affrontato secondo un approccio sperimentale e numerico.

Gli esperimenti hanno riguardato la misura dei campi di moto che si realizzano sul piano equatoriale di un modello in scala amplificata della camera vitrale attraverso l'utilizzo della tecnica PIV. La camera vitreale è stata modellata per mezzo di una cavità ricavata in un contenitore rigido di perspex. I contenitori usati hanno due diverse geometrie: una sferica e una sferica deformata che riproduce verosimilmente la reale conformazione della camera. Il modello è stato riempito con glicerina, un fluido newtoniano ad elevata viscosità, e montato sull'albero di un motore che è stato fatto ruotare secondo una generica legge di tipo.

Attraverso visualizzazioni del moto tridimensionale si è osservato come il moto principale si sviluppi su piani ortogonali all'asse di rotazione. Sono stati anche osservati moti secondari la cui intensità risulta essere di tre-quattro ordini di grandezza inferiori al moto principale. Nel caso di modello sferico i risultati sperimentali sono stati confrontati con una teoria semplificata mostrando un buon accordo con quest'ultima. Il massimo valore della tensione tangenziale è risultato non dipendere in modo significativo dall'ampiezza del movimento ma di essere influenzato dalla viscosità del fluido interno alla camera. I campi di moto sono fortemente influenzati dalla geometria del dominio. Usando il contenitore deformato in tutti gli esperimenti è stata osservata la formazione di un vortice. Il percorso descritto dal vortice durante un periodo di oscillazione è risultato dipendere dal numero di Womersley del moto.

La seconda parte della Tesi è dedicata alla formulazione di un modello nu-

merico consistente in una simulazione numerica diretta del moto di un fluido viscoso all'interno di una sfera oscillante. Il problema è formulato nelle variabili primitive pressione-velocità. Sostituendo l'equazione di continuità con l'equazione di Poisson per la pressione, associata ad una condizione al contorno indipendente di tipo integrale, si è ricavata una formulazione disaccoppiata dl problema. Le equazioni sono state discretizzate nel tempo per mezzo di uno schema non frazionario. Il problema originale è stato trasformato attraverso una tecnica spettrale in una sequenza di problemi alle derivate ordinarie per i coefficienti di espansione in armoniche sferiche delle variabili del problema. Le tecniche spettrali sono state testate per la soluzione di un problema semplificato fornendo risultati confortanti.

Preface

The anatomical shape of the human eyes is approximately spherical. They consist of two main chambers: the anterior and the posterior chamber separated by the lens and enclosed by the eye wall. The anterior chamber is filled with aqueous humour a physiological liquid whose mechanical characteristics are similar to those of water. In the posterior chamber a substance called vitreous humour is located. The vitreous humour is an homogeneous gel but ageing can produce significant mechanical changes in vitreous consistency, leading to a partial liquefaction of the vitreous. This process may involve the entire posterior chamber and frequent occurs in eyes that are nearsighted (myopia). Vitreous cavity can be filled with a liquid also in vitrectomy surgery, when the surgeon replace a cloudly vitreous with a salt solution or a silicon oil. Most of the eye globe volume is occupied by fluids, therefore many of the problems related with the eye's functioning are of a strictly fluid dynamic nature. Aqueous humour circulation in the anterior chamber has been studied by several authors. Vitreous dynamics has not been extensively studied yet, even though many authors have postulated a connection between the actions exerted by vitreous motion during eyes movements and the pathogenesis of retinal detachment. On the other hand the vitreous motion may influence heat and drug transport between the anterior and the posterior segments of the eye. Quite a few papers in the medical literature have faced the problem of analyzing the transport phenomena within the vitreous cavity.

In the first part of the present work we present the results of an experimental model of vitreous motion, firstly within a spherical cavity and then in a deformed sphere chamber that closely resembles the real vitreous cavity shape. The experimental conditions reproduce the case of liquefied vitreous or tamponade fluid used in some surgical treatments (e.g. vitrectomy). In these cases hydrodynamic events within the vitreous cavity are expected to be the most intense. A Newtonian fluid is, in fact employed, thus the elastic properties of the vitreous are disregarded at this stage. Visualizations of the whole flow field were made in order to prove some theoretical assumptions for the model. Measures of velocity fields on the equatorial plane were made by

means of PIV technique to investigate the vitreous fluid dynamic and derive the shear stress on wall during eye movements.

With the P.I.V. experimental set-up at our disposal we were able to reconstruct the vitreous motion on the equatorial plane of the model, but we cannot quantify the fully 3D velocity field inside the entire chamber (only qualitative visualizations were made experimentally). In order to exceed the limits of the experimental model we use a numerical model.

In the second part of the present work we describe the some preliminary results of numerical model which solves, with a spectral technique, the 3D problem of motion inside a rotating spherical crown. The numerical model solves the Navier Stokes equations in the primitive formulation form, where the incompressibility constraint is replaced by the Poisson equation. A boundary condition of integral character are derived for the pressure, obtaining a system of split-equations for the pressure and velocity to be solved independently and in sequence, after having solved the metharmonic problem providing the metharmonic function to be inserted in the integral condition for the pressure. The shape of the investigated domain suggest the expansion of the unknown in series of spherical harmonics. The original time dependent problem reduce to a sequence of stationary problem for the coefficients of the expansion that are function only of the radial coordinate. The problem is solved by means of a tau projection spectral method after having used a further expansion of the coefficients in terms of Chebyshev polynomials.

Part I

An experimental model

Introduction to part I

The largest structure of the eye is the posterior chamber the shape of which is approximately spherical. The vitreous humour is the substance that fills the posterior chamber. The vitreous body gives the eye its shape. Being in contact with the retina it helps to keep it in place by pressing it against the under layers. Moreover, it acts as a barrier between the anterior and the posterior segments of the eye for heat and drug transport [14]. In normal condition the vitreous is a gel, Lee et al.(1992)[12] have studied the rheology of vitreous, showing that it behaves like a viscoelastic material. However, with advancing age the vitreous partially or completely loses its elastic properties, as a consequence of a liquefaction process. In the medical literature many authors have supposed a connection between the fluid dynamic within the vitreous chamber and the retinal detachment (RD). In fact, during the rapid rotations of the eyes, the actions exerted by the relative motion of the eye wall respect to the substance filling the posterior chamber can be quite intense and constitute an important factor in inducing the rhegmatogenous retinal detachment. This pathology occurs when the fluid infiltrates into the subretinal space through a hole, tear, or break in the retina raising it from the underlying layers. Myopia is a risk factor for the RD, especially when for aging or pathologies the vitreous is liquified or replaced by salt solution or a silicon oil during a surgery called vitrectomy. The vitreous motion may also influence heat and drug transport between the anterior and the posterior segments of the eye. This phenomenon is poorly studied and all authors have always interpreted it in terms of a diffusion process disregarding the convective transport due to eye motion. Medical observations show that if vitreous mixing is intense enough, advection may be by far a more efficient process.

The above mentioned reasons have motivated the research on the dynamics of the vitreous motion induced by eye rotations. Regarding to the vitreous actions on the retina Lindner (1933) [13] performed hydrodynamic experiments aimed at understanding the structure of the flow field induced by eye rotations within the vitreous chamber, once a retinal tear has occurred. The above experiments were later reconsidered and extended by Rosengren and

Östrelin(1976) [21]. Both authors point out the importance of eye rotations on the vitreous dynamics. However, their works only provide a qualitative picture of the hydrodynamic events within the vitreous chamber because their experimental approach did not permit quantitative measurements of the fluid velocity to be performed. They also underline the importance of the lens, the presence of which confers an irregular shape to the chamber and hence may significantly modify flow field characteristics. Nevertheless, both authors used a cylindrical glass container, filled with water, with a thin transparent membrane covering the inner wall with the aim to model the vitreous chamber, disregarding the real shape of the posterior cavity and the scale effects arising from the magnified size of the model. Greater insight into real geometry configuration in vitreous dynamic is provided by the work of David et al (1998)[7]. The authors studied analytically the motion of a viscoelastic fluid within a periodically rotating sphere, adopting simplifying assumptions on the flow characteristics. Their results clarify most of the features of the flow field, showing, in particular, that the fluid motion in the inner part of the domain is out of phase with respect to the wall motion. Moreover, David et al (1998)computed the shear stress at the wall and showed that it increases with the sphere radius. The latter observation was interpreted by the authors as a possible explanation of the more frequent occurrence of retinal detachments in myopic eyes, typically characterized by a larger size. Recently Repetto et al.(2004)[19] have proposed an analytical model attempting to interpret one of the possible mechanisms whereby tensile stresses may be generated on the retina, possibly leading to a retinal break. The authors have studied the liquefied vitreous' dynamics in the presence of vitreous membranes. Their results suggest that vitreous membrane oscillations may be resonantly excited by small-amplitude, high-frequency eye movements, called 'microsaccades', which may thus induce high stresses on the retina and eventually cause its detachment.

In the present Thesis, we present the results of a series of experiments on vitreous motion. The vitreous chamber has firstly been modelled as a spherical cavity, magnified with respect to the real geometry, carved within a perspex cylinder and able to rotate according to a prescribed time law. The cavity has been filled with glycerol, which is a high viscosity Newtonian fluid; thus, the elastic properties of physiological vitreous are disregarded in this first experimental approach to the problem. However, David et al (1998)[7] reported that the elastic component plays a fairly minor role on the flow field within the eye globe with respect to the viscous one. Moreover, according to the authors' results, the elastic component of the vitreous behaviour does not influence the maximum shear stress at the wall, which is the most important quantity to be evaluated in order to understand the connection between vitreous motion and

retinal detachment. The flow field has been experimentally measured, on the equatorial plane orthogonal to the axis of rotation, through the PIV technique. This allowed us to obtain the spatial structure of the flow field in such a plane with great detail. Moreover, the time evolution of the flow can also be suitably described. Two sets of experiments have been carried out. The first one is aimed at verifying the results of David et al (1998)[7] for purely viscous fluids; therefore, the eye model is rotated according to a sinusoidal time law. Real saccades have a much more complex time behaviour, though, and can hardly be described in terms of a simple sine function. Therefore, in a second set of experiments it has been adopted, a time law for the container angular velocity that correctly reproduces the main features of real large amplitude saccadic movements. Saccades with different amplitudes (from 10° to 50°) have been studied. The second set of experiments provides important insight both into the structure of the flow field occurring within the eye and into the shear stresses exerted by the vitreous on the retina during real saccadic movements.

As a matter of fact, however, the vitreous chamber is not spherical but rather concave in the anterior part, due to the presence of the lens. Even a weak departure from the spherical shape may induce strong changes in the flow field and, in turn, it may significantly affect the shear stress distribution at the wall; this evidence was already underlined by Lindner (1933)[13] and Rosengren and Östrelin(1976) [21] studies. The only attempt to account for this effect is due to Repetto (2006)[18]. The latter analysis is carried out analytically in the limit of low viscosity fluid and the vitreous cavity is described as a weakly deformed sphere. The author's results show that the irregular shape of the domain produces a complex and strongly three-dimensional flow field, which is responsible of an intense mixing. This result shows that the motion of the vitreous induced by eye rotation is responsible for a convective transport and, if vitreous mixing is intense enough, advection may be by far a more efficient process than the diffusive one. The stress distribution at the wall, which is not computed in the author's model since the flow within the boundary layer at the wall is not accounted for, is also expected to be strongly affected by the shape of the vitreous chamber.

With regard to the vitreous motion effect on heat and drug transport between the anterior and the posterior segments of the eye the reader is referred to H.Lund-Andersen (2003)[14] for a thorough review. As far as the transport of solutes is concerned in vivo measurements are possible by using fluorescein as a tracer and performing vitreous fluorophotometry. Such measurements are invariably interpreted in terms of a diffusion process, with the aim of determining an overall diffusion coefficient. The diffusive process of heat transfer within the vitreous body has been studied numerically by many authors [22] ,[6]e.g..

All the existing works consider the transport processes within the vitreous as being purely diffusive. Neverthless, the convective transport may be important in transport of solutes: this idea is somehow supported by the clinical observation that, in normal conditions, the vitreous is much more efficient as a transport barrier than in the case of liquefied vitreous or in vitrectomised eyes and, in both the aforementioned circumstances, the fluid dynamic processes in the vitreous cavity are much more intense even though admittedly, however the diffusive coefficient of the healthy vitreous may also be different from the one of the liquefied vitreous or of tamponade fluids.

The second kind of experiments related in the present contribution were made with a model which closely resembles the real shape of vitreous chamber, showing an indentation which simulate the presence of the lens. The cavity is filled with the same fluid used in spherical model experiments and measurements of quantitative flow fields on the equatorial plane of the model were made with PIV technique. The result of this second model experiments were published in 2007 on "Physics in Medicine and Biology" [23].

The first part of this Thesis is organized as follow: in the *first chapter* the eye and its movements are described with particular attention on the characteristics of vitreous humour and the metric of the saccadic movement; in *chapter 2* a theoretical overview of the problem is given both in the case of the spherical geometry and the deformed geometry; *chapter 3* describes the principles of PIV technique; in *chapter 4* the experimental set-up used during the experiment is presented; finally *chapter 5* is dedicated to the experiments made for the spherical and deformed sphere model and their results; the experimental model description ends with conclusions and future developments of the study.

Chapter 1

The human eye and its movement

The eye is our window on the world. In this chapter the human eye is described under the anatomical and the physiological aspect. Then, the attention is focused on the characteristics of the possible movement, in particular the saccade, and on the dynamics of vitreous humour, the fluid filling the posterior chamber of this organ.

1.1 Anatomy of the human eye

The eye is the organ of sight (vision) in humans and animals. The eye transforms light waves into visual images. Eighty percent of all information received by the human brain come from the eyes. These organs have almost spherical shape and are housed in the eye (orbital) sockets in the skull. The human eyeball is not perfectly round as its anterior-posterior axis (about 2.3 centimeters) is slightly longer than the crossing axis. The eye wall consists of three layers: the sclera, the choroid, and the retina.

The sclera, the outer fibrous layer, encases and protects the eyeball. In the center of the anterior portion of the sclera there is the cornea, which projects slightly forward. A delicate membrane, the conjunctiva, covers the cornea, that is the visible portion of the sclera. Underneath the sclera lies the choroid. It is composed of a dense pigment and numerous blood vessels that nourish the internal tissues of the eye. At the front end of the choroid, there is the ciliary body that runs like a ring around the visible portion of the eye and contains some muscles that are connected by ligaments to the lens behind the iris. The iris is the visible portion of the choroid. It gives the eye its color, which varies depending on the amount of pigment present in the choroid. Dense pigment

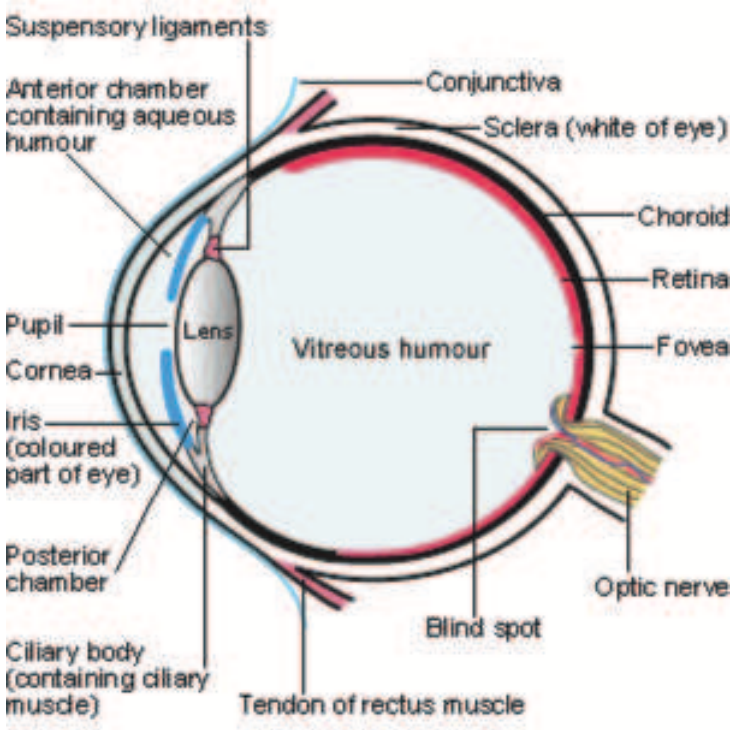

Figure 1.1: Anatomical Structure and eye components

makes the iris brown, while little pigment makes the iris blue. If there is no pigment the iris is pink, as in the eye of a white rabbit or in albino persons. In bright light, muscles in the iris constrict the pupil, reducing its radius and governing the amount of light entering the eye. Conversely, the pupil dilates (enlarges) in dim light, increasing the amount of light entering. The pupil also dilates in consequence of extreme fear, head injuries, and certain drugs. The lens is a crystal-clear, flexible biconvex body sitting behind the iris. The entire surface of the lens is smooth and shiny, contains no blood vessels, and is encased in an elastic membrane. The muscles of the ciliary body hold the lens in place and contracting and relaxing cause the lens to become fatten or thin. The lens' shape changes focuses light on the retina, the photo-sensitive part of eye. The retina is the innermost layer of the eye, it begins at the ciliary body and encircles the entire interior portion of the eye. The retina contains millions of sensitive nerve cells, (called rods and cones) which convert light energy into electrical signals sent to the brain via the optic nerve. Rods cannot detect color, but they are the first cells to detect movement, function chiefly in dim light, allowing limited night vision. Cones function best in bright light and are sensitive to color. In each eye there are about 126 million rods and 6

million cones. In the posterior part of retina, in the center of macula there is the fovea (1.5 mm) where the highest concentration of cone photo-receptors is located. The cavities between the cornea and the iris and between the iris and the lens are filled with a clear watery fluid known as aqueous humor. This fluid aids good vision by helping to maintain eye's shape, providing support for the internal structures, supplying nutrients to the lens and cornea, and disposing of the eyes' cellular waste. The large cavity in back of the lens (the center of the eyeball) is filled with a clear gel-like substance called vitreous humor. It helps to maintain the shape of the eye and to support its internal components. The vitreous humor is 99 percent water and contains no cells. It is transparent as it is crossed by light on its way to the retina. Sight process begins when light waves enter the eye through the cornea, pass through the pupil, then come across the lens behind the iris. The lens focuses light onto the retina, and nerve cells in retinas, convert light energy into electrical impulses. These impulses are then carried by the optic nerve to the brain where they are interpreted as images.

1.1.1 Characteristics of vitreous humour

The posterior chamber is the largest structure of the eye and its shape is approximately round. This cavity is delimitated by the retina behind and by the lens in front. The vitreous humour is the substance which fills the posterior chamber. The vitreous body gives the eye its shape; being in contact with the retina helps to keep it in place by pressing it against the choroid. Moreover, it acts a barrier between the anterior and the posterior segments of the eye for heat and drug transport (Lund-Andersen 2003 [14]). It is crossed by light waves and its refractive index is 1.34, like the aqueous humour. The gel in the vitreous chamber is quite stagnant, unlike the aqueous humour which is continuously replenished. Therefore, if blood, cells or other byproducts of inflammation get into the vitreous, they will remain there unless removed surgically. Vitreous humour property were investigated by Sebag (1992) and other studies by Balazs (1961) and Buchsbaum et al.(1984). It shows like a transparent, colorless, gelatinous mass, composed mainly by water with very few cells and macromolecules (salts, sugars, vitrosin, hyaluronic acid and a network of collagen fibers). Hyaluronic acid and collagen are responsible of gelatinous consistency of vitreous. No blood vessels are present in this body but in axial direction the ialoid channel, heritage of ialoid artery that nourishes the lens in foetus, crosses the cavity. The normal vitreous is an homogeneous gel in youth, but its ageing can produce significant mechanical changes in vitreous consistency, leading to a progressive degeneration of the gel structures, mainly in the center of eye. In fact the collagen fibres are held apart by electrical

charges that tend to reduce. with ageing, and the fibres may clump together. Cells and other organic clusters float freely within the vitreous humour and are perceived in the visual field as spots or fibrous strands. This process of liquefaction may involve the entire posterior chamber and it's reasonable to suppose that a similar phenomenon can modify the action of the vitreous on the eye wall due to the relative motion induced by rapid eye's rotations. Vitreous liquefaction frequently occur in eyes that are nearsighted (myopia). It can also occur after injuries to the eye or inflammation in the eye (uveitis). Vitreous cavity can be filled by a liquid also in vitrectomy surgery, when the surgeon replace a cloudly vitreous with a salt solution or a silicon oil. When vitreous is partially or completely liquefied its motion can be quite intense and, as a consequence of the intense action of vitreous on the retina, retinal detachment can occur. Many authors have supposed a connection between the retinal detachment and the fluid dynamic within the vitreous chamber. Vitreous rheology was investigated by Lee et al.'s(1992)[12], who measured the viscoelastic property of human, bovine and swine vitreous humour in three different location. They found remarkable difference in these properties in the center, in the top and in the bottom of the cavity. The rheology of vitreous can be explained with a model of four parameters called Burgers' model. It deals with a Maxwell element with elastic modulus μ_M and viscosity η_M connected with a Kelvin element with elastic modulus μ_K and viscosity η_M like shown in (1.2).

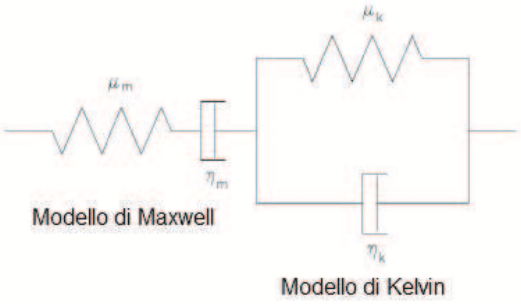

Figure 1.2: Rheological Burgers' model

Maxwell's element represents the elastic and the viscous response of the material while the Kelvin element describes the elastic delayed response, produced by the rendering of the previous deformation of big polymeric chains. The deformation time law J(t) can be derived measuring the deformations induced by a constant load using a creep test:

$$J(t) = J_M + J_K[1 - exp(-t/\tau_K)] + t/\eta_M \tag{1.1.1}$$

Where:

- J_M is the instantaneous elastic response (given by the Maxwell element);
- J_K is the delayed response (given by the Kelvin element);
- τ_K is the time delay.

The qualitative shape of this function is shown in 1.3:

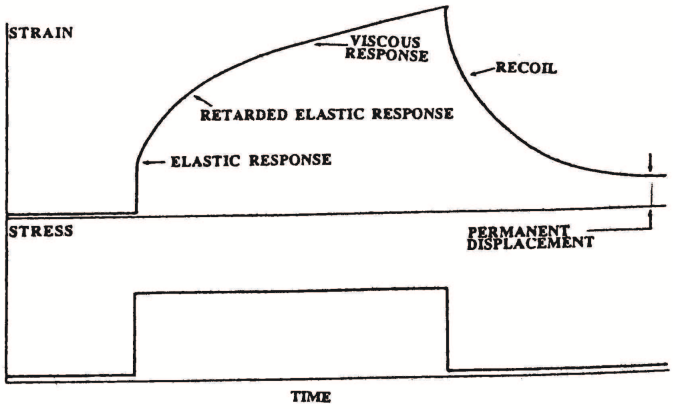

Figure 1.3: Time response of a viscoelastic fluid to a constant stress

Lee et al used a magnetic microrheometer, able to take measurement in little volumes, to perform the creep test. To stress the vitreous humour they put inside it, metallic micro-spheres (Ø approximately 200 m) and then move them by the action of a magnetic field produced by a couple of electromagnets located in the extremes of the sample. The results of Lee et al. are shown in Tab. 1.1.

		Anterior segment	Central segment	Posterior segment
Kelvin	μ_K	2.5	1.27	1,21
Model	η_K	3.67	7.27	3.01
Maxwell	μ_M	1.398	2.18	4,86
Model	η_M	0.313	0.352	4.7

Table 1.1: Values of the Burgers' rheological model of the human vitreous for different eye regions

1.2 Common eye's pathologies

Farsightedness and nearsightedness are common vision disorders. They occur because of a defect in the shape of the eyeball or in the refractive power (ability to bend light rays) of the lens. In these cases, the image perceived from the eye is distorted because the parallel rays of light that enter the eye do not fall perfectly on a tiny hollow (called the fovea) in the retina at the back of the eye.Options for correcting these problems are standard eyeglasses, contact lenses or refractive surgery.

Astigmatism, another common vision disorder, happens when the cornea (sometimes the lens) has an uneven curvature. As a result, some light rays entering the eye focus on the fovea while others focus in front or behind it. Like farsightedness and nearsightedness, astigmatism can be corrected with eyeglasses or contact lenses.

A common pathology is cataract which consists in opacity or cloudiness of the lens which alters the amount of light entering the eye. The most common cataracts are senile cataracts. If not treated, cataract may cause blindness. Clear vision can be restored by a relatively simple surgical procedure in which the entire lens is removed and an artificial lens is implanted.

A serious vision disorder is the glaucoma characterized by increased pressure within the eyeball, caused by a buildup of aqueous humor, which is prevented for some reason from properly draining. Connections are present between pathologies like diabetes or a malfunctioning thyroid gland and this kind of eye disease. Glaucoma leads to permanent blindness if not treated. The condition can be controlled with drugs that either increase the outflow of aqueous humor or decrease its production.

A widely diffused eye disorder is retinal detachment. It happens when the retina peels away from its underlying layer. Initial detachment may be localized, but without rapid treatment the entire retina may detach, leading to vision loss and blindness. Some diseases like myopia or diabetes predispose to the development of a retinal detachment.

Many of the previously described vision disorder are age-related.

1.2.1 The retinal detachment

The retina adheres firmly to the vitreous body but it does not present any adherent to the choroid. A retinal detachment is a pathology consisting in separation of the retina from its attachments to its underlying tissue within the eye. Most retinal detachments are a result of a retinal break, hole, or tear. These retinal breaks may occur when the vitreous gel pulls loose or separates from its attachment to the retina, usually in the peripheral parts of the retina.

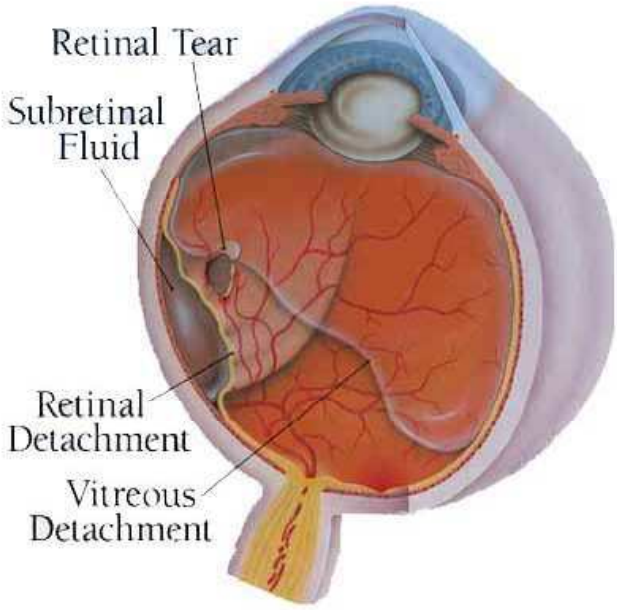

Figure 1.4: Eye affected by vitreous and retinal detachment

This mechanism is called vitreous detachment and is frequent in middle-aged people, in which the vitreous loses its consistence going to a partial liquefaction. Sometimes during this process the vitreous exerts traction on the retina, and if the retina is weak, it will tear. Most retinal breaks are not a result of injury, apart from vitreous detachment, also intense flow to the eye or hereditary predisposition can cause retinal tears. When there is a retinal tear or break retinal detachment easily occurs. The presence of the tear in fact allows the liquefied vitreous humour to pass through the break and to penetrate between the retina and the choroid.

In medical literature we can find three kind of different retinal detachments:

1. The **rhegmatogenous retinal detachment (RRD)**, is the most common form of retinal detachment and occurs when a hole, tear, or break in the retina allows fluid to pass from the vitreous space into the subretinal space. The symptoms of this kind of pathology are floating spots, light flashes, peripheral visual loss and central blurring of vision.

2. The **Exudative retinal detachment (ERD)**, occurs due to inflammation, injury or vascular abnormalities that results in fluid accumulating underneath the retina without the presence of a hole, tear, or break. Diabetes is a risk factor for this form of retinal detachment. Exudative retinal detachment can derive from many other diseases like tumours,

Figure 1.5: Example of vision in a subject with rhegmatogenous retinal detachment

inflammation and macular degenerative conditions. The symptoms associated with exudative retinal detachment are moderate vision loss or a visual field deficit.

3. The **Tractional retinal detachment (TRD)** is caused by the pull of the fibrovascular tissue to the sensory retina from the retinal pigment epithelium. Tractional retinal detachments are typically asymptomatic. Diabetic retinopathy is considered to be the most common cause of tractional retinal detachment.

A rapid diagnosis and treatment of vitreous detachment and retinal tear can prevent a retinal detachment. One time retinal detachment has occurred there are several methods of treating this pathology, mainly dealing with the reattaching retinal strip around a tear to the choroid.

The most recent and mainly employed treatments are laser and cryopexy, which both avoid surgical operations. These treatments create a scar/adhesion around the retinal hole to prevent fluid from entering the hole and accumulating behind the retina. This is made burning the area surrounding the break (laser) or freezing it (cryopexy).

The surgical approaches to to repairing a retinal detachment are substantially three:

1. **Scleral buckle** This treatment consists in positioning one or more silicone bands (bands, tyres) to the outside of the eyeball in correspondence to the retinal hole. The bands push the wall of the eye inward against the retinal hole, closing the hole and allowing the retina to re-attach.

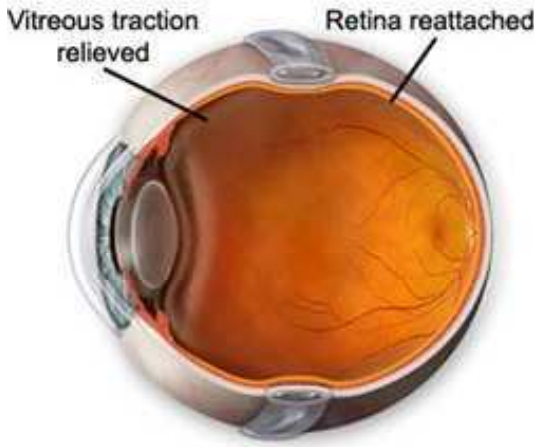

Figure 1.6: Eye with a scleral bulkle

2. **Vitrectomy** is a treatment consisting in the removal of the vitreous gel and its substitution with gas bubble (SF6 or C3F8 gas) or silicon oil.

3. **Pneumatic retinopexy** This treatment involves injecting a gas bubble into the vitreous cavity after which laser or freezing treatment is applied to the retinal hole. The patient's head is positioned in a way such that the gas bubble reasts against the retinal tear. This operation is generally performed in the doctor's office under local anesthesia.

After treatment patients gradually regain their vision over a period of a few weeks, although the visual acuity may not be as good as it was prior to the detachment. However, if left untreated, total blindness could occur in a matter of days after detachment occurrence.

1.3 Movements of human eye

Eyes are subject to various movements. They can move both in the same direction or in different direction: in any case they play like an unique system thanks to the perfect coordination of ocular muscles. Moreover, the nerve centers, make the fusion of retinoic images felt by the two eyes, allowing the stereoscopic vision of the world. Each eye is controlled by six individual muscles called extraocular muscles (to distinguish them from the interior ones). The extraocular muscles attach via tendons to the sclera at different places. Each muscle is responsible for keeping the eye fixated in a direction and also for coordinating the individual eye with the other eye. Normally, in fact, both eyes move together, receive the same image on corresponding locations on both retinas, and the brain fuses these images into one three-dimensional image. The exception is in strabismus which is a disorder where one or both eyes deviate out of alignment, most often outwardly (exotropia) or toward the nose (esotropia). There are four basic muscles for up-and-down and in-and-out motion. These are the rectus muscles (superior, inferior, lateral and medial). The other two muscles are the oblique muscles (superior and inferior). These muscles are responsible for rotational movements that, when combined with the rectus muscles, allow us to look in any direction.

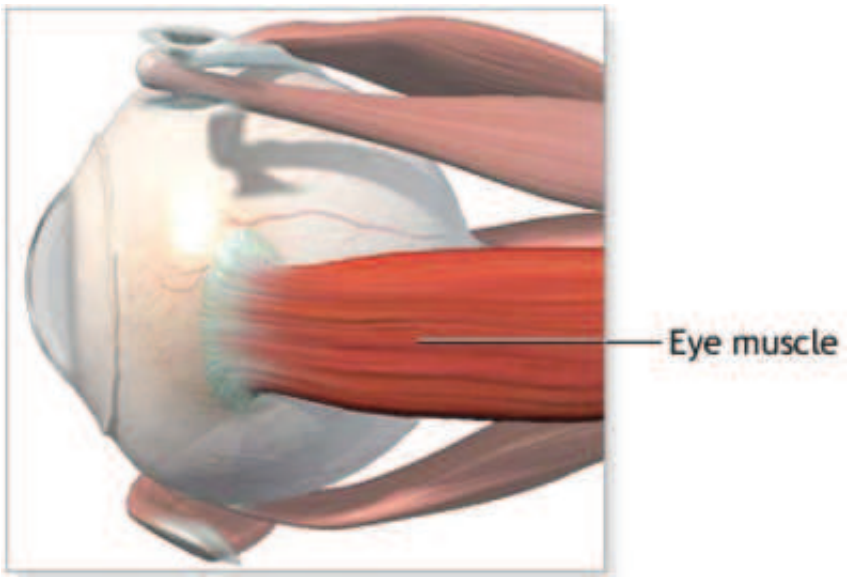

Figure 1.7: Extraocular eye muscles

Each eye can rotate around its three axis doing the movements specified in Tab.(1.2.

Muscle	Main action	Secondary action
Lateral Rectus	Abduction	-
Medial Rectus	Adduction	-
Superior Rectus	Elevation	Adduction - Intortion
Inferior Rectus	Depression	Adduction - Extortion
Superior Oblique	Depression	Abduction - Intortion
Inferior Oblique	Elevation	Abduction - Extortion

Table 1.2: Ocular muscles and their actions

Rotations around vertical axis are called adduction when they are in nose direction, while they are called abduction when they are directed toward the external.

1.3.1 The saccade

In this work we are concerned with a special kind of movement called "saccadic movement", that, for their characteristics effect the most severe condition of motion within the vitreous chamber. The saccades are little, fast concurrent movements that happen when the eyes go from a point of the visual field to another . You do it, for example, reading this page, moving your eyes from the beginning of the line to successive points of the line stopping the look for short instants on them and then coming back to the subsequent line.

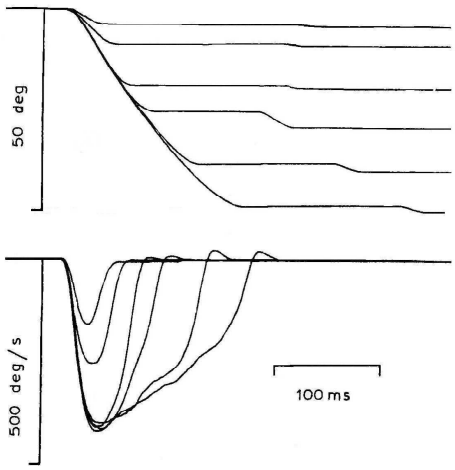

Figure 1.8: Saccadic movements of amplitude 5, 10, 20, 30, 40 e 50 degrees; position (upper graphic) and velocity (lower graphic).

Their kinematic characteristics are reported in detail in Becker(1989). The main features of saccadic movements are a very high initial angular acceleration (up to 30000 deg/s^2) that suddenly starts the movement, a somewhat less intense deceleration that is nevertheless capable of inducing a very efficient stop of the movement. The velocity peak increases correspondingly to the saccade amplitude up to a saturation value corresponding to 400 - 600 deg/s. Saccade amplitude ranges from 0.05 deg (microsaccades, usually spread in a random manner around the peered object) to 80 - 90 deg, which is the physical limit of the orbit (notice that saccades of large amplitude are unusual as such extension are usually accompanied by head rotation). We can describe the metric of saccade using three fundamental characteristics: the amplitude A, the duration D, the peak angular velocity ω_p and the acceleration time t_p i.e. the time required to reach the peak velocity starting to the rest.

Duration

Saccade duration is related to the amplitude of movements. According to Becker (1989) if we consider the whole interval of possible amplitudes in physiological condition the relationship between D and A is non linear, but narrowing the amplitude range from 5. to 50 deg the relationship is very well described by the following linear law:

$$D = D_0 + dA \tag{1.3.1}$$

In equation 1.3.1 D is measured in seconds and A in degrees.

According to measurements performed by the author the slope d approximately assumes the value of 2.5 ms/deg and the intercept D_0 typically ranges between 20 and 30 ms.

Some authors propose for microsaccades a non linear dependence in accordance with the following power law

$$D = CA^k \tag{1.3.2}$$

with $C > 0$ and $A < 1$.

For big amplitude saccades linear relationship can be used with a good accordance, taking in account that in normal conditions, such movements are accompanied by head rotations.

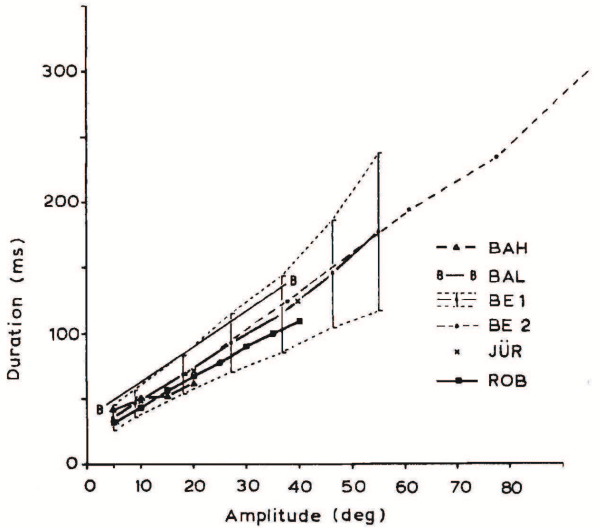

Figure 1.9: Saccades duration related to their amplitude, according to some authors who have measured them with different techniques:BAH Bahill (1981), measurements on 13 subjects with infrared technique; BAL Balloh (1975), observation on 25 subjects, EOG; BE1 Becker (not published), 26 subjects population, EOG, the function is contained inside a 95% confidence band ; BE2 Becker (not published), 6 subjects, EOG, measurements with head able to rotate; JUR Jurgens e Becker (not published), 10 subject; ROB Robinson (1964), 3 subjects.

Velocity

The average angular velocity during a saccadic movement depends on amplitude and saccade duration and is defined as:

$$\bar{\omega} = \frac{A}{D} \tag{1.3.3}$$

On the contrary the peak velocity is independent of the duration and measurements suggest that the ratio $\frac{\omega_p}{\bar{\omega}}$ between the peak and the mean velocities attains a fairly constant value. Becker (1989) suggests the following estimate:

$$\frac{\omega_p}{\bar{\omega}} \approx 1.64 \tag{1.3.4}$$

The Figure 1.10 below shows that the peak velocity initially increases proportionally to saccade amplitude, then it asintotically tents to a limit value for larger amplitudes.

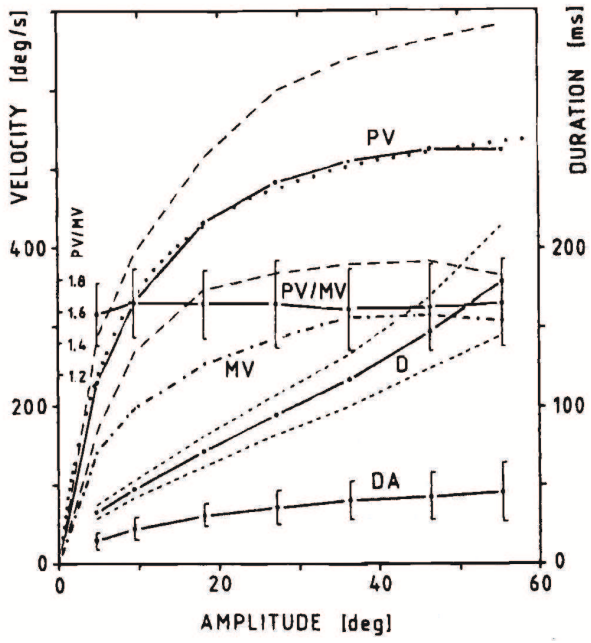

Figure 1.10: Peak velocity related to saccade amplitude. The lines are referred to the authors of figure 1.3.1 and BOG is referred to Boghen et al.(1974), 15 subjects observations, EOG with infrared, function contained inside a 95% confidence band.

Acceleration

The experimental data show that small amplitude saccades (smaller than 10 *deg*) follow an almost symmetrical time law, the acceleration time being approximatively equal to 0.45 D. The dimensionless acceleration time $\frac{t_p}{D}$ varies linearly with increasing saccade amplitudes, to the value $\frac{t_p}{D} \approx 0.25$ for saccades of 50 *deg*. Experimental studies show that, for amplitudes of about 10 *deg*, peak acceleration goes about 30000 deg/s^2 while for larger amplitudes it can reaches 35000 deg/s^2.

Time law for saccades

Some author (see David (1989) [7]) have suggested to use a periodic time law to reproduce the saccade. This approximation is quite good for low saccades amplitude.

In order to reproduce the time law $\varphi(t)$, in the present work, it is described the angular displacement of the eye during the saccade, employing a fifth

degree polynomial function, of the form:

$$\varphi(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 + c_5 t^5 \tag{1.3.5}$$

The six coefficients are computed imposing the following constraints:

$$\begin{aligned} &\varphi(0) = 0 \\ &\varphi(D) = A \\ &\dot{\varphi}(0) = 0 \\ &\dot{\varphi}(t_p) = \omega_p \\ &\ddot{\varphi}(t_p) = 0 \end{aligned} \tag{1.3.6}$$

where D, ω_p and t_p are evaluated by the relationships described above and superscript dots denote time derivatives.

The resulting functions $\varphi(t)$ and $\omega(t) = \frac{\partial \varphi(t)}{\partial t}$ are shown in Figure 1.11 a) and b) and they satisfactorily reproduce the main features of real saccadic movements as reported by Becker (1989).

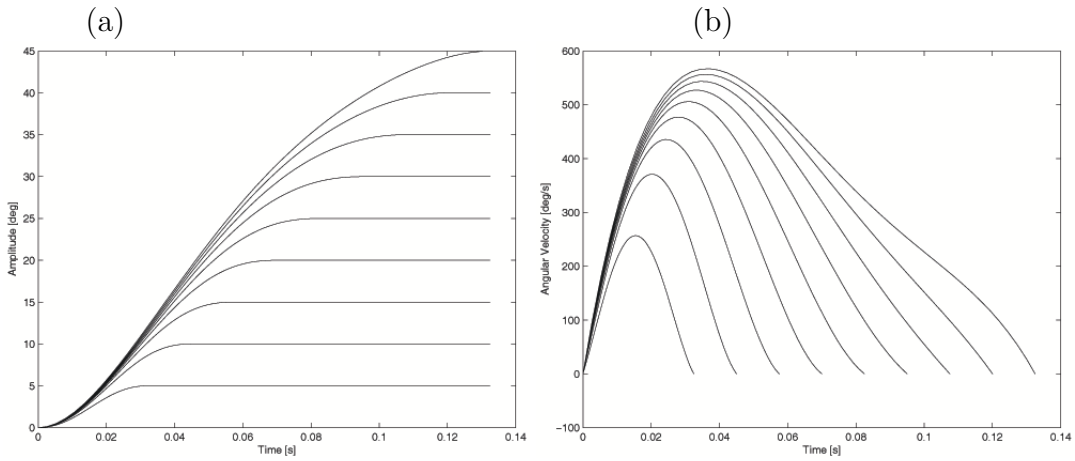

Figure 1.11: Polynomial saccade position (a) Polynomial saccade velocity (b)

Chapter 2

Theoretical background

In this chapter a mathematical model is presented both for the motion inside a sphere and a deformed sphere. The cavity rotates around its axis with a defined time law, under simplified conditions listed below. The treatment refers to David et al (1998)[7] for the spherical model and to Repetto (2006)[18] for the sphere with the indentation simulating the lens presence.

2.1 A spherical model

David et al (1998)[7] have proposed a simplified analytical model for the motion of a viscoelastic fluid within a periodically rotating sphere, which will shortly be described in the following for the case of a purely viscous fluid. The eye is supposed to be a sphere, with interior radius R_0. The indentation caused by the lens presence is disregarded as the different dimensions of three ocular axis. Let us consider a system of spherical coordinates (r, θ, φ), with θ and φ the elevation angular coordinate and the azimuthal coordinate, respectively. Let us also assume that the sphere rotates about the vertical axis $\theta = 0, \pi$, as shown in Fig. 2.1.

If the flow field is described in terms of the velocity components $(u_r, u_\theta, u_\varphi)$ the Navier-Stokes equations for a purely viscous fluid are:

$$\left[\frac{\partial u_r}{\partial t} + u_r\frac{\partial u_r}{\partial r} + \frac{u_\theta}{r}\frac{\partial u_r}{\partial \theta} + \frac{u_\varphi}{r\sin\theta}\frac{\partial u_r}{\partial \varphi} - \frac{u_\theta^2 + u_\varphi^2}{r}\right] = F_r - \frac{\partial P}{\partial r} + \nu\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_r}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial u_r}{\partial \theta}\right) + \frac{1}{r^2\sin\theta^2}\frac{\partial u_r^2}{\partial \varphi^2} - \frac{2u_r}{r^2} - \frac{2}{r^2}\frac{\partial u_\theta}{\partial \theta} - \frac{2u_\theta\cot\theta}{r^2} - \frac{2}{r^2\sin\theta}\frac{\partial u_\varphi}{\partial \varphi}\right] \quad (2.1.1)$$

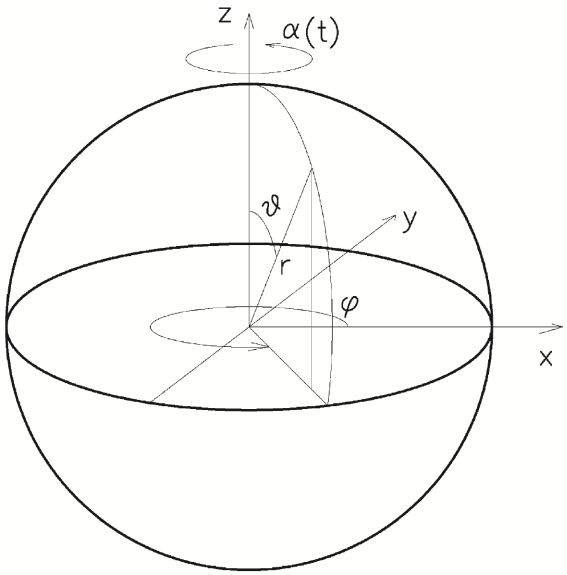

Figure 2.1: Spherical system of coordinate used in the analytical model of David et al. (1998)[7]

$$\left[\frac{\partial u_\theta}{\partial t}+u_r\frac{\partial u_\theta}{\partial r}+\frac{u_\theta}{r}\frac{\partial u_\theta}{\partial \theta}+\frac{u_\varphi}{r\sin\theta}\frac{\partial u_\theta}{\partial \varphi}+\frac{u_r u_\theta}{r}-\frac{u_\varphi^2}{r}\cot\theta\right]= F_\theta-\frac{1}{r}\frac{\partial P}{\partial \theta}+\nu\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_\theta}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial u_\theta}{\partial \theta}\right)+\frac{1}{r^2\sin\theta^2}\frac{\partial u_\theta^2}{\partial \varphi^2}-\frac{2}{r^2}-\frac{u_\theta}{r^2\sin\theta^2}-\frac{2\cos\theta}{r^2\sin\theta^2}\frac{\partial u_\varphi}{\partial \varphi}\right] \quad (2.1.2)$$

$$\left[\frac{\partial u_\varphi}{\partial t}+u_r\frac{\partial u_\varphi}{\partial r}+\frac{u_\theta}{r}\frac{\partial u_\varphi}{\partial \theta}+\frac{u_\varphi}{r\sin\theta}\frac{\partial u_\varphi}{\partial \varphi}+\frac{u_r u_\varphi}{r}-\frac{u_\varphi u_\theta\cot\theta}{r}\right]= F_\varphi-\frac{1}{r\sin\theta}\frac{\partial P}{\partial \varphi}+\nu\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_\varphi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial u_\varphi}{\partial \theta}\right)+\frac{1}{r^2\sin\theta^2}\frac{\partial u_\varphi^2}{\partial \varphi^2}-\frac{u_\varphi}{r^2\sin\theta^2}+\frac{2}{r^2\sin\theta^2}\frac{\partial u_r}{\partial \varphi}-\frac{2\cos\theta}{r^2\sin\theta^2}\frac{\partial u_\theta}{\partial \varphi}\right] \quad (2.1.3)$$

where F represents mass forces acting on the fluid volume and P is the pressure.

To (2.1.1), (2.1.2) and (2.1.3) equations no slip boundary conditions on the wall have to be associated (in r=0 symmetry conditions have to be considered that lead to $u_k = 0$, for $k = r, \theta, \phi, \varphi$:

$$u_r = 0 \qquad (r = R_0) \tag{2.1.4}$$
$$u_\theta = 0 \qquad (r = R_0) \tag{2.1.5}$$
$$u_\varphi = u \sin(\omega t) \sin\theta \qquad (r = R_0) \tag{2.1.6}$$

Then introducing the dimensionless variables:

$$u^* = \frac{u}{U} \tag{2.1.7}$$
$$r^* = \frac{r}{R_0} \tag{2.1.8}$$
$$t^* = t\omega \tag{2.1.9}$$
$$P^* = \frac{P}{\rho U^2} \tag{2.1.10}$$

where R, $U = \omega_p R$ and $\omega = 1/T$ are respectively a length, velocity and time scales.

In the following P is assumed to be dynamic pressure and the dimensionless variables are intended without *. The equations (2.1.1) , (2.1.2) , (2.1.3) in a dimensionless form result:

$$\begin{aligned}\alpha^2 \frac{\partial u_r}{\partial t} + Re \left[u_r \frac{\partial u_r}{\partial r} + \frac{u_\theta}{r} \frac{\partial u_r}{\partial \theta} + \frac{u_\varphi}{r \sin\theta} \frac{\partial u_r}{\partial \varphi} - \frac{u_\theta^2 + u_\varphi^2}{r} \right] = \\ - Re \frac{\partial P}{\partial r} + \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial u_r}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial u_r}{\partial \theta} \right) \right. \\ \left. + \frac{1}{r^2 \sin\theta^2} \frac{\partial u_r^2}{\partial \varphi^2} - \frac{2u_r}{r^2} - \frac{2}{r^2} \frac{\partial u_\theta}{\partial \theta} - \frac{2u_\theta \cot\theta}{r^2} - \frac{2}{r^2 \sin\theta} \frac{\partial u_\varphi}{\partial \varphi} \right]\end{aligned} \tag{2.1.11}$$

$$\begin{aligned}\alpha^2 \frac{\partial u_\theta}{\partial t} + Re \left[u_r \frac{\partial u_\theta}{\partial r} + \frac{u_\theta}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_\varphi}{r \sin\theta} \frac{\partial u_\theta}{\partial \varphi} + \frac{u_r u_\theta}{r} - \frac{u_\varphi^2}{r} \cot\theta \right] = \\ - Re \frac{1}{r} \frac{\partial P}{\partial \theta} + \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial u_\theta}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial u_\theta}{\partial \theta} \right) \right. \\ \left. + \frac{1}{r^2 \sin\theta^2} \frac{\partial u_\theta^2}{\partial \varphi^2} - \frac{2}{r^2} - \frac{u_\theta}{r^2 \sin\theta^2} - \frac{2\cos\theta}{r^2 \sin\theta^2} \frac{\partial u_\varphi}{\partial \varphi} \right]\end{aligned} \tag{2.1.12}$$

$$\alpha^2\frac{\partial u_\varphi}{\partial t}+Re\left[u_r\frac{\partial u_\varphi}{\partial r}+\frac{u_\theta}{r}\frac{\partial u_\varphi}{\partial\theta}+\frac{u_\varphi}{r\sin\theta}\frac{\partial u_\varphi}{\partial\varphi}+\frac{u_r u_\varphi}{r}-\frac{u_\varphi u_\theta\cot\theta}{r}\right]=$$
$$-Re\frac{1}{r\sin\theta}\frac{\partial P}{\partial\phi}+\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_\varphi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u_\phi}{\partial\theta}\right)\right.$$
$$\left.+\frac{1}{r^2\sin\theta^2}\frac{\partial u_\varphi^2}{\partial\varphi^2}-\frac{u_\varphi}{r^2\sin\theta^2}+\frac{2}{r^2\sin\theta^2}\frac{\partial u_r}{\partial\varphi}-\frac{2\cos\theta}{r^2\sin\theta^2}\frac{\partial u_\theta}{\partial\varphi}\right] \quad (2.1.13)$$

where $\alpha=\sqrt{\frac{\omega R_0^2}{\nu}}$ is the Womersley dimensionless number, ratio between the characteristic length scale of the flow domain (R_0), and the thickness of the boundary layer at the wall, which order is $\sqrt{\frac{\nu}{\omega}}$. $Re=\frac{uR_0}{\nu}$ is the Reynolds number. At low Reynolds numbers, viscous effect are predominant and allow to neglect all the terms multiplying Re.

Moreover, for the symmetry of the domain all the φ derivatives are neglected. Equations (2.1.11), (2.1.12) and (2.1.13) can be rewritten in a simplified form:

$$\frac{\partial u_r}{\partial t}=\frac{1}{\alpha^2}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_r}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial u_r}{\partial\theta}\right)+\right.$$
$$\left.-\frac{2u_r}{r^2}-\frac{2}{r^2}\frac{\partial u_\theta}{\partial\theta}-\frac{2u_\theta\cot\theta}{r^2}\right] \quad (2.1.14)$$

$$\frac{\partial u_\theta}{\partial t}=\frac{1}{\alpha^2}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_\theta}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial u_\theta}{\partial\theta}\right)+\right.$$
$$\left.-\frac{2}{r^2}-\frac{u_\theta}{r^2\sin\theta^2}\right] \quad (2.1.15)$$

$$\frac{\partial u_\varphi}{\partial t}=\frac{1}{\alpha^2}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_\varphi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial u_\varphi}{\partial\theta}\right)+\right.$$
$$\left.-\frac{u_\varphi}{r^2\sin\theta}\right] \quad (2.1.16)$$

The authors assume that the flow field is axis-symmetrical, as a consequence of the symmetry of the domain, and that the velocity vectors lay anywhere on planes perpendicular to the axis of rotation. We will prove that this assumption is correct, being the secondary flow induced by the curvature of the wall,

consisting in two toroidal vortices, at least three or four orders of magnitude smaller than the circumferential flow both in the case of periodic rotations and in the case of saccadic movements. The simplifying assumptions adopted by David et al (1998) [7] imply that $u_r = u_\theta = 0$ and that u_φ does not depend on φ. Notice that the above assumptions can be formally justified on a mathematical ground only in the case of small values of the Reynolds number, i.e. when the nonlinear terms in the Navier Stokes equations can be neglected. In the experiments performed the Reynolds number (computed adopting the maximum wall velocity and the sphere radius) ranges between 1 and 10. This implies that the above assumptions are not strictly mathematically justified.

David et al.(1998) considered the case of periodic sinusoidal rotations, according to experimental results of Weber e Daroff (1972):

$$u_\varphi = u \sin(\omega t) \sin\theta \tag{2.1.17}$$

where ω is the angular frequency of movement and U is the maximum tangential velocity. As the u_φ problem is linear and uncoupled, the authors sought a separable solution in the form:

$$u_\varphi = e^{i\omega t} g(r) \sin\theta + c.c. \tag{2.1.18}$$

where c.c. denotes the complex conjugate and the dependance of the solution from $\sin\theta$ arises from the boundary condition at the wall (2.1.4). The solution for the function g is found in the following form:

$$g(r) = c j_1(kr) \tag{2.1.19}$$

where j_1 is the spherical Bessel function of the first kind and order 1. The coefficient k in the equation 2.1.19 is defined as:

$$k = \frac{\sqrt{2}}{2}(1-i)\sqrt{\frac{\omega}{\nu}} \tag{2.1.20}$$

The constant c can be determined imposing the boundary condition at the side wall and takes the following values:

$$c = \frac{1}{2 j_1(kR)} \tag{2.1.21}$$

$$c = -\frac{i}{2 j_1(kR)} \tag{2.1.22}$$

depending on whether F(t) is equal to $cos(\omega t)$ or $sin(\omega t)$. As pointed out by David et al (1998), the above basic solution can be used to build up more

complicated ones induced by sphere rotations that are not sinusoidal in time. Indeed, periodic functions F(t), with period T , can be Fourier decomposed in the form:

$$F(t) = c_0 + \sum_{n=1}^{\infty} c_n \cos\left(\frac{2\pi n}{T}t\right) + s_n \sin\left(\frac{2\pi n}{T}t\right) \tag{2.1.23}$$

In the present case, in order to reproduce saccadic eye movements, we have employed a time law consisting of a sequence of rotations described by the fifth-order polynomial function 1.3.5 followed by a period of no motion long enough to allow for a complete stop of the fluid motion. In this case, the coefficients c_0, c_n and s_n can be readily determined analytically and the solution for the φ component of the velocity can be written in the form:

$$\begin{aligned} u_\varphi = \sin\vartheta \left[c_0 \frac{r}{R} + \sum_{n=1}^{\infty} c_n \left(e^{\frac{2\pi n i t}{T}} \frac{1}{2j_1(k_n R)} j_1(k_n r) + c.c. \right) \right] + \\ + \sin\vartheta \sum_{n=1}^{\infty} s_n \left(-e^{\frac{2\pi n i t}{T}} \frac{1}{2j_1(k_n R)} j_1(k_n r) + c.c. \right) \end{aligned} \tag{2.1.24}$$

2.1.1 Theoretical predictions

Examples ofu_φ profiles on the equatorial plane of the sphere, along the radius for different instants during a period are reported in figure 2.2, 2.3 and 2.4. All the variables have been made dimensionless with their maximum [1].

As we can see in figure 2.2, 2.3 and 2.4 the thickness of boundary layer at the wall increases decreasing the Womersley number, this result corresponds to the weakening of viscous effects, whose relative magnitude with respect to the inertial forces is contained in α definition. Velocity profiles highlight the existence of a phase lag between the motion of the fluid in the deeper regions of the sphere and that close to the wall.

Tangential stresses at the wall can be derived by the velocity profiles trough the relation:

$$\tau_{r\varphi} = \frac{\partial u_\varphi}{\partial r} - \frac{u_\varphi}{r} \tag{2.1.25}$$

Notice that in the 2.1.25 the term $\frac{1}{r}\frac{\partial u_r}{\partial \varphi}$ was omitted having assumed that the radial component is negligible respect to the circumferential one.

[1] Dimensionless time, obtained dividing the dimensional time by the period of rotation, ranges between 0 e 2π.

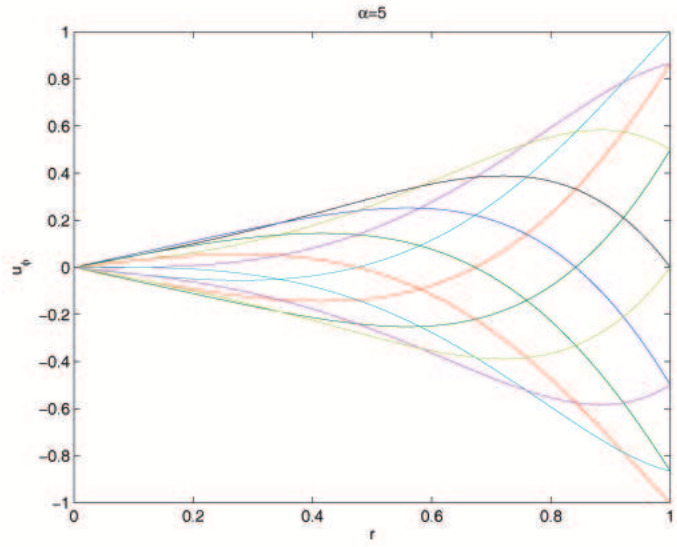

Figure 2.2: Dimensionless u_φ profiles for $\alpha = 5$ along the radius

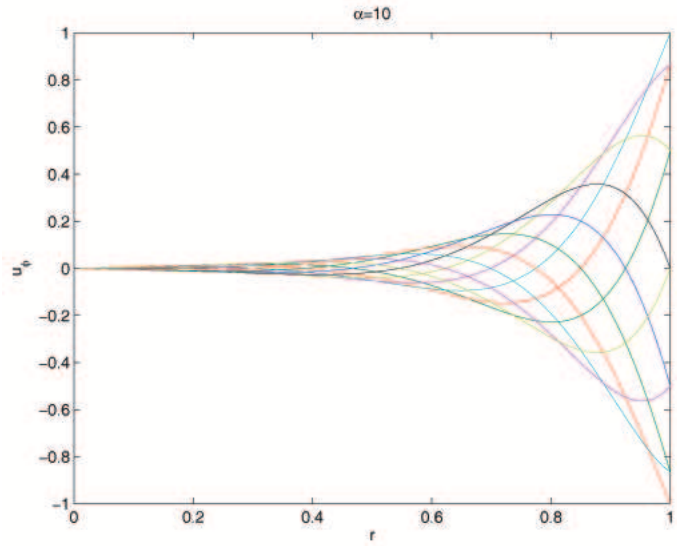

Figure 2.3: Dimensionless u_φ profiles for $\alpha = 10$ along the radius

Figure 2.5 shows the dimensionless tangential stresses:

$$\tau^*_{r\varphi} = \frac{\tau_{r\varphi} R_0}{\mu U} \tag{2.1.26}$$

at the wall versus the time, for different values of α. The superscript * in the 2.1.26 denotes the dimensionless stress.

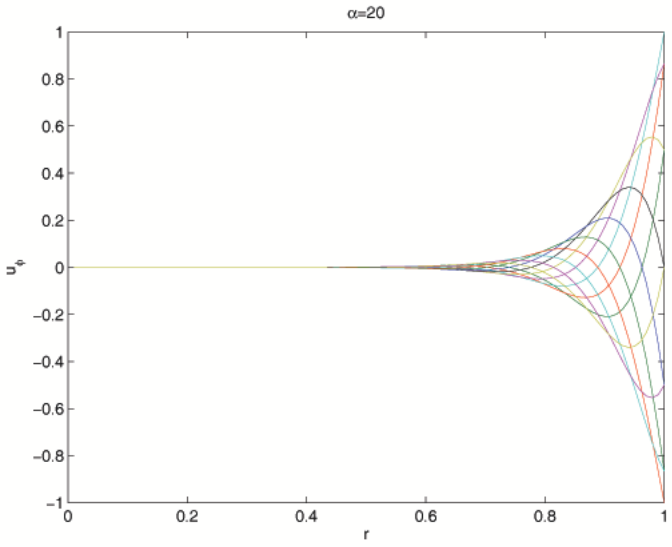

Figure 2.4: Dimensionless u_φ profiles for $\alpha = 20$ along the radius

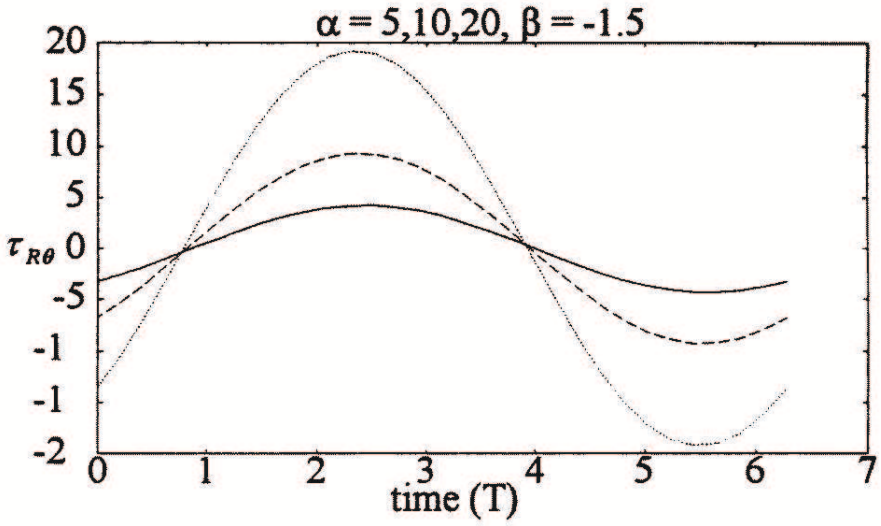

Figure 2.5: $\tau_{r\varphi}$ for different α versus time

David et al.(1998) supposed that the maximum tangential stress is proportional to the Womersley parameter according to the following relation:

$$max(\tau_{r\varphi}) = k * \alpha \tag{2.1.27}$$

in which k is a constant ~ 1.

Maximum tangential stresses obtained by the authors with the (2.1.27) are listed in table 2.1 for $\alpha = 5$, 10 and 20 respectively.

α	**Maximum value of** $\tau_{r\varphi}$
5	4.25
10	9.25
20	19.26

Table 2.1: Maximum tangential stress at the wall for different α.

The theoretical predictions highlight that tangential stresses increase proportionally to the radius:

$$\alpha = \sqrt{\frac{\omega R_0^2}{\nu}} \propto R_0 \tag{2.1.28}$$

according with the fact that RD often occurs in myopic eyes characterized by an antero-posterior axis dimension longer than the normality.

2.2 A deformed sphere model

As previously explained in chapter 1 the shape of the anterior part of the vitreous chamber of a real eye significantly differs from a sphere due to the presence of the lens. Repetto (2006) [18] investigated from an analytical point of view the vitreous motion inside a weakly deformed sphere assuming the vitreous to be liquefied. The analysis is carried out in the limit of low viscosity Newtonian incompressible fluid so that the boundary layer is thin close to the wall and the core flow is irrotational[2]. Because of the previous hypothesis,the boundary layer does not separate during a saccadic movement. The eye globe is modelled as a rigid deformed sphere rotating about its vertical axis. The vertical cross section of the considered domain is shown in figure 2.6

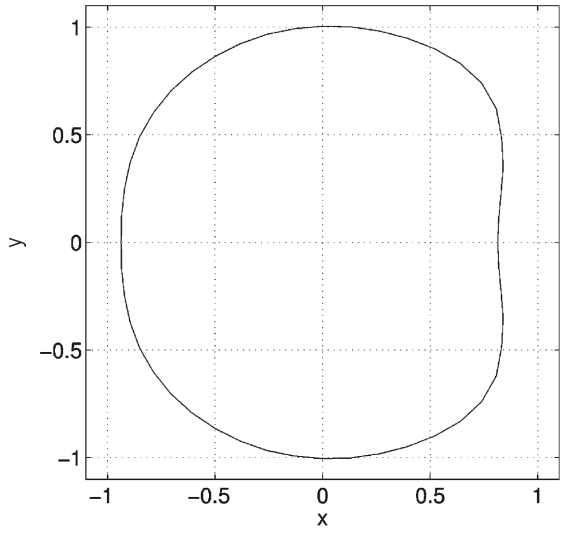

Figure 2.6: Vertical cross section of the considered domain

The mathematical problem is formulated employing the system of polar spherical coordinates (r, θ, φ) shown in figure 2.1 and referring to a fixed frame. We can define the velocity potential Φ as follow:

$$\mathbf{u} = \nabla\Phi \tag{2.2.1}$$

where $\mathbf{u} = (u_r, u_\theta, u_\varphi)$ is the velocity. Fluid incompressibility implies that

[2]The viscous layer at the wall spreads over a thickness of the order of $\sqrt{\nu t}$, where ν and t are the kinematic viscosity of the fluid and time, respectively. For a low viscosity fluid -say like water- and a time scale like saccadic one the viscous layer is at least one order of magnitude lower than eye radius.

the velocity potential must be a harmonic function, i.e.:

$$\nabla^2\Phi = 0 \tag{2.2.2}$$

The boundary conditions impose vanishing flux through the eye wall, defined by a radial coordinate $R(\theta, \varphi, t)$. The equations governing the fluid flow, in term of velocity potential, can then be written in the dimensional form:

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\Phi}{\partial\varphi^2} = 0 \tag{2.2.3}$$

$$-\frac{\partial R}{\partial t} + \frac{\partial\Phi}{\partial r} - \frac{1}{r^2}\frac{\partial\Phi}{\partial\theta}\frac{\partial R}{\partial\theta} - \frac{1}{r^2\sin^2\theta}\frac{\partial\Phi}{\partial\varphi}\frac{\partial R}{\partial\varphi} = 0 \quad [r = R(\theta, \varphi - \alpha(t))] \tag{2.2.4}$$

$$\frac{\partial\Phi}{\partial t} + \frac{P}{\rho} + \frac{1}{2}\left[\left(\frac{\partial\Phi}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial\Phi}{\partial\theta}\right)^2 + \left(\frac{1}{r\sin\theta}\frac{\partial\Phi}{\partial\varphi}\right)^2\right] = f(t) \tag{2.2.5}$$

having introduced the dynamic pressure P through the Bernoulli equation (2.2.5) in the case of an unsteady irrotational flow. In the previous equations $\alpha(t)$ is the angular position in time with respect to a reference position, ρ is the fluid density and $f(t)$ an arbitrary function eventually vanishing by suitably redefining the potential function.

Then introducing the dimensionless variables:

$$r* = \frac{r}{R} \quad R* = \frac{r}{R} \tag{2.2.6}$$

$$\Phi* = \frac{\Phi}{\omega_p R^2} \tag{2.2.7}$$

$$P* = \frac{p}{\rho\omega_P^2 R^2} \tag{2.2.8}$$

$$t* = \omega_p t \tag{2.2.9}$$

where R is the sphere radius and ω_p is the peak angular velocity during a saccadic movement (see chapter 1). Here dimensionless variables present a superscript asterisk that, for simplicity, will be omitted in the follow. Making the following change of variable:

$$\phi = \varphi - \alpha(t) \tag{2.2.10}$$

the position of the wall is no longer time dependent.

The set of equations (2.2.3), (2.2.4) and (2.2.5) can be rewritten in the dimensionless form with the above assumptions.

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial \Phi}{\partial r}\right)+\frac{1}{\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial \Phi}{\partial \theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 \Phi}{\partial \phi^2}=0 \tag{2.2.11}$$

$$\frac{\partial \alpha}{\partial t}\frac{\partial R}{\partial \varphi}+\frac{\partial \Phi}{\partial r}-\frac{1}{r^2}\frac{\partial \Phi}{\partial \theta}\frac{\partial R}{\partial \theta}-\frac{1}{r^2\sin^2\theta}\frac{\partial \Phi}{\partial \varphi}\frac{\partial R}{\partial \varphi}=0\,[r=R(\theta,\varphi)] \tag{2.2.12}$$

$$\frac{\partial \Phi}{\partial t}-\frac{\partial \alpha}{\partial t}\frac{\partial R}{\partial \varphi}+P+\frac{1}{2}\left[\left(\frac{\partial \Phi}{\partial r}\right)^2+\left(\frac{1}{r}\frac{\partial \Phi}{\partial \theta}\right)^2+\left(\frac{1}{r\sin\theta}\frac{\partial \Phi}{\partial \varphi}\right)^2\right]=0 \tag{2.2.13}$$

The shape of the domain considered is that of a sphere with unit radius, conveniently deformed. Being the sphere center located in the origin of a cartesian system of axis, let us consider the $R(\theta,\varphi)$ function such that:

$$R(\theta,\varphi)=1+\delta R_1(\theta,\varphi) \tag{2.2.14}$$

where $\delta R_1(\theta,\varphi)$ represents the deviation of the actual domain geometry from the sphere. The parameter δ corresponds to the maximum value of this deviation and was assumed equal to 0.15. The function $R_1(\theta,\varphi)$ can be expanded in spherical harmonics (see for further information Appendix B):

$$R_1=\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}a_{mn}\cos(m\varphi)\,P_n^m(\cos\theta) \tag{2.2.15}$$

in which $P_n^m(\cos\theta)$ are the associated Legendre functions (see for further information Appendix D):

$$P_n^m(x)=\left(1-x^2\right)^{m/2}\frac{d^m}{dx^m}P_n(x) \tag{2.2.16}$$

As shown in appendix D the expansion coefficients a_{mn} can be easily computed by integrating the (2.2.15), using the orthogonal properties of Legendre functions. Now, let us expand the unknowns p and Φ in a power series of δ:

$$\Phi=\Phi_0+\delta\Phi_1+O\left(\delta^2\right) \tag{2.2.17}$$

$$P=P_0+\delta P_1+O\left(\delta^2\right) \tag{2.2.18}$$

The above expansion allow us to linearize the problem (2.2.11), (2.2.12) and (2.2.13) by substituting the variables with their expansion 2.2.17 restricted to

the first order of approximation. This procedure is strictly valid only for $\delta \ll 1$, while $\delta = 0.15$, nevertheless, the solution in this case is worth to perceive the principal features of the fluid dynamic and to be analytical. Solutions at further order of approximation, will be conveniently found in a numerical way. The solution at $O(\delta^0)$ order is the trivial: $\Phi_0 = 0$ and p_0 =contant. This agree with the fact that in absence of the adherence condition at the wall, no motion is generated inside a rotating sphere. The equation at the first order of approximation are:

$$\nabla^2 \Phi_1 = 0 \tag{2.2.19}$$

$$\frac{\partial \Phi_1}{\partial r} = -\frac{\partial \alpha}{\partial t}\frac{\partial R_1}{\partial \varphi} \qquad (r = 1) \tag{2.2.20}$$

$$P_1 = -\frac{\partial \Phi_1}{\partial t} + \frac{\partial \alpha}{\partial t}\frac{\partial \Phi_1}{\partial \varphi} \tag{2.2.21}$$

Expanding the first order coefficient Φ_1 of the velocity potential such that:

$$\Phi_1 = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty} \Phi_{mn} \sin(m\varphi)\, P_n^m(\cos\theta) \tag{2.2.22}$$

and substituting the expansion in the 2.2.19, the problem reduces to a sequence of ordinary differential equations in the expansion coefficient Φ_{mn}:

$$\frac{d}{dr}\left(r^2 \frac{d\Phi_{mn}}{dr}\right) - n(n+1)\Phi_{mn} = 0 \tag{2.2.23}$$

$$\frac{d\Phi_{mn}}{dr} = m a_{mn} \frac{\partial \alpha}{\partial t} \qquad (r = 1) \tag{2.2.24}$$

$$\Phi_{mn} = 0 \qquad (r = 0) \tag{2.2.25}$$

Notice that to solve the first equation of the second order we have to impose two boundary conditions, one is provided at the external boundary and the other is a regularity condition at the center of the chamber. The problem 2.2.23 admits the following solutions:

$$\Phi_{mn} = a_{mn} \frac{\partial \alpha}{\partial t} \frac{m}{n} r^n \tag{2.2.26}$$

From which it is possible to compute the velocity potential and derive the

components of the velocity fields.

$$u_{r1} = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty} a_{mn}\frac{\partial\alpha}{\partial t} m r^{n-1}\sin(m\varphi)\,P_n^m(\cos\theta) \tag{2.2.27}$$

$$u_{\theta 1} = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty} a_{mn}\frac{\partial\alpha}{\partial t}\frac{m}{n} r^{n-1}\sin(m\varphi)\,\frac{d}{d\theta}P_n^m(\cos\theta) \tag{2.2.28}$$

$$u_{\varphi 1} = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty} \frac{a_{mn}\frac{\partial\alpha}{\partial t}\frac{m^2}{n} r^{n-1}\cos(m\varphi)\,P_n^m(\cos\theta)}{\sin\theta} \tag{2.2.29}$$

Finally from the linearized Bernoulli equation we obtain the solution for the pressure:

$$p = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\left[-a_{mn}\frac{\partial^2\alpha}{\partial t^2}\frac{m}{n} r^n\sin(m\varphi)\,P_n^m(\cos\theta) + a_{mn}\left(\frac{\partial\alpha}{\partial t}\right)^2\frac{m^2}{n} r^n\cos(m\varphi)\,P_n^m(\cos\theta)\right] \tag{2.2.30}$$

Notice that the p solution is composed by a term proportional to the angular acceleration and a term dependent from the angular velocity.

2.2.1 Theoretical predictions

In the following, some results of the above theoretical model are summarized, for further information the reader is referred to Repetto (2006) [18]. Theoretical predictions show that the real shape of the domain plays a significant role, being responsible of an intense flow with remarkably different characteristics from the viscous flow occurring within a rotating sphere. The presence of the lens indentation actually produces a high three-dimensional motion within the vitreous chamber and generates significant vertical velocity components which would not be observed if the eye globe were purely spherical. A flow particularly intense is generated close to the lens as we can observe in figure 2.7 reporting the absolute and relative flow field induced on the horizontal plane by a counterclockwise rotation of the eye globe.

Notice that in the region close to the lens, the velocity is in the opposite direction with respect to the motion of the wall. Moreover a less intense velocity field is also generated in the posterior region of the vitreous chamber having velocity vectors again directed in the opposite sense with respect to the wall velocity. The three-dimensional character of the flow may play an important role in producing vitreous mixing. The velocity distribution at the wall suggests that, if the non-spherical character of the domain is accounted for, also the shear stress at the wall may be significantly modified with respect

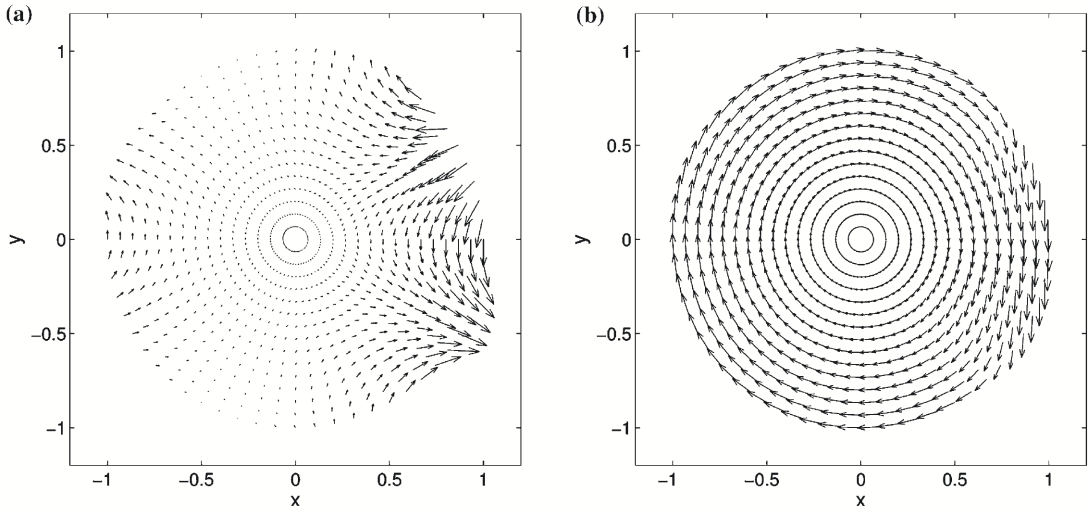

Figure 2.7: Absolute (a) and relative (b) velocity fields on the horizontal plane, induced by a counterclockwise rotation of the eye globe.

to the perfectly spherical case. Moreover an indication of the possible occurrence of boundary layer separation can be simply obtained considering the two-dimensional flow on the equatorial plane orthogonal to the axis of rotation produced by an impulsive rotation. In the case of low viscosity vitreous, eye rotation with amplitude larger than about 15° are expected to produce boundary layer separation with consequent formation of a wake region within the eye globe.

Chapter 3

Principles of PIV

PIV is a visualization techniques for quantitative measurement in flow fields. Features of PIV make it unique as it allows to capture the whole velocity field instantaneously without interfering with it. The PIV technique, in fact, works non intrusively. This allows the application of PIV even in high speed flows with shocks or in boundary layers close to the wall, where the flow may be disturbed by the presence of probes.

Moreover PIV can be used to indirectly obtain information on other important physical quantities of a fluid such as its density, temperature, concentration or test optical instrument performances. Although the theoretical principle of this technique has already been known for a long time, only recent technical progress in optics, lasers, video and computers makes it suitable to be employed in measurements of complex flow fields.

PIV supplies indirect velocity measurements of tracer particles seeded inside the observed fluid, by means of comparison of position that particles assume in photo images taken at subsequent time instant. The experimental set-up of a PIV system typically consists of several sub-systems.

An accurate system of illumination is required to light a plane inside the flow fields at least twice within a short time interval. Illumination is usually provided by a laser that produces light beams, which are appropriately adapted by a system of lens and glasses to be a light plane. The light scattered by the particles is recorded either on a single frame or on a sequence of frames by one or more cameras. All components (motor, laser, camera) of PIV system must be synchronized.

The displacement of the particle images between the light pulses has to be determined through evaluation of the PIV recordings. In order to elaborate images to obtain required information P.I.V. technique avails of sophisticated processing algorithms. PIV images are divided in small subareas called "interrogation areas", supposed to have homogeneous velocity. For each interroga-

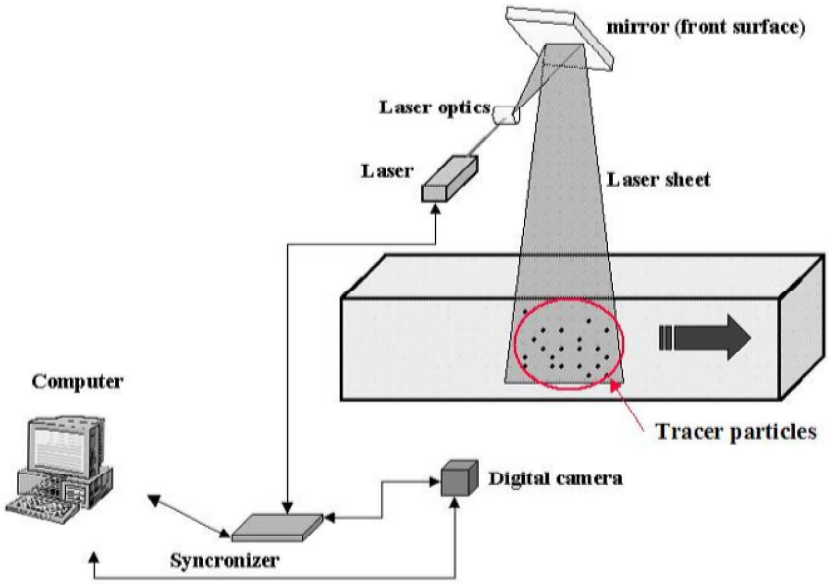

Figure 3.1: Main components of a PIV experimental set-up

tion area variation of particles position between subsequent images is observed by means of statistical methods (auto- and cross- correlation). The projection of the vector of the local flow velocity into the plane of the light sheet is calculated taking in account the time delay between the two images. The great amount of data needs a powerful post-processing system.

Features which mainly influence the quality of PIV acquisitions are:

- **Illumination** A compromise between the dimension of particles (big particles better scatter light) end their faithfully reproduction of flow field (small particles less interfere with the flow) has to be found. This problem is quite significant for application in gas flows while in liquid flows larger particles can usually be accepted. Thus, light sources of lower peak power can be used here.

- **Duration of illumination pulse** The duration of the illumination light pulse must be short enough that the motion of the particles is "frozen"during the pulse exposure to avoid blurring of image ("no streaks").

- **Time delay between illumination pulses Δt:** It must be long enough to be able to determine the displacement between the images of tracer particles with sufficient resolution and short enough to avoid particles

with an out-of-plane velocity component to leave the light sheet between subsequent illumination.

- **Distribution of tracer particles in the flow**. For qualitative flow visualization certain areas of the flow are made visible by marking a stream tube in the flow with tracer particles (smoke, dye). For PIV the situation is different: a homogeneous distribution of medium density is desired for high quality PIV recordings in order the obtain optimal evaluation. No structures of the flow field can be detected on a P.I.V. recording of high quality.

- **Density of tracer particles on the PIV recording** It's a fundamental factor in clear understanding of images. Three different types of image density can be distinguished. If the particles concentration is low, the image of individual particle can be detected. This kind of images requires tracking methods for evaluation (Particle Tracking Velocimetry PTV). A medium image particles density is required in standard PIV technique. In this case the image of particles can be detected as well, however is no longer possible to identify image pairs by visual inspection of the recording, statistic method are used. If particles density is high is not even possible to detect single particles as they overlap forming speckles. This situation is called Laser Speckle Velocimetry (LSV).

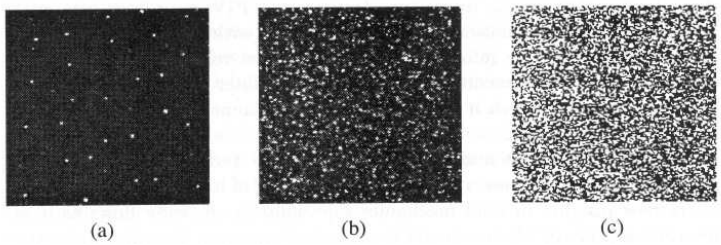

Figure 3.2: (a) Low density of tracer particles (P.T.V.) (b) Mean density of tracer particles (P.I.V.) (c) High density of tracer particles (L.S.V)

- **Size of interrogation areas** The size of the interrogation area must be small enough to neglect velocity gradients inside it. As the processing of an interrogation area furnish one velocity vector the number of the independent velocity vector is determined by the number of interrogation window.

3.1 PIV components set-up

Here we briefly explain the main features of the PIV components with particular reference to the experimental set-up used for the experiments below described.

3.1.1 Tracer particles

PIV technique is based on the direct determination of the two fundamental dimension of the velocity: length and time. On the other hand the velocity of fluid is indirectly determined being the particles velocity seeded in the flow to be measured. Therefore, fluid mechanical properties of the particles have to be accurately checked in order to avoid significant discrepancies between fluid and particles motion.

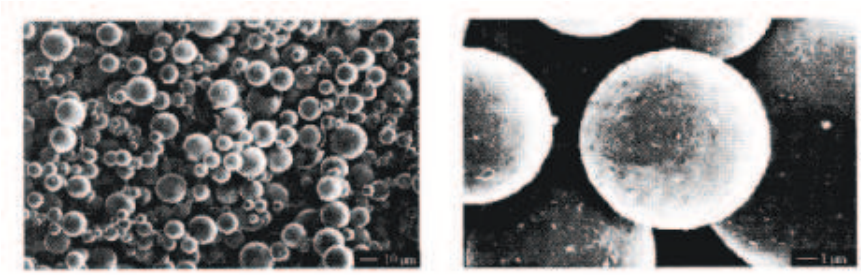

Figure 3.3: Images of hollow glass particles to an electronic microscope

A primary source error is in fact the different response to the gravitational forces in the cases the particles and the fluid have different density. In practical situation this can be avoided by a correct choice of particles. For a single spherical particle moving in a viscous fluid at very low Reynolds numbers the velocity due to gravitational forces is derived from the Stokes drag law:

$$U_g = d_p^2 \frac{(\rho_p - \rho)}{18\mu} g \tag{3.1.1}$$

where g is the acceleration due to gravity, μ is the dynamic viscosity of the fluid and d_p is the diameter of the particle. This velocity has to be added to the velocity of fluid which drags the particle, resulting delayed.

$$U_s = U_p - U = d_p^2 \frac{(\rho_p - \rho)}{18\mu} a \tag{3.1.2}$$

where U_p is the particle velocity, that typically follows an exponential law if particle density is greater than fluid density.

$$U_p(t) = U\left[1 - \exp\left(\frac{-t}{\tau_s}\right)\right] \tag{3.1.3}$$

being τ_s the relaxation time given by:

$$\tau_s = d_p^2 \frac{\rho_p}{18\mu} \tag{3.1.4}$$

τ_s is a convenient measure for the adaptation of the particles to the fluid velocity even in the cases in which Stokes law does not apply.

When applying PIV to liquid flows the problem of finding particles with matching densities are usually not severe, solid particles with adequate mechanical properties can often be found. Difficulties arise in applications in gases, taking in account that the particle diameter should be not too small to allow a good light scattering. Since the intensity and the contrast of the PIV recording are proportional to the light scattered by particles required quality of images can be reached by acting on laser light power or optical properties of tracer particles. Light scattered by a particle depends on its refractive indices, its size, shape and orientation. The light intensity as can be seen from all Mie scattering diagrams, spread in all directions. When a large number of particles are affected by light multiscattering effect magnify the local scattering.

3.1.2 Laser

Lasers are devices able to produce an intense beam by amplifying light through a process called stimulated emission. Lasers are widely used in PIV application, as light sources, for their ability to emit monochromatic[1], spatially coherent and collimated[2] light with high energy density, easily converted in light sheets for illuminating and recording particles image without chromatic aberrations.

A fundamental understanding of laser physics is based on the basic relationship between electrons and the nucleus in an atomic model. The current model of stimulated emission is described by quantum physics, which defines different energy levels of electrons while revolving around the nucleus in different levels of orbit. In this model, a stable electron is a normal state makes a transition to a higher but unstable energy level by absorbing a photon whose

[1] Radiation emitted by lasers has a well defined frequency coincident whit the characteristic frequency of atoms of the active substance present in the laser.

[2] The light either is emitted in a narrow, low-divergence beam, or can be converted into one with the help of optical components such as lenses

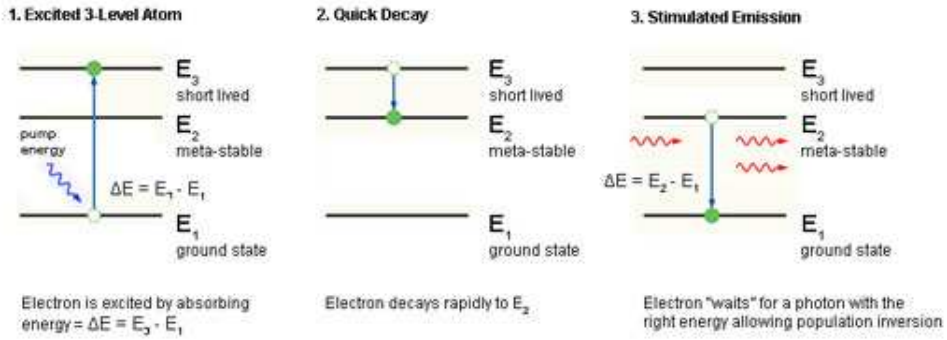

Figure 3.4: Typical scheme of working of a three level laser

frequency is $\nu = \Delta E/h$, where ΔE is the difference of energy between the two levels and h is the Plank constant. This unstable electron with high energy ultimately may return spontaneously to the original stable level. Alternately, this emission can be induced by a forced interaction between one photon and the unstable electron in order to release a new photon (stimulated emission), which is the basis of laser energy.

A laser primarily consists of three main components:

1. an **active medium** that is the component where the laser radiation is generated;

2. a **stimulation (excitation) mechanism** that is the power source of the laser pump;

3. an **optical chamber** that provides to amplify and collimate the light.

The function of **active medium** is to supply a source of stimulated atoms, molecules, and ions. It is a material of controlled purity, size, concentration, and shape, which amplifies the beam through the process of stimulated emission. It may be in a solid, gaseous or liquid state. The active medium absorbs pump energy, which raises some electrons into higher-energy ("excited") quantum states. Particles can interact with light both by absorbing photons or by emitting photons. If the emission is stimulated the photon is emitted in the same direction as the light that is passing by. When the number of particles in an excited state exceeds the number of particles in some lower-energy state, population inversion is achieved and the amount of stimulated emission due to light that passes through is larger than the amount of absorption. Hence, the light is amplified. Different types of lasers are named according to what is used as an active medium. Lasers, with a solid state of active medium are the

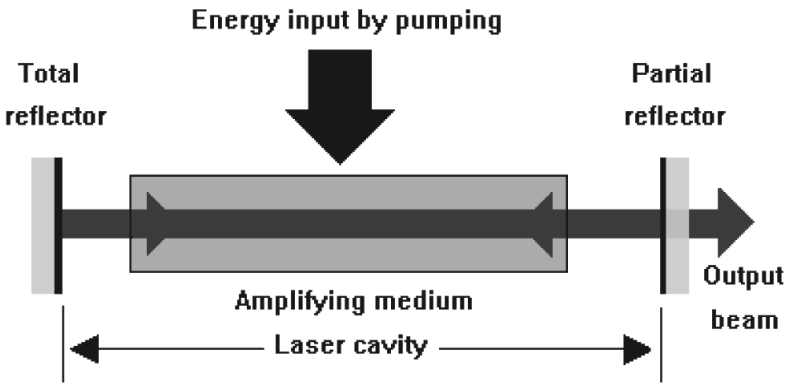

Figure 3.5: Scheme of main components of a laser

Nd:YAG (Yttrium-Alluminium-Garnet crystal with incorporated Nd^{3+} ions), ruby, and diode lasers. Lasers using a gaseous active medium are the CO_2, argon, and helium-neon lasers. The helium-neon laser is used as an aiming beam in lasers with an invisible[3] beam (as in the CO_2 lasers) in order to create a visible beam. A laser with a liquid active medium uses organic dye. The activation status of the laser medium is operated by the operation mode of the laser device. Three operational modes are currently available. In the *continuous mode*, the active medium is kept in a stimulated mode, which provides constant and stable energy. In the *pulsed mode*, the active medium is intermittently activated for a very short time. However, a much higher maximum of instantaneous energy is delivered with pulses compared to the continuous mode in which average power output is greater. In *Q-switched mode*, very short pulses of the laser are produced in a controlled manner, which is obtained with a shutter positioned between the active medium and the partial reflector. In this mode, pulses range from 10 ns to 10 μs.

The second component devoted to supply the energy required for the amplification is called **pump**. It provides the energy that is used to activate the medium. The energy is supplied as an electrical current or as light at a different wavelength or have chemical form. Such light may be provided by a flash lamp or perhaps another laser.

The **optical chamber** contains the active medium and it is used to direct the output and also to provide feedback from amplification and collimation. The cavity consists in reflective surfaces so that the light passes through the active medium more than once before it is emitted from the output aperture or lost to diffraction or absorption. As light circulates through the cavity, passing

[3]Not all the electromagnetic radiations be able to emit with a laser are in the visible spectrum: hence there are infrared lasers, ultraviolet lasers, X-ray lasers, etc.

through the active medium, if the amplification in the medium is stronger than the resonator losses, the power of the circulating light can rise exponentially. The result is the generation of stationary waves under the condition:

$$L = \frac{m\lambda}{2n} \tag{3.1.5}$$

in which n is the refractive index, m is an integer number and L the resonator length. Since the frequency ν, according to the transition $\nu h = E_2 - E_1$ does not correspond to an exactly one wavelength, but rather to a spectrum of a certain band width $\Delta\nu$ which can be filled by different wavelengths or frequencies.Therefore, the resonator can oscillate in many axial modes with distinct frequencies.

Besides these major components of the laser, it must contain a cooling system, a delivery system from the laser to the operative field, a control unit in order to operate power output (frequency and duration of exposure and spot size), and a remote control. Delivery system are important in the selection of the laser. They can be articulated arm (for the CO_2 laser), optical fibers (for near-infrared and visible lasers) or a connection between laser and microscope.

3.1.3 Light sheet optics

The domain of measurement must be lighted with a thin sheet of light. The essential element for the generation of a light sheet is a cylindrical lens. When using lasers with a sufficiently small beam diameter and divergence one cylindrical lens can be sufficient in order to create a light sheet of appropriate characteristics. With Nd:YAG lasers a combination of cylindrical and spherical lenses are usually required. As usual combination consists in a diverging lens, used to refract the light beam in a fan shape, associated with a converging lens devoted to focus the light to an appropriate thickness. A third lens can be inserted in order to generate a light sheet of a constant height.

An adequate choice of the optic system prevent the formation of focal lines, to be avoided especially with high power lasers as the air close to the focal point will be ionized and the dust particles will be burned if the area in the vicinity of the line is not covered or evacuated. Sometimes the experimental set-up requires the presence of other optical device such as glasses or optical groups.

3.1.4 Digital camera

The recent progress in digital imaging have provided an attractive alternative to the photographic methods of PIV recording. Modern digital cameras have substituted traditional cameras because of many advantages they present,

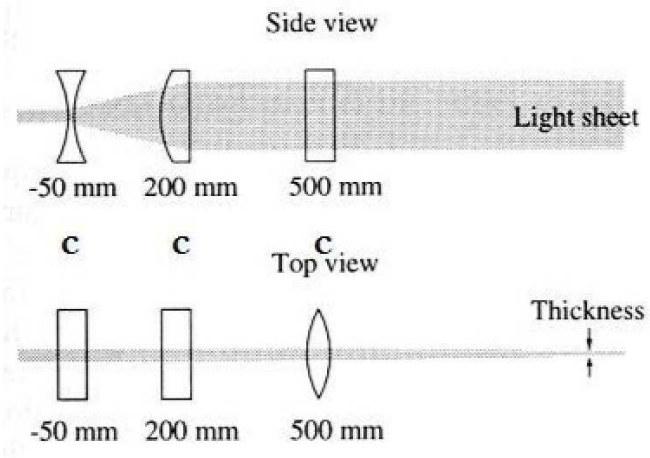

Figure 3.6: Example of combination of lens for laser beam adapting to a sheet

starting from the immediate availability of images as well as a complete avoidance of photochemical processing. Focusing phase, exposure regulation and choice of framing are the same in both type of camera. Image in traditional camera was fixed on photo film while in digital camera it is stored on a Charge Couple Device (CCD). The CCD is an electronic sensor that can convert light into electrical charge. Today there is a variety of electronic image sensors available. They are arranged either in form of a line or in a rectangular array. The individual CCD element in the sensor is the pixel whose dimension are approximately 10x10 μm^2. The CCD is built on a semiconducting substrate (typically silicon) with metal conductors on the surface, an insulating oxide layer, an n-layer (anode) and a p-layer (cathode) below that. When a photon of proper wavelength enters the p-n junction of the semiconductor an electronic hole pair is generated. While the hole is absorbed by the p layer, the electron migrates following the gradient of electric field toward its minimum where it is stored. Electrons continue to accumulate for the duration of the pixel exposure to light. A pattern of points with different light intensity is created in this way, this constitute the digitalized image, saved in camera memory or transferred to a computer. Pixel storage capability is limited, when the number of electrons exceeds the maximum , the additional electrons migrate to the neighboring pixels (overexposure) leading to image blooming.

3.2 PIV recording techniques

Different approaches to PIV recording are possible:

- methods which capture the illuminated flow on a single frame (**single-frame/single-multi-exposure**);
- methods which provide a single illuminated image for each illumination pulse(**multi-frame/ single-multi exposure**).

The principal distinction between the two branches is that the single-frame method, without any additional effort, does not retain information on their temporal order of the illumination pulse giving rise to a directional ambiguity in the recovered displacement vector. Using this approach on the same image we find the marks left by the particles during one or more period of exposure. For clarity in the fig. (3.7) examples of single-frame acquisition are reported:

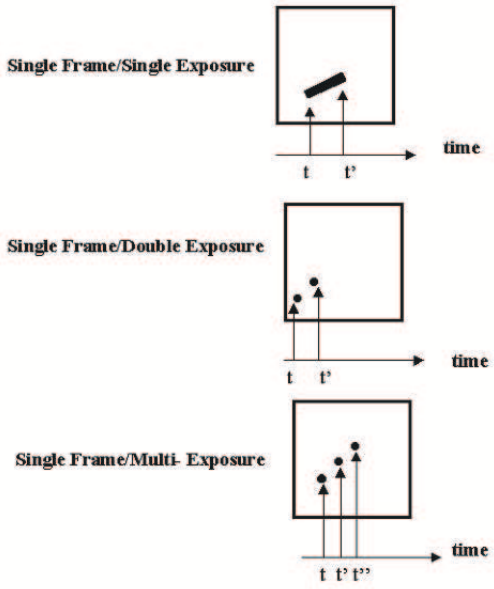

Figure 3.7: Single frame technique of recording

Multi-frame recording preserves the temporal order of the particles images. Particles position in subsequent instants is registered on different image.

The choice of the recording method to be employed in recordings is strictly related to the techniques used in the subsequent phase of image evaluation. The most frequent approach consists in registration of images separated by a time interval (Δt) called "pulse separation". This parameter is decided

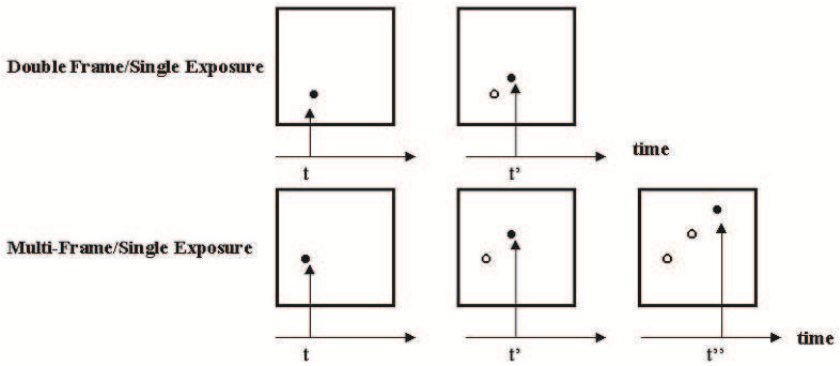

Figure 3.8: Multi frame technique of recording

by the operator according to the magnitude of velocity fields, the domain of measurement dimensions and the camera spatial resolution. A correct value of pulse separation avoid that particles leave their interrogation window and allow to easily identify their displacement.

Sampling frequency of images depends exclusively on the technical characteristics of camera and laser.

3.3 Image evaluation methods

In order to extract the displacement between two pattern of particle images some sort of interrogation scheme is required. For the image evaluation statistical approach is used: the measurement domain is divided in interrogation areas, each resulting vector represents the displacement of the group of particles contained inside the interrogation window supposed to have an homogeneous behaviour.

3.3.1 Pre processing operations

Before proceed to the image evaluation one has to decide the dimension of the grid superimposed to the images and the number of grid's nodes on which velocity vectors have to be referred.

Moreover, when the PIV measurements are performed over domains smaller than the entire image plane, or with irregular shape (i.e. the recordings shape is in this case rectangular), one has to superimpose to the images a mask consisting in a black figure covering the parts of the image out of the measurement domain, before starting in images processing.

3.3.2 Particle Image Locations

A single exposure recording consists of a random distribution of particle images, which correspond to the pattern of N tracer particles inside the flow. At the time instant t the statistical distribution of particles can be described by the random variable:

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \\ . \\ . \\ . \\ \mathbf{X_N} \end{pmatrix} \tag{3.3.1}$$

where the vector:

$$\mathbf{X_i} = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} \tag{3.3.2}$$

are the coordinate of the position of a tracer particle in a $3N$-dimensional space. On the image plane the coordinate of the image position vector of the particle i at time t are given by:

$$\mathbf{x_i} = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \tag{3.3.3}$$

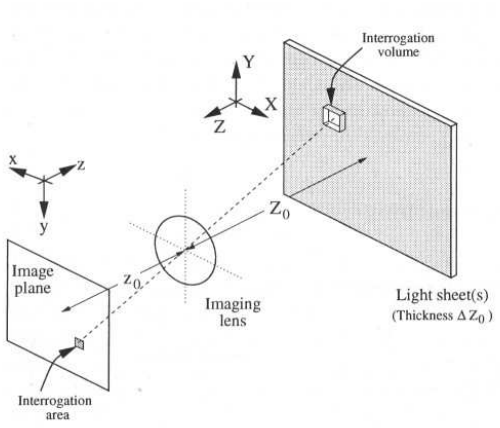

Figure 3.9: Schematic representation of geometric imaging

For simplicity we assume that the particle position coordinates and the image position coordinates are related by a constant magnification factor M

such that:

$$X_i = \frac{x_i}{M} \tag{3.3.4}$$

$$Y_i = \frac{y_i}{M} \tag{3.3.5}$$

3.3.3 Image Intensity Field

The image is interpreted like the result of the convolution of the geometric image and the impulse response of the imaging system. For infinite small particles and perfectly aberration-free well focused lenses the amplitude of the response can mathematically be described by the square of the first order Bessel function the so-called *Airy function* A good approximation of this function for a real system of lenses is represented by a Gaussian function:

$$\tau = K \exp\left(\frac{-8|x|^2}{d_r^2}\right) \tag{3.3.6}$$

where:

$$K = \frac{8\tau_0}{\pi d_r^2} \tag{3.3.7}$$

The convolution product of $\tau(\mathbf{x})$ with the geometric image of the tracer particle at the position $\mathbf{x}_i$ therefore describes the image of a single particle located at position $\mathbf{X}_i$. As the distribution of the real point is discrete, to describe the geometric image of each particle a Dirac delta-function shifted to position $\mathbf{x}_i$ is used. Thus, the image intensity field of the first exposure may be expressed by:

$$I(x, \Gamma) = \tau(x) \sum_{i=1}^{N} V_0(X_i)\delta(x - x_i) \tag{3.3.8}$$

Because of $\tau(x)$ is a real function:

$$\tau(x - x_i) = \tau(x) * \delta(x - x_i) \tag{3.3.9}$$

the 3.3.8 can be rewritten such that:

$$I(x, \Gamma) = \sum_{i=1}^{N} V_0(X_i)\tau(x - x_i) \tag{3.3.10}$$

being $V_0(\mathbf{X}_i)$ the transfer function giving the light energy of the image of an individual particle i inside the interrogation volume V_i and its conversion into an electronic signal or optical transmissivity. $\tau(\mathbf{x})$ is considered to be

identical for every particle position. The visibility of a particle depends on many parameters as, for example, the scattering properties of the particle, the light intensity at the particle position, the sensitivity of the recording optics and the sensor or film at the corresponding image position.$V_0(\mathbf{X})$ just describes the shape, extension, and location of the actual interrogation volume:

$$V_0(X) = W_0(X,Y)I_0(Z) \tag{3.3.11}$$

where $I_0(Z)$ is the intensity profile of the laser light sheet in the Z direction and $W_0(X,Y)$ is the projection of the interrogation window on the light sheet.

3.3.4 Correlation of PIV recording

The main mathematical tool in evaluating PIV images is the correlation. The auto-correlation (i.e. the degree of similarity of the image with itself), of the single-frame/single-exposure intensity field can be described by the auto-correlation function, derived by associating to each point a numerical value corresponding to its light intensity:

$$R_I(s,\Gamma) = \langle I(x,\Gamma)I(x+s,\Gamma)\rangle = \frac{1}{a_I}\int_{a_I}\sum_{i=1}^{N} V_0(X_i)\tau(x-x_i)\sum_{j=1}^{N} V_0(X_j)\tau(x-x_j)dx \tag{3.3.12}$$

where a_I is the size of the interrogation area and s is the displacement vector on the correlation plane. The terms $i = j$ represent the correlation of each particle with itself while the terms $i \neq j$ represent the correlation of images of different particle and therefore constitute a randomly distributed noise in the correlation plane.

The auto-correlation function has a primary peak in correspondence to the s=0 displacement (self-correlation) and secondary peaks. One of these corresponds to the effective displacement of the particle.

Taking in account the multiple-frame/single-exposure PIV recordings, the images are evaluated by locally cross-correlation. We consider, now two recordings made at a time instant t and $t' = t + \Delta$. In the following, a constant displacement $\mathbf{D}$ of all particles inside the interrogation volume is assumed, so that the particle locations during the second exposure at time t' are given by:

$$X_i' = X_i + D = \begin{pmatrix} X_i + D_x \\ Y_i + D_y \\ Z_i + D_z \end{pmatrix} \tag{3.3.13}$$

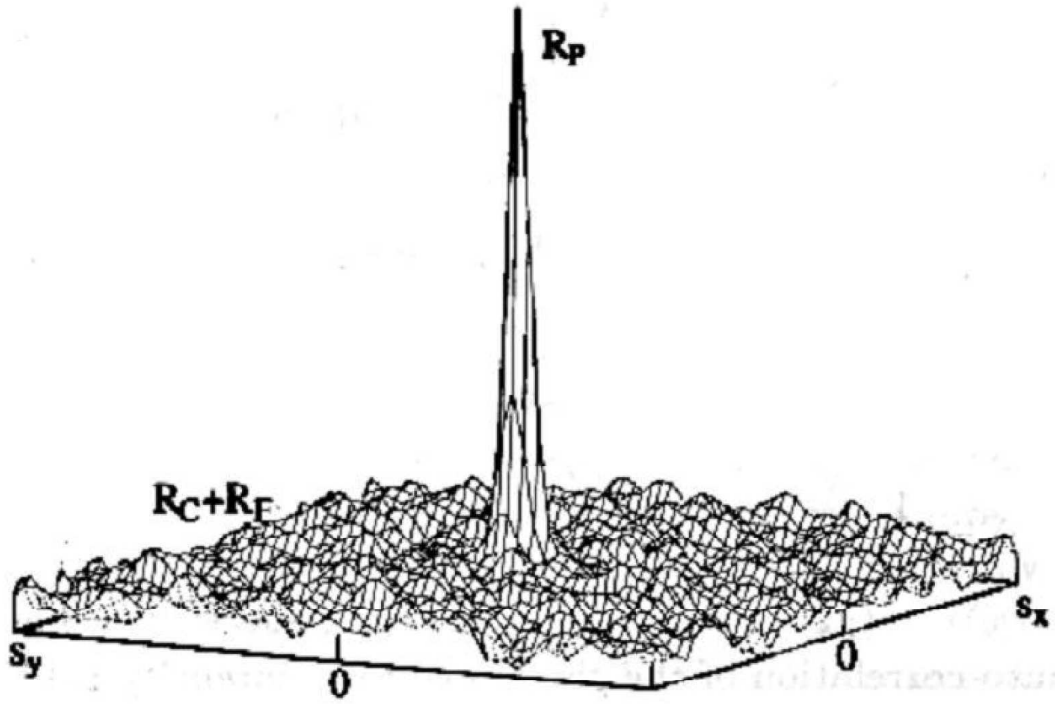

Figure 3.10: Composition of peaks in the auto-correlation function, according to Adrian: $R_I(s,\Gamma) = R_C(s,\Gamma) + R_F(s,\Gamma) + R_P(s,\Gamma)$, where $R_C(s,\Gamma)$ is the mean intensity I convolution, $R_F(s,\Gamma)$ is representative of the noise components for the terms with $i \neq j$ and $R_P(s,\Gamma)$ is the auto-correlation peak positioned in (0,0) in the correlation plane

We furthermore assume that the particle image displacements are given by:

$$d = \begin{pmatrix} MD_x \\ MD_y \end{pmatrix} \tag{3.3.14}$$

If we first consider identical light sheet and windowing characteristics, the cross-correlation function of two interrogation areas can be written as:

$$R_{II}(s, \Gamma, D) = \frac{1}{a_I} \sum_{i,j} V_0(X_i) V_0(X_j + D) \int_{a_I} \tau(x - x_i) \tau(x - x_j + s - d) dx \tag{3.3.15}$$

Hence, for a given distribution of particles inside the flow, the displacement correlation peak reaches a maximum for $\mathbf{s} = \mathbf{d}$.

Therefore the location of this maximum yields the average in-plane displacement, and thus, as we know the time interval between the two images, we are able to evaluate the U and V components of the velocity inside the flow.

In the case of a unique image containing the particle position in subsequent time instant (single-frame/multiple-exposure), we have to evaluate cross-correlation of intensity of the first set of particles with itself immortalized at the successive instant.

This technique has the disadvantage of ambiguity in displacement direction. Moreover if the time interval between two successive recording is low, superimposition of particles mark are possible, making image interpretation hard to do.

3.4 Post processing of images

Investigations employing the PIV technique usually result in a great number of images which must be further processed. Thus, it is quite obvious that a fast, reliable and fully automatic processing of the PIV data is essential. Before incoming to the final results PIV data follow a process consisting in the following steps.

- **Validation of the raw data.** After automatic evaluation of the PIV recordings, a visual inspection of images shows a certain number of incorrectly determined velocity vectors (so-called *outliers*). Vectors that have to be substituted have clearly anomalous characteristics respect to the neighboring vectors, like length or direction significantly different.

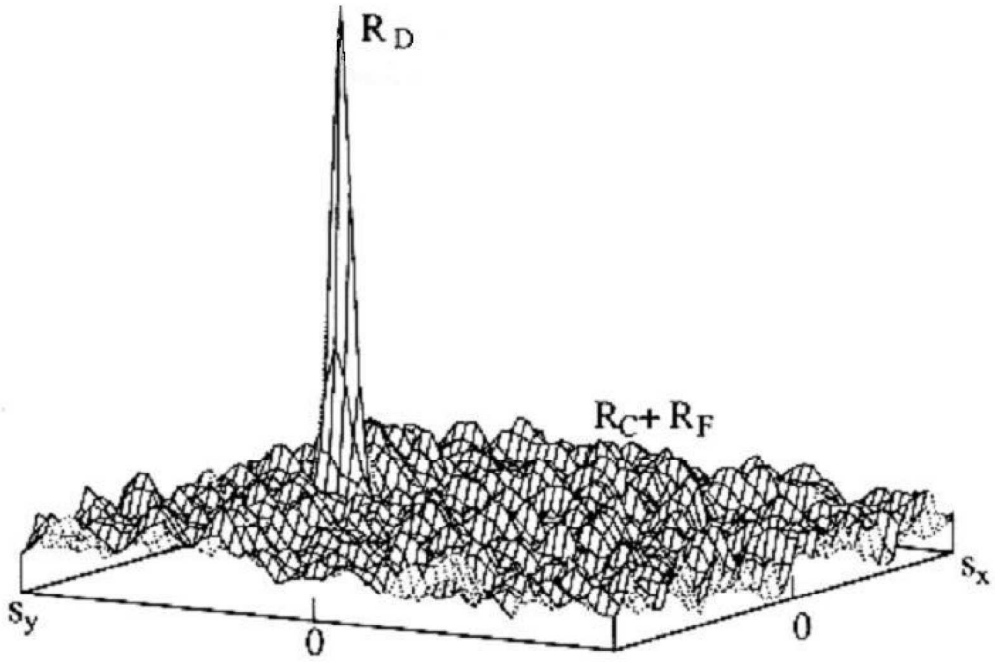

Figure 3.11: Composition of peaks in the cross-correlation function. where $R_D(s, \Gamma, D)$ represents the correlation of the images with itself (terms with i=j)

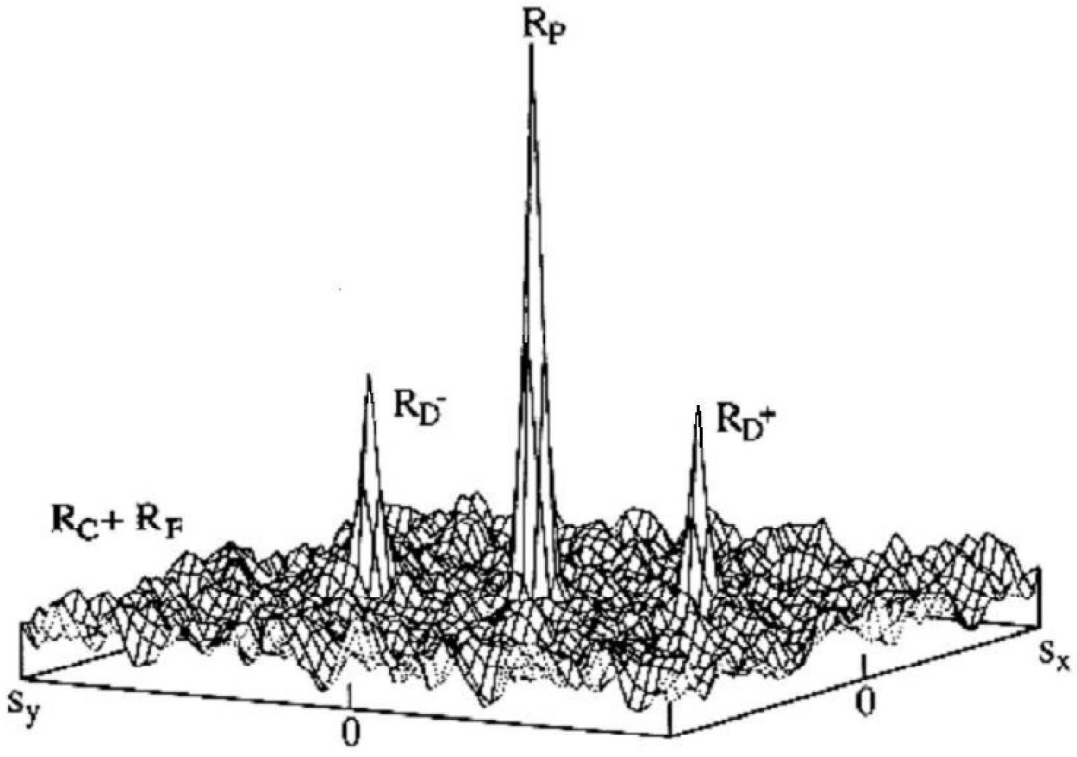

Figure 3.12: Auto-correlation function components for the single-frame/double-exposure mode

The presence of incorrect vectors inside the flow fields derive from error in evaluating the cross-correlation function. Scatters, fault in surfaces crossed by the beam of light, rumor generated from too many particles, velocity gradients are common sources of error. In order to detect these incorrect data many validation methods have been proposed in literature. Some of these methods are based on correlation function peaks, on knowledge of general pattern and values of flow field, or apply filters or mobile mean processes. The main characteristic these algorithms must have is to work automatically.

- **Replacement of the incorrect data.** For most post-processing algorithms (e.g. calculation of vector operators) complete data fields are required. Therefore, the gaps (data drop-outs) in the experimental data left by the previous procedure have to be filled with other data. This purpose can be achieved with procedures of bilinear interpolation using the vectors in the vicinity to the one that has been removed or weighted mean with the "Gaussian window" technique.

- **Data reduction.** PIV is a largely productive technique, as result of PIV evaluations one obtains thousands of data. It is quite difficult to inspect such a number of velocity vector maps and to describe their fluid mechanical features. Usually techniques like *averaging* (in order to extract the information about the mean flow and its fluctuations),*conditional sampling* (in order to distinguish between periodic and non periodic parts of the flow), and *vector field operators* (e.g. vorticity, divergence in order to detect structures in the flow) are applied.

- **Analysis of the information.** At present this is the most challenging task for the user of the PIV technique. PIV being the first technique to offer information about complete instantaneous velocity vector fields allows new insights in old and new problems of fluid mechanics. New tools for analysis such as *proper orthogonal decomposition* (POD) or *neural networks* are applied to PIV data.

- **Presentation and animation of the information.** A number of software packages—commercially available as well as in-house developed—are obtainable for the graphical presentation of the PIV field data. It is also very important to support the understanding of a human observer of the main features of the flow field. This can be done by *contour plotting*, *color coding*, etc. Animation of PIV data is very useful for a better understanding in the case of time series of PIV recordings or 3D data.

3.5 Flow fields analysis

The results of the previous elaborations are vector maps, on a regular grid, usually referred to a cartesian system of coordinate. Components of velocity vector are registered for each node of the grid. Sometimes commercial software used to elaborate PIV measurements supply further information like: spatial derivative of velocity, mean of data on all registered velocity fields, standard deviation and other statistical information. By elaborating velocity data is then possible evaluate shear stress, actions on boundary, pressure fields...

Chapter 4

Experimental set-up

In this chapter we give a description of the experimental set-up. The aim of the experiments is the measurements of the flow fields inside a magnified model of vitreous cavity. The eye globe has been modeled by means of two semi-spherical cavities of 40 mm radius (scaling about 4:1 the real dimensions of normal human eye), carved into two halves of a plastic cylinder of external radius equal to 120 mm. The cavity was filled with a 98 % pure glycerol solution, a Newtonian fluid with high viscosity. The cylinder was mounted on the shaft of a computer controlled motor which rotates according to prescribed time law. The rotation of the shaft was synchronized with a two-dimensional Particle Image Velocimetry (PIV) acquisition system that was employed to measure two-dimensional velocity fields on the equatorial plane of the model. For a description of the main feature of this technique the reader is referred to chapter 3; a thorough examination of this subject is given in [17].

4.1 Design of the experiments

4.1.1 Scaling effect

The experiments have been performed with a magnified model of the vitreous chamber; therefore, scale effects need to be properly accounted for. If we make equation 2.1.1, 2.1.2 and 2.1.3 dimensionless, scaling the variables (u by $U = \omega_p R$, being ω_p the peak angular velocity, t by T, where T a characteristic time scale of globe rotation, and r by R) like previously shown in 2.1.7 we found only one relevant dimensionless parameter, namely the Womersley number:

$$a = \sqrt{\frac{R^2}{\nu T}} \tag{4.1.1}$$

In the case of periodic rotations it is a natural choice to replace T by $2\pi/\omega$, where ω is the angular velocity, as the characteristic time scale. With this substitution a reduces to the form:

$$\alpha = \sqrt{\frac{R^2\omega}{\nu}} \tag{4.1.2}$$

as it appears in David et al (1998)[7]:

For the above reasons α is the parameter which has been conserved in the passage from the prototype to the model. Thus we impose:

$$\frac{R^2\omega}{\nu} = \frac{R_p^2\omega_p}{\nu_p} \tag{4.1.3}$$

where the subscript p indicates the variables referred to the prototype.

Once a particular eye movement has been selected equation (4.1.3) allows us to determine a relationship between T and T_p, i.e. the scaling of time, which reads:

$$T = \frac{R^2}{R_p^2}\frac{\nu_p}{\nu}T_p \tag{4.1.4}$$

4.2 The experimental models and the rotating system

4.2.1 Experimental model of vitreous chamber

The experiments described in the present work have been made with two different model of posterior chamber:

- a **spherical model**
- a **deformed sphere model** which takes in account the presence of the lens in the anterior part of the chamber

Both models consist in a cavity of 4.08 *cm* radius carved in a perspex cylinder (refractive index 1.48), with external diameter 12 cm and 18 cm high. The model is divided in two halves.

A hole on the top of the cylinder allow its filling with the experimental fluid seeded whit tracer particles. The cavity insulation is assured by two strips put around the model, while the contact surface of the two halves hare covered by vaseline.

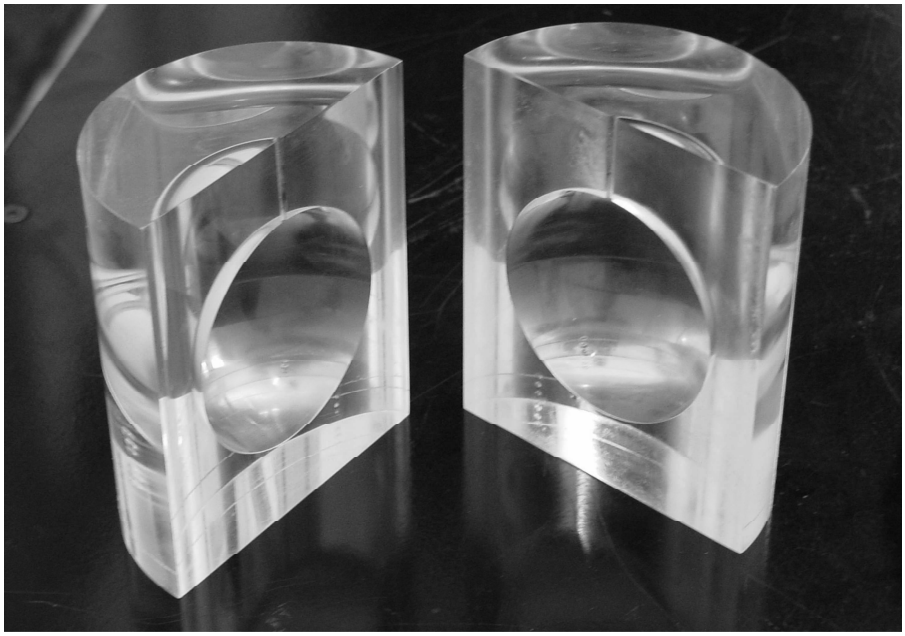

Figure 4.1: Picture of the spherical model

4.2.2 Experimental fluid

Interior cavity model has been filled with a 98 % pure glycerol solution. This liquid has newtonian behaviour and high viscosity. The solution relative density is almost 1,256 and its refractive index is 1,474. This value matches the index of refraction of the perspex used for the container, thus excluding any deformation in the images due to refraction of light rays when they cross the curved interface between two materials. Strong dependence of glycerol (i.e. the fluid used in experimentation) viscosity on the external temperature (see table 4.1) has taken in account by periodical measurement of liquid viscosity by a viscometer, during experiments, in order to continuously monitor its possible variations. Moreover each session of experiments have been made employing the time strictly necessary, in order to assure the same environment conditions avoiding viscosity fluctuations during the experiment.

	0^o C	20^o C	60^o C
Water	1.8	1.0	0.65
Glycerol	10000	1410	81

Table 4.1: Kinematic viscosity values related to external temperature for water and glycerol

Viscosity measurements were hold on with a sphere drop viscosimeter, model 356-0001 Haake, a picture of which is taken in figure 4.2.

Figure 4.2: Picture of the viscometer used to monitor viscosity values

4.2.3 Tracer particles

The tracer particles used during the experiments are hollow glass sphere with diameter variable between 2 and 20 μm and refractive index of 1.52. The particles density is about 1.15 g/cm^3 nearly matching the density of the working fluid, assuring the drag velocity of the particles to be negligible with respect to the motion of the fluid. Particles were seeded into the glycerol and carefully mixed to guarantee an uniform distribution inside the fluid before the fluid was inserted inside the cavity. The adequate quantity of tracer particles to be introduced inside the fluid was evaluated by means of the analysis of images taken in the initial test experiments.

To fill the model chamber a medical drip suitably modified was employed. The system consists in a glass container of 500 cm^3 capability, with a gum cap, on which a surgical needle and a small metal pipe have been inserted and connected with a gum duct (interior diameter 5 mm)ending with a needle, to be introduced into the cavity.

4.2.4 Model rotating system

The models were mounted on a support directly connected with the shaft of an electrical brushless motor (mod. $G404-550$ MOOG with T200 control) by means of four screws. On the top of the support we put a black disc directed to increase the contrast between the equatorial images of the model and the background. The motor allowed the eye model to rotate around its vertical axis, according to an assigned temporal law. The motor was governed by the computer, and allows oscillations from $\pm 5°$ to 20 Hz and from $\pm 50°$ to 1 Hz.

The engine can operate in:

- **Current (Torque) Mode:** requiring a torque couple in Volt (until 10 Volt);
- **Velocity Mode:** requiring the velocity law to follow;
- **Position Mode:** requiring a position law.

4.3 Feature of the PIV apparatus

In the figure 4.6 a scheme of the experimental apparatus is reported. Equatorial plane of the model rotating with a prescribed time law was enlighten with a pulse laser beam. At the same time of laser pulse a digital camera mounted on the model rotation axis at about 1 m over its equatorial plane takes couples of picture to be processed with cross-correlation analysis as described in chapter 3.

4.3.1 Laser and optics

Measurements domain enlightenment was provided by a couple of lasers. During the PIV experimentation we used two 30 Hz Nd:Yag lasers called "Gemini YAG", individually manageable through two different consoles. The light beam has a wavelength of 532 nm (green) and was converted into a sheet of 0.5 mm thickness and with planar dimensions sufficient to illuminate the area of interest.

In order to adapt the beam to experimental requirement a system of two lens and two glasses with a 45° slope has been used. Both lens are cylindrical and have rectangular shape. The characteristics of the lens are listed below:

- plane-convex lens: focal f=700 mm
- plane-concave lens: focal f=70 mm

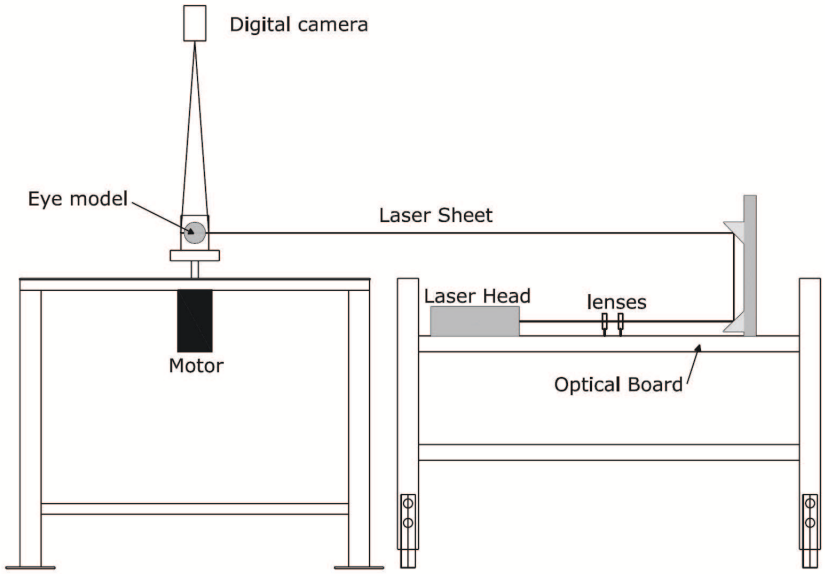

Figure 4.3: Sketch of the lateral view of the experimental apparatus

Figure 4.4: Picture of Gemini YAG lasers used in experiments

Laser optics needs to be placed with care in a stable position. This aim was reached by the use of an optical table consisting a aluminium frame of 600 x 1200 mm and 51.5 mm thickness, covered with a steel layer. On the surface of the optical table a square pattern of hole spaced out of 25 mm allows to place the optical components with screw, trolley or magnetic connections.

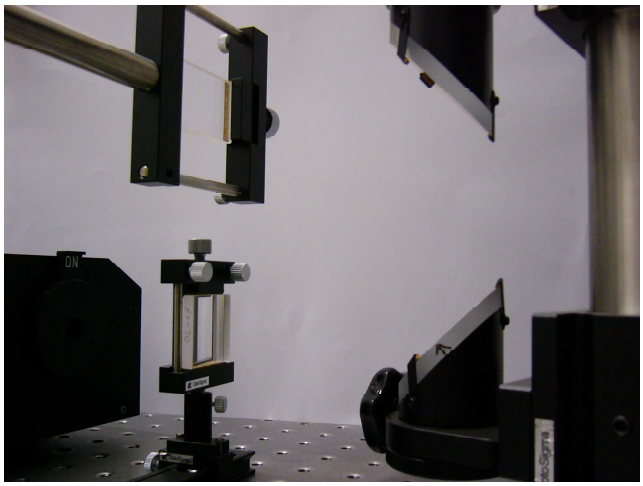

Figure 4.5: Experimental set-up of laser optics

4.3.2 Digital camera

The images used in PIV measurements were taken with the digital camera SONY,SharpVISION 1300DE. Camera was positioned on the model axis 100 cm over the measurement plane. The characteristics of CCD sensor Sony ICX085AL mounted on digital camera are listed in the table 4.2:

Pixel dimension	6.7×6.7 $[\mu m]$
Sensor area	$10.0(H) \times 8.7(V)$ $[mm]$
Sensor dimension	$11[mm]$ diagonale
Spatial resolution	$1300(H) \times 1030(V)$
Read-out noise	7-8 e^-
Sensor capability	16000 e^-
Spectral response	290 - 1000 $[nm]$
Dynamic range	12 bit
Antiblooming	¿ 1000
Dark current	¡ 0.1 $[e^-/pixel/s]$
Sensitivity	4350 $[e^-/lux/mm^2/s]$

Table 4.2: CCD sensor characteristics

4.3.3 Synchronization devices

A perfect synchronization among the different component of a PIV apparatus is fundamental. Connection of components are realized with cables connected

to two terminals. These devices constitute an expansion of gates of computer controlling the components.

Figure 4.6: Terminal used in connecting PIV apparatus components

4.4 Image acquisition

The recording of images have been taken with a multi-frame single-exposure approach. The sampling rate employed in the experimental measurements was set such to obtain 40 vector fields within a period of sinusoidal movement and 20 flow fields in a single duration of a saccadic movement, in the case of experiments performed with a spherical model, while we obtain 100 vector fields within a sine period for the lowest frequency experiments and 50 flow fields in a single sine period for the highest frequency runs with the deformed sphere model.

4.5 Post Processing

The results of the previous elaborations are vector maps, on a regular grid, usually referred to a cartesian system of coordinate. Components of velocity vector are registered for each node of the grid. The PIV setting used to analyze the images in terms of cross correlation yielded to a spatial resolution of about one velocity vector per 1.5 mm^2. Each velocity field $(u(x,y,t);v(x,y,t))$ obtained by PIV recordings was measured 60 times. An ensemble average $(U(x,y,t);V(x,y,t))$ was then determined and the corresponding root mean square values u_{RMS} and v_{RMS} evaluated. Averaging of the velocity fields permit to obtain feasible measured velocity fields by means of the random errors correction.

4.6 Creation of a polar grid

Vector fields were transformed into polar coordinates, which is a more natural choice for the geometry of the investigated domain. The polar grid reproduces the mask domain. The grid consist in evenly spaced radial sectors with an unevenly spacing along the radius (i.e. Δr is not a constant) in order to assure that all the elements have the same area. The velocity components resulting from PIV processing of the images have been arranged on the polar grid by means of interpolation. Each interpolated vector has been computed using a number of 12-20 vectors on the original cartesian grid. Notice that the presence of the mask produces vectors identically equal to zero in the areas external to the measurement domain. This could false the interpolation results near the external boundary. This effect was controlled introducing a great number of velocity vectors honoring the boundary condition:

$$U(R) = U_{boundary} \tag{4.6.1}$$

The interpolated vectors were obtained considering a little area S_s centered in the node x_b containing N_s velocity vectors. Different interpolation techniques were tested. Finally we adopted the method suggested in Agui and Jemenez (1987)[2] and Stuer and Blaser (2000) [24]. The interpolated vector in the generic node x_b is evaluates using the Adaptive Gaussian Window technique:

$$\mathbf{u}(x_b) = \frac{\sum_{i=1}^{N} \alpha_i u_i}{\sum_{i=1}^{N} \alpha_i} \tag{4.6.2}$$

where:

$$\alpha_i = \exp\left(\frac{-(x_b - x_i)^2}{H^2}\right) \tag{4.6.3}$$

and the dimension of the gaussian window H is:

$$H = 1.24\Delta_s \tag{4.6.4}$$

being $\Delta_s = \frac{S}{N}$ the dimension of the single areal element, ratio between the whole domain area and the number of interpolated vectors (remind that the polar grid had constructed with elements with the same extension).

According to Miller (1976) [15] the error on the estimate of the interpolated vectors has been evaluated by the Jackknife technique, resulting lower than that obtained in computing the averaged U and V. This result guarantees that the interpolation procedure do not introduce further errors on experimental data. Jackknife technique consists in evaluating N_s times the vectors $\hat{\mathbf{u}}_\mathbf{b}$

using time by time the algorithm 4.6.2 having neglected one of the N_s vectors contained inside the area S_s. This allow to compute the variance $\sigma(\mathbf{u}_b)$:

$$\sigma(\mathbf{u}_b) = \sqrt{\frac{N_s - 1}{N_s} \sum_{i=1}^{N_s} (\hat{\mathbf{u}}_b - \mathbf{u}_b)^2} \tag{4.6.5}$$

In spherical model simulations, in view of the axial symmetrical shape of the domain, we have performed an average of the circumferential velocity u_φ along φ. From the measured velocity profiles we have derived tangential stresses, by differentiation of the velocity using a finite least-squares scheme.

Chapter 5

Experiments description and results

In this chapter we describe the experiments brought to completion using a magnified model of vitreous chamber and their results. The experimentation was carried out with two different models of vitreous chamber. The first model has a spherical shape. The second model has a shape closely resembling to the real one, being deformed on one side in order to simulate the presence of the lens. Sinusoidal and saccadic time law were used to move the models around its axis. Velocity fields on the equatorial plane of the models were measured by means of the PIV technique.

5.1 Flow visualization

The choice of taking measurements of velocity fields on the equatorial plane orthogonal to the axis of rotation was justified by the assumption that the secondary flow generated during the model rotations can be neglected respect to the tangential flow. Moreover such a plane is a plane of symmetry of the motion; therefore, one expects the velocity vectors to lie on it. In order to evaluate the magnitude of secondary flow and to prove the accuracy of the assumption of purely axis-symmetrical flow we made qualitative flow visualizations by means on a colored dye inserted into the cavity during its motion. A blue dye was injected close to the wall just above the equatorial plane in the spherical model. The evolution of dye motion during the model rotations allowed us to visualize the secondary flow induced by the curvature of the wall. This secondary flow, observed both for periodic and saccadic movements, consists of two toroidal mirror vortices, located in the northern and southern hemispheres, respectively. In the upper hemisphere, particles close

to the equatorial plane move towards the globe center, then proceed upwards to descend back close to the curved wall. In the lower hemisphere, particles after moving towards the globe center, descend to the bottom of the cavity and then go up again close to the wall. All the experiments were recorded by a digital camera. Propagation of such a vortices was found of some importance only with repeated high frequency rotations in conditions far from the real characteristics of eye movements. The intensity of the secondary flow, however, is at least three to four orders of magnitude smaller than that of the circumferential flow in the case of periodic rotations and even smaller for saccadic movements. Thus, flow measurements on the equatorial plane of a rotating cavity can be considered a good representation of the motion even if they cannot perceive the out of plane components of the secondary flow, which can be neglected respect the circumferential flow. In figure 5.2 frames of vertical planes at different times are shown for a periodic experiment.

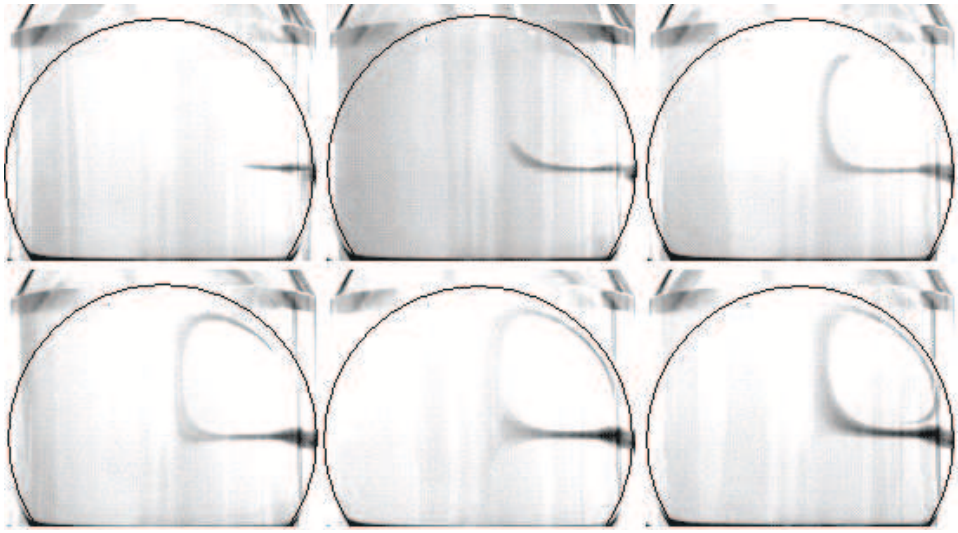

Figure 5.1: Snapshots of vertical planes, in false colors, at different times of the visualization in a periodic experiment. Time grows from top left to bottom right, the time interval between two successive frames is 8 s, which corresponds to 20 periods of the sinusoidal rotation ($A = 20°$, $f = 2.5Hz$).

5.2 A spherical model

A first set of experiments has been performed with a spherical model consisting in a cavity of 4.08 cm radius carved in a perspex cylinder (refractive index 1.48), with external diameter 12 cm and 18 cm high. For practical reasons connected with the model realization and cleaning after experiments the model has been built in two halves.

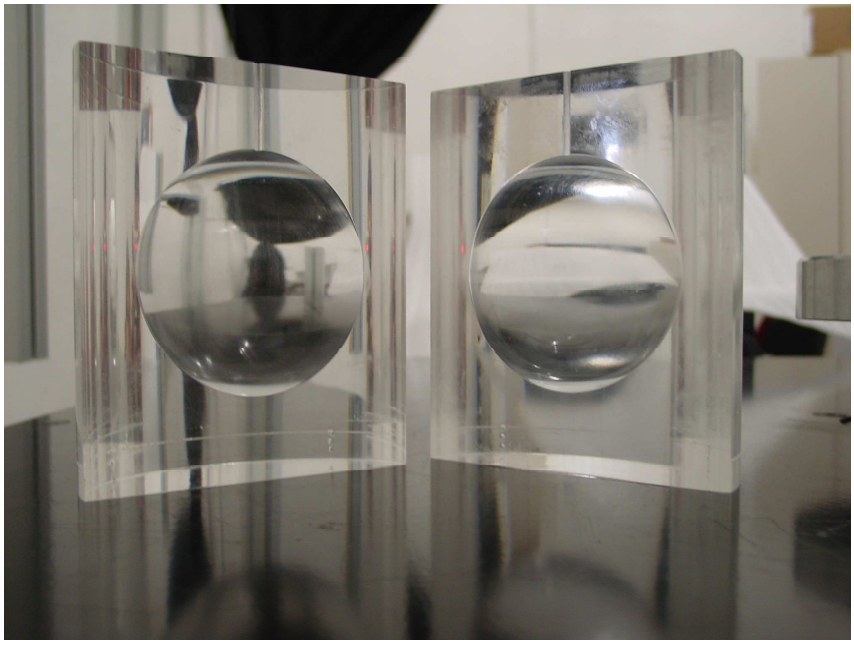

Figure 5.2: Picture of the spherical model

A thin pipe dug on the top of the cylinder allow its filling with the experimental fluid seeded with tracer particles. The two halves of the model were recomposed by two plastic and iron strips put around the cylinder and the contact surface of the two halves have been covered by vaseline to guarantee the cavity has been tight closed. The movement of the eye globe was simulated assigning either:

- a **sinusoidal time law**
- or a **polynomial temporal law**.

5.2.1 Periodic movements

The first set of experiments, performed with periodic rotations of the model, was aimed at validating the experimental apparatus under relatively simple

conditions and verifying the feasibility of theoretical results of David et al (1998)[7] for purely viscous fluids. Notice, moreover, that a sinusoidal rotation may be thought in a highly simplified way, as a sequence of saccadic movements with alternate directions and with a duration of half the period of rotations (see figure**??**). In table 5.1, the relevant parameters of the whole set of the periodic experiments performed are reported.

Exp. #	A [deg]	a	f [Hz]	ν [10^{-4}m^2/s]
sin-1	10	5.25	1.0	3.79
sin-2	20	3.99	1.0	6.58
sin-3	5	4.10	1.0	6.25
sin-4	10	2.77	0.5	6.79
sin-5	10	4.27	1.25	7.16
sin-6	30	3.94	1.0	6.73
sin-7	40	4.01	1.0	6.50
sin-8	10	5.07	1.5	6.11
sin-9	10	5.41	1.75	6.26
sin-10	10	3.75	0.75	5.58
sin-11	20	5.0	1.5	6.29
sin-12	10	5.52	2.0	6.87
sin-13	10	6.14	2.5	6.94
sin-14	5	3.79	1.0	7.33
sin-15	10	4.21	1.25	7.37
sin-16	20	5.80	2.5	7.76
sin-17	10	5.61	2.0	6.64
sin-18	10	4.66	1.5	7.22
sin-19	10	5.01	1.75	7.13
sin-20	20	4.71	1.5	7.08
sin-21	10	6.12	2.5	6.98

Table 5.1: Main parameters of the experiments performed with periodic rotations.

The sampling rate employed in the experimental measurements was set such to obtain 40 vector fields within a period. Each velocity field $(u(x, y, t), v(x, y, t))$ on the equatorial plane, was measured 60 times. Examples of some resulting velocity field are shown in figure 5.3 and 5.4.

Moreover, we have performed an average of the circumferential velocity u_φ along φ justified by the axial symmetrical shape of the domain. Figure 5.5 shows the experimental profiles of the averaged circumferential velocity along the radial direction plotted at different times for experiments sin-10 and sin-16.

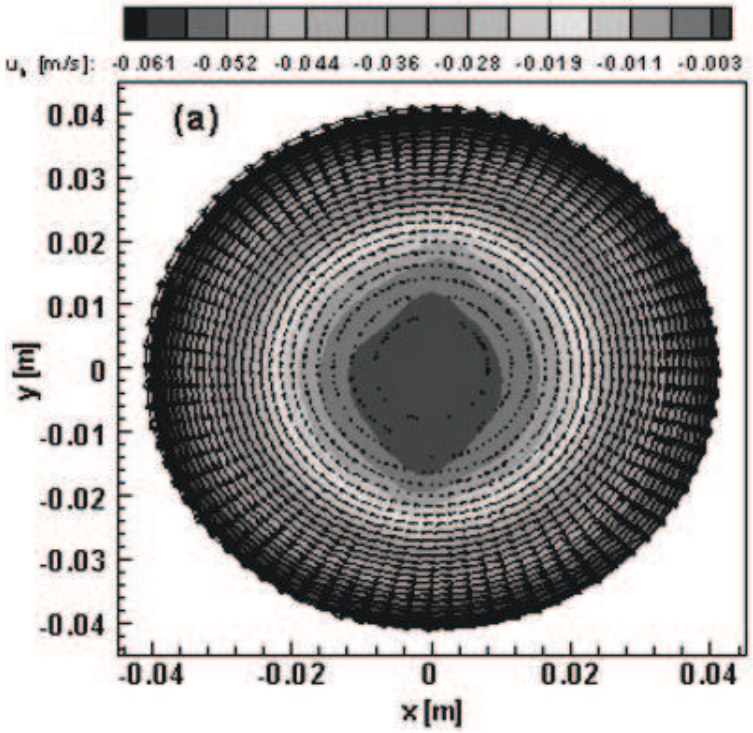

Figure 5.3: Example of two-dimensional interpolated velocity field of periodic experiment (A=10°, f=1.75 Hz) together with contours of the circumferential velocity: velocity field corresponding to the maximum peripheral velocity

5.2.2 Discussion of the results

The experimental results show that a significant phase lag characterizes the motion of fluid in the most internal regions with respect to the motion close to the wall. The motion of the fluid during a period can be split into three different phases. As we can see in figure 5.6, in the first phase (red lines) velocity profiles are in phase and the maximum velocity along the radius is located at the wall. In the second phase (green and blue lines) the boundary velocity inverted and the maximum velocity migrate toward the most internal regions of the fluid domain. The last phase (yellow lines) shows a sensitive phase lag between the fluid layers close to the wall and the internal ones characterized by circumferential velocity with opposite directions. Experimental observations agree with the theoretical results found by David et al (1998)[7]. It also appears that, for larger values of the Womersley number (experiment sin-16), the oscillating boundary layer at the wall is thick and the fluid particles in the most inner region are subjected to small amplitude oscillations. As far as the Womersley number decreases (experiment sin-10) the envelope of

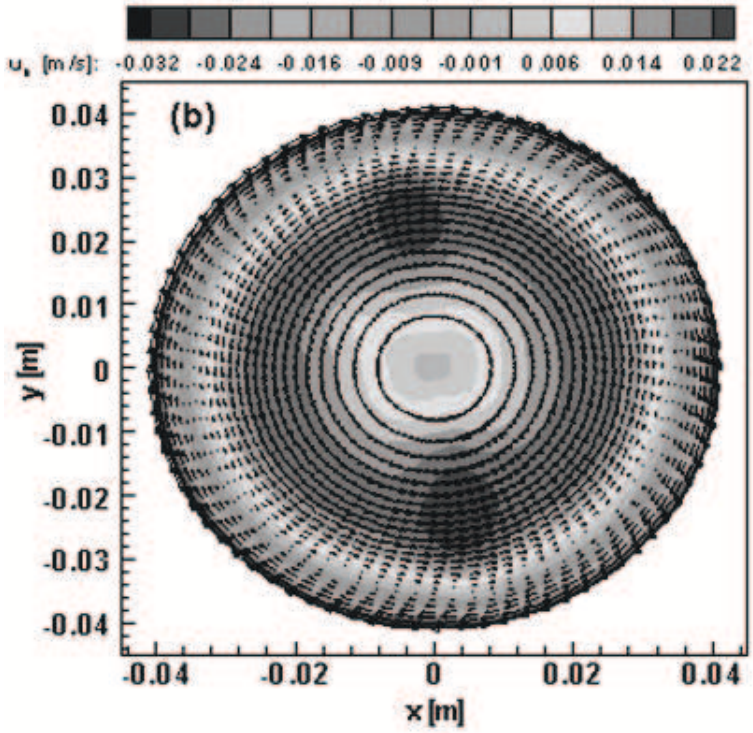

Figure 5.4: Example of two-dimensional interpolated velocity field of periodic experiment (A=10°, f=1.75 Hz) together with contours of the circumferential velocity: velocity field measured immediately after the inversion of the peripheral velocity

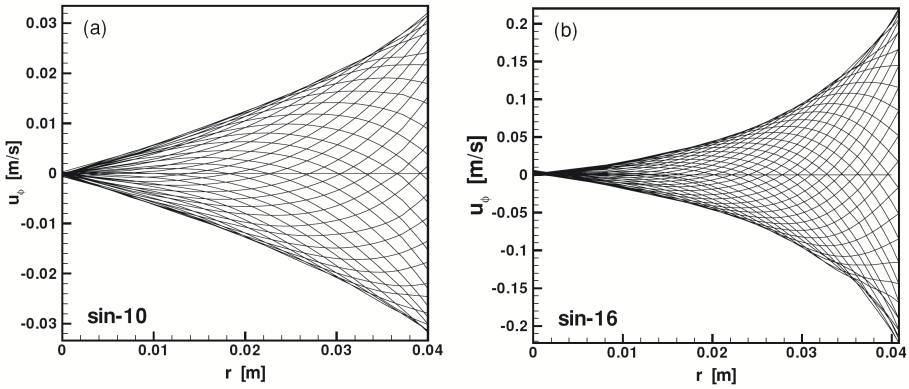

Figure 5.5: Experimental profiles of the averaged circumferential velocity u_φ along the radius at different times, with interval equal to 1/40 of the period. (a) Experiment sin-10; (b) Experiment sin-16.

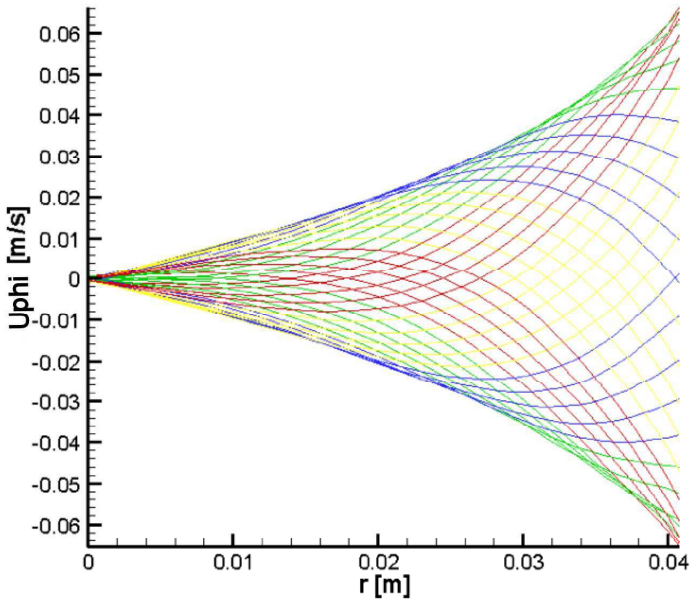

Figure 5.6: Picture of velocity profiles along the radius. With red, green, blue and yellow color have been highlighted different phases of profiles evolution.

the velocity profiles progressively tends to a linear shape. Figure 5.7 reports the amplitude of speed oscillations, scaled by its maximum value as a function of the dimensionless radial coordinate r/R for some experiments. Referring to the notation employed in section 2.1 the curves in figure 5.7 present the modulus of the dimensionless function $g* = g(r)/\left|g(R)\right|$ with g defined in equation (2.1.19). Each profile is associated with a different value of α, varying from 2.77 (experiment sin-4) to 6.12 (experiment sin-21). The agreement between measurements and theoretical predictions is fairly good for all values of a. Weak discrepancies may be partly due to the simplified assumptions adopted in the theoretical model and partly to measurement errors. In figure 5.8 a more detailed comparison is shown for experiments sin-10 and sin-16. In this case both the real Re and imaginary Im parts of the dimensionless function $g*$ are plotted. Notice that the ratio $Im(g*)/Re(g*)$ denotes the tangent of the phase lag ϕ with respect to the wall motion at each point in the radial direction. The agreement between measurements and theoretical predictions appears quite satisfactory for all the experiments.

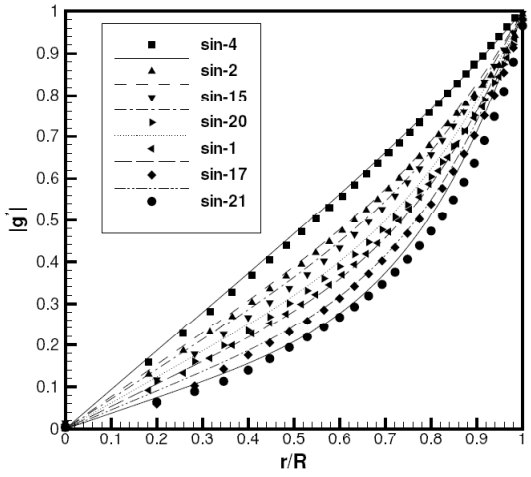

Figure 5.7: Radial profiles of the dimensionless function $|g*| = g(r)/\,|g(R)|$, for different runs. Symbols represent experimental measurements and continuous lines the corresponding theoretical distributions.

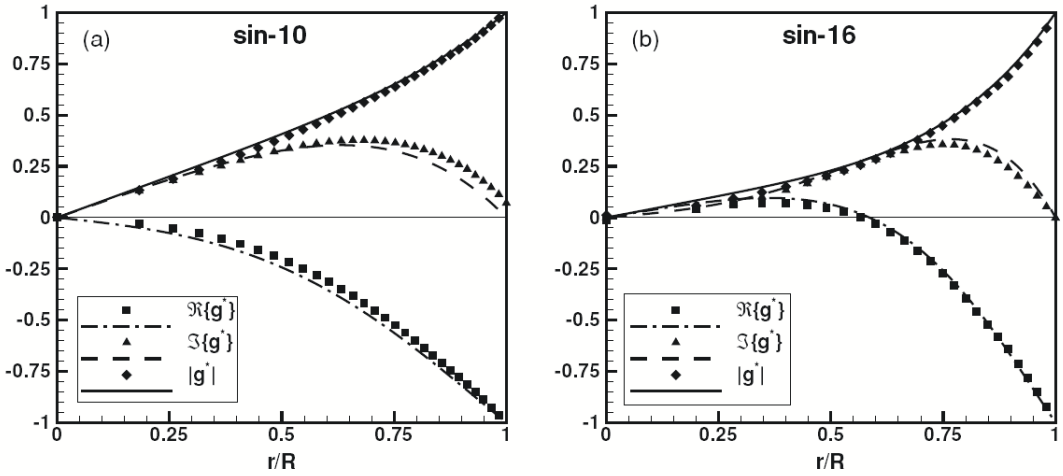

Figure 5.8: Comparison between the experimental and theoretical distributions of the functions $|g*| = g(r)/\,|g(R)|$, its real part $Re(g*)$ and its imaginary part $Im(g*)$. Symbols denote experimental data, continuous lines theoretical predictions.

5.2.3 Saccadic movements

The second set of experiments performed with the spherical model concerned rotations according to saccadic movements. The angular time law employed to describe the saccades is that in subsection 1.3.1 in chapter 1. One time it has

been decided saccade amplitude the duration of the movement is computed adopting equation 1.3.1. The corresponding duration that had to be used in the model is obtained multiplying the duration D of the prototype by the factor $(R/R_p)^2(\nu_p/\nu)$. Each saccadic movement is followed by a time of rest with a duration of five saccade duration, thus allowing a complete stop of the fluid motion. In the table 5.2 are collected the main characteristics of this set of experiments.

		Model			Prototype	
Exp. #	A [deg]	D [s]	ω_p [deg s^{-1}]	ν [10^{-4}m^2/s]	D [s]	ω_p [deg s^{-1}]
sac-1	10	0.1024	160.11	7.9	0.050	328.0
sac-2	15	0.1297	189.70	7.8	0.0625	393.6
sac-3	20	0.1536	213.48	7.9	0.075	437.33
sac-4	25	0.1793	228.73	7.9	0.0875	468.57
sac-5	30	0.2023	243.20	8.0	0.10	492.0
sac-6	35	0.2334	245.91	7.8	0.1125	510.22
sac-7	40	0.2529	259.42	8.0	0.1250	524.80
sac-8	45	0.3007	245.41	7.4	0.1375	536.73
sac-9	50	0.3468	236.45	7.0	0.1375	546.67
sac-10	30	0.2043	240.77	7.92	0.10	492.0
sac-11	40	0.2470	265.58	8.19	0.1250	524.80
sac-12	20	0.1428	229.69	8.5	0.0750	437.33
sac-13	10	0.1018	161.12	7.95	0.050	328.0
sac-14	15	0.1386	177.54	7.3	0.0625	393.60

Table 5.2: Main parameters of the experiments performed with saccadic movements.

The sampling rate employed in the experimental measurements was set such to obtain 20 vector fields within a saccade. According to the periodic experiments each velocity field $(u(x, y, t), v(x, y, t))$ on the equatorial plane, was measured 60 times.

5.2.4 Discussion of the results

As we can see in figure 5.9 reporting the experimental velocity profiles along the radial direction at different times for experiments sac-12 and sac-11 which reproduce a small amplitude (10 deg) and a large amplitude (40 deg) saccade, starting from rest, fluid motion is initially confined within a narrow region close to the wall and the thickness of such a region increases with time. As in periodic rotations three different phases of motion can be identified during

a single saccade. During the first acceleration phase of duration t_p, velocity profiles have the same concavity and the maximum velocity is located at the wall (solid lines). During the following phase of deceleration of the wall, the maximum of the velocity detaches from the boundary and progressively moves towards the center (dashed lines). The last phase describes the fluid motion dissipation from the time $t = D$ when the wall stops and the velocity profile go to the rest. During the decay phase velocity profiles become more symmetrical because the maximum velocity moves towards the point $r = R/2$.

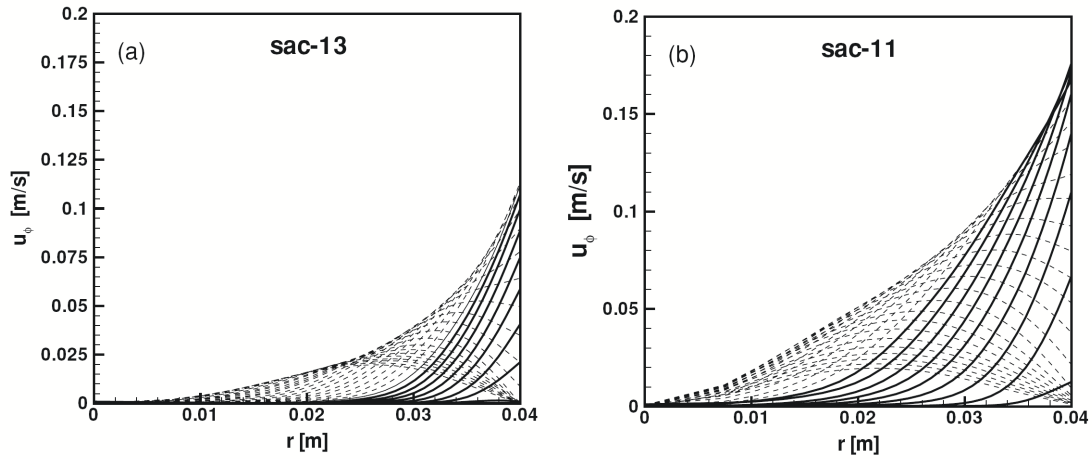

Figure 5.9: Experimental profiles of the averaged circumferential velocity ((a) and (b)) along the radius at different times with interval equal to 1/20 of the saccade duration D. (a): run sac-13 (D = 0.1018 s); (b) run sac-11 (D = 0.2470 s). Solid lines correspond to profiles measured during the acceleration phase of the saccade ($t < t_p$), while dashed lines to profiles measured during the deceleration phase.

Figures 5.10 (a) and (b) show the shear stress $\tau_{r\varphi}$ along the radial coordinate, corresponding to the velocity distributions previously shown in figures 5.9(a) and (b). The shear stress is computed according to the relationship:

$$\tau_{r\varphi} = \mu \left(\frac{\partial u_\varphi}{\partial r} - \frac{u_\varphi}{r} \right) \tag{5.2.1}$$

using the experimental circumferential averaged velocity profiles. In the above relationship μ represents the dynamic viscosity of the fluid. It appears that, during a saccadic movement, the wall is subject both to positive and negative shear stresses, though the maximum shear stress at the wall is invariably positive. In other words, quite intuitively, the fluid acts at the boundary with stresses mainly in the opposite direction with respect to the wall motion.

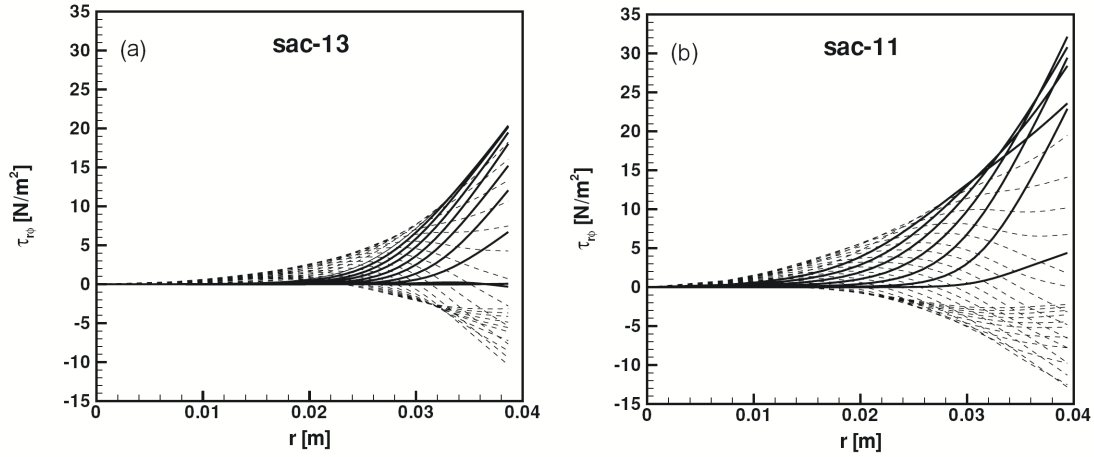

Figure 5.10: Experimental profiles of the shear stress $\tau_{r\varphi}$ along the radius at different times with interval equal to 1/20 of the saccade duration D. (a): run sac-13 (D = 0.1018 s); (b) run sac-11 (D = 0.2470 s). Solid lines correspond to profiles measured during the acceleration phase of the saccade ($t < t_p$), while dashed lines to profiles measured during the deceleration phase.

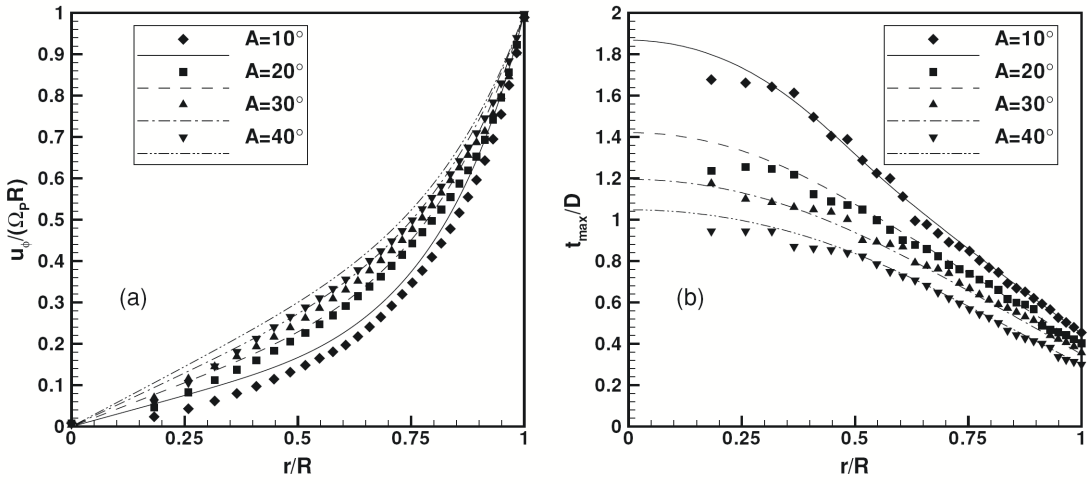

Figure 5.11: Dimensionless maximum value reached in time by the circumferential velocity u_φ at each point in the radial direction for different experiments. (b) Dimensionless time t_max/D at which the maximum tangential velocity is reached at each point in the radial direction for different experiments. Symbols denote experimental data, lines theoretical predictions.

The comparison between experimental results and theoretical predictions of equation (2.1.24) are in good agreement as we can see in figure 5.11 (a)

where the maximum velocity scaled by the corresponding maximum velocity at the wall $\Omega_p R$ is plotted along the dimensionless radial coordinate r/R. Each curve corresponds to saccades with a different amplitude. Continuous lines represent results obtained through the theoretical solution 2.1.24, symbols indicate the correspondent experimental results. Notice that in spite of the general agreement the theory slightly overestimates the maximum velocities. This figure shows that the envelope of the velocity profiles is thicker for larger saccades: small saccades induce in the deeper regions of the cavity a motion weaker than one driven by larger saccades. This is mainly due to the dependence of saccade duration from amplitude, thus larger saccades are longer and the boundary layer starting at the wall has more time to grow. In figure 5.11 (b) we report the dimensionless time t_{max}/D at which the maximum value of velocity is attained, for each value of the dimensionless radius. The agreement appears fairly good; the weak irregularity of experimental data is due to the relatively coarse temporal discretization of the saccadic movement (only 20 flow fields during a saccade duration). Note that the agreement is less satisfactory in the inner part of the domain. We can explain this results with the fact that in such a region the velocity is very small and affected by a greater relative experimental error. Obviously, at the wall ($r/R = 1$) the maximum velocity is attained at the acceleration time, so that $t_{max} = t_p$.

Since the agreement between theoretical predictions and experimental results has proved to be satisfactory for all the experiments, the theoretical model can be employed to obtain some information on problems that are of potential clinical relevance. Let us first consider the dependence of the maximum shear stress at the wall τ_{max} from the saccade amplitude. In figure 5.12 (a) the dimensional value of τ_{max} is plotted versus the saccade amplitude A. All quantities refer to the prototype, i.e. to the real scale eye. The three curves are relative to different values of the kinematic viscosity of the fluid. In all cases it appears that the maximum shear stress at the wall does not strongly depend on the amplitude of eye rotations. This result, which is not obvious, bears some practical importance. Indeed, small amplitude saccadic movements are, by far, more frequent than large ones. Therefore, according to the present results, small amplitude eye rotations stress the retina more than large movements and should be regarded as the most potentially dangerous ones. Finally, in figure 5.12 (b) the dependence of the maximum shear stress at the wall τ_max is plotted versus the viscosity of the fluid for different saccade amplitudes. Kinematic viscosity ν ranges between 10^{-6} (approximatively the viscosity of the completely liquefied vitreous) and $5 \times 10^{\{} - 3\}$ (relative to the most viscous silicon oil employed during vitrectomy). The figure shows that τ_max significantly increases with the fluid viscosity up to a value, for the most

viscous case, which is more than twice than that reached in the physiological case ($\nu \sim 10^{\{} - 4\}$). Thus, from a purely mechanical point of view, in order to minimize the stresses on the retina it would be preferable to employ low viscosity tamponade fluids.

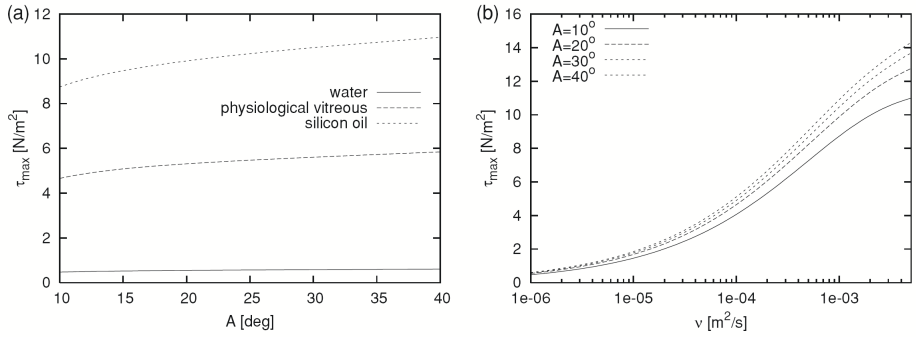

Figure 5.12: (a) Dependence of the maximum shear stress at the wall τ_max from saccade amplitude A for different fluid viscosities. (b) Maximum shear stress at the wall τ_max as a function of the kinematic viscosity of the fluid ν for different saccade amplitudes.

5.3 A deformed sphere model

The second type of model that has been employed in the experiments has a shape alike to the vitreous chamber of the real eye . Information coming from medical observations and theoretical results (see [18] and section 2.2), show that a weak deviation from the spherical shape may induce great difference in flow fields and consequently in fluid mixing. A set of chamber models have been specifically designed for this purpose from a careful analysis of several images of normal human eyes, taken from ultrasound and magnetic resonance scans. As we can observe in figure 5.13, a normal eye shape moves aside from a sphere because the antero-posterior axis is slightly shorter than the other two and the lens produces an indentation such that, in the anterior segment, the domain concavity faces inward.

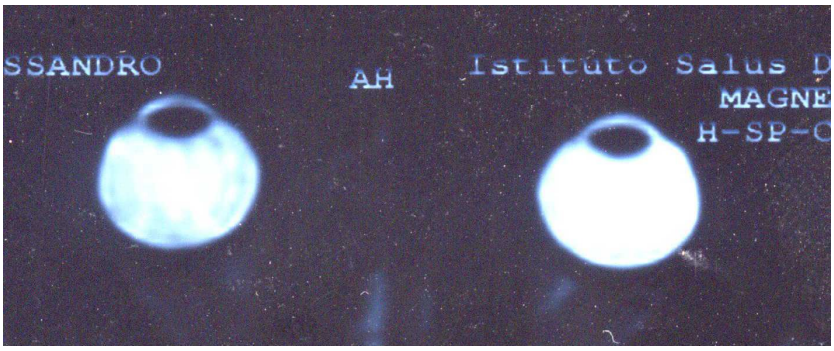

Figure 5.13: Equatorial section of normal eyes taken with RNM

This latter feature is significant on the flow characteristics more than the different length of eye axis, thus we decide to take it in account into the experiments. The eyes models consist in a spherical cavity of radius $R = 40.8$ mm carved in a perspex cylinder, with a deformation on one side simulating the lens, modeled as a spherical indentation. We have at our disposal three deformed model with "lens" jutting toward the center of the cavity of a length δ, (i.e. the maximum radial distance of the actual geometry from the outer spherical boundary) equal to the 15, 20 and 30 % of the model radius R.

Most of the experiments were conducted employing the "large lens" container ($\delta/R = 0.3$) since it is the one which most closely resembles the real shape of the vitreous cavity. In order to understand the influence of cavity shape on the basic features of the flow field we restrict the experimental conditions to **sine rotations**.

(a) (b)

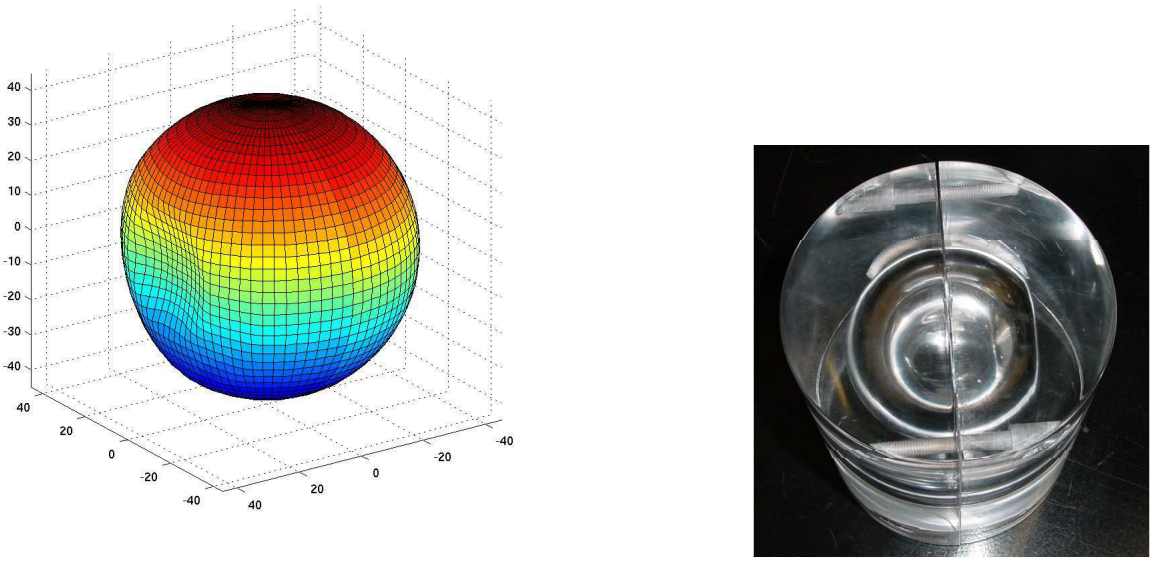

Figure 5.14: Deformed sphere model geometrical shape and realization

5.3.1 Periodic movements

We focus our attention on sinusoidal rotations of the eye model, with period T, frequency f and amplitude A. We have performed **three series of experiments**:

- In the **first set**,denoted as s1, we have reproduced sinusoidal rotations with the same amplitude and variable frequency, spanning in this way a range of Womersley number values from about 4.5 up to 15. This range is representative of real eye movements with the viscosity of the normal or liquefied vitreous; the motion in eyes in which the vitreous has been replaced by tamponade fluids may fall within this range as well. In the case of vitrectomized eyes, when the vitreous cavity is eventually filled with aqueous humour, much higher values of the Womersley number may be reached. Limitations of the experimental apparatus, however, precluded the possibility to investigate these conditions.

- A **second set** of experiments, denoted as s2, had the purpose to study the role of the amplitude of rotations for a fixed Womersley number

- The **third set** s3 was aimed at investigating the effect of different lens sizes.

In table 5.3 the dimensional and dimensionless parameters for each of the performed experiments are reported. The sampling rate chosen in the present experiments was such to produce 100 vector fields within a period for the lowest frequency experiments and 50 flow fields in a single period for the highest frequency runs.

Exp. #	A [deg]	α	f [Hz]	ν [10^{-4}m^2/s]	δ/R
1-s1	20	4.32	1	5.59	0.3
2-s1	20	5.97	2	5.87	0.3
3-s1	20	7.60	3	5.43	0.3
4-s1	20	8.74	4	5.48	0.3
5-s1	20	9.79	5	5.45	0.3
6-s1	20	10.87	6	5.31	0.3
7-s1	20	11.74	7	4.98	0.3
8-s1	20	12.96	8	4.50	0.3
9-s1	20	14.49	9	4.48	0.3
10-s1	20	12.39	10	6.81	0.3
11-s1	20	15.03	11	5.01	0.3
1-s2	10	3.63	1	7.91	0.3
2-s2	10	7.27	4	7.91	0.3
3-s2	10	9.62	7	7.91	0.3
4-s2	20	3.71	1	7.60	0.3
5-s2	20	7.37	4	7.60	0.3
6-s2	20	9.82	7	7.60	0.3
7-s2	30	3.69	1	7.70	0.3
8-s2	30	7.37	4	7.70	0.3
9-s2	30	9.82	7	7.70	0.3
10-s2	40	3.71	1	7.70	0.3
11-s2	40	7.37	4	7.70	0.3
12-s2	40	9.82	7	7.70	0.3
1-s3	20	8.48	4	5.40	0.2
2-s3	20	8.48	4	5.40	0.15

Table 5.3: Main parameters of the experiments performed.

Notice that the necessary image masking before image processing in this case has been almost complicated, because the irregular boundary of the measurements domain is rotating during the experiment. Each pair of images has been preprocessed before performing the cross-correlation analysis, applying an edge detection algorithm able to recognize the actual boundary of the measurement domain. In particular, a technique based on the search of the

zero-crossings of the image intensity Laplacian has been implemented with very good results.

Each flow fields was measured 50 times in order to obtain the averaged velocity fields $[u(x, y, t), v(x, y, t)]$.

5.3.2 Results discussion

Description of flow fields

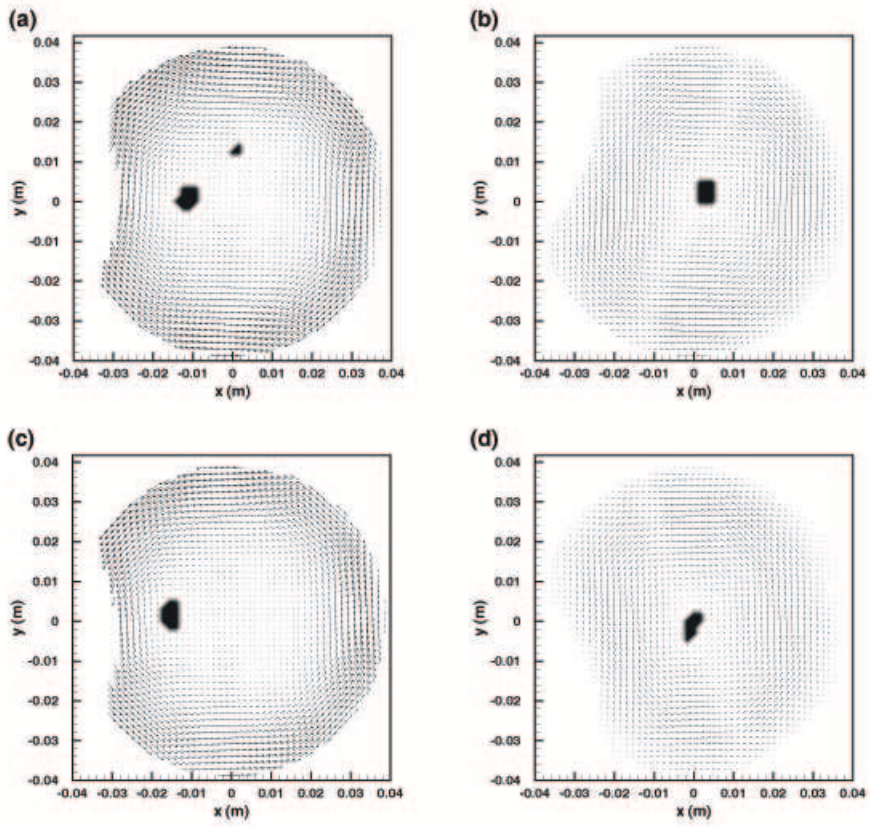

Figure 5.15: Flow fields on the equatorial plane at 4 different times ($\alpha = 4.32$, $a = 20°$). Contour lines refer to the scalar function Γ defined by equation (5.3.1). (a) $t \simeq 0$ (maximum container angular velocity), (b) $t \simeq 1/4T$, (c) $t \simeq 1/2T$, (d) $t \simeq 3/4T$.

In figure 5.15a-d, two-dimensional velocity vector fields are shown at different times for the run exp-1-s1 ($\alpha = 4.32$ and $a = 20°$). At a first glance it appears that the indentation at the anterior part of the vitreous cavity model significantly affects the flow field. In particular, approximately at time $t = 0$,

i.e. when the container reaches the maximum angular velocity, a vortex structure is present close to the lens. Spatial and temporal evolution of the vortex was enlightened by plotting contour lines of the function Γ defined in a generic point P by equation:

$$\Gamma(P) = \iint_S \sin\theta_M dS, \tag{5.3.1}$$

where S is an area surrounding P, the generic point M lies in S, θ_M represents the angle between the velocity vector in M and the radius vector $\overrightarrow{PM}$. This scalar function presented in Graftieaux (2001) [11] ranges between $[-1, 1]$ and its local maximum identifies the vortex center. Among the different techniques of vortex identification available in the literature this algorithm has proved to be particularly efficient in the present case. Other methods based on the Okubo-Weiss parameter and the swirling strength value (i.e. the eigenvalue of the tensor of the velocity gradients) that have been implemented, not being invariant to the reference system rotations, proved to be incapable of distinguishing the vortical structure from the average motion, rotating itself. In our PIV measurements the area S has a square shape, contains 20 measured velocity vectors and is centered at the point P. Notice that, as stated by Graftieaux (2001), the number of points contained in the area S has a weak influence on the location of the maxima of the function Γ.

From the observation of the sequence of plots reported in the figures 5.15 a-d it appears that the vortex is generated close to the lens and then it progressively migrates towards the center of the eye where it is eventually dissipated. A new vortex is then generated close to the lens with an opposite sign of rotation during the second half of the period. The generation of such a vortical structure is invariably observed during all the experiments performed, regardless of the amplitude of the sinusoidal motion, of the Womersley number, and of the size of the lens. An example of flow field at higher values of the Womersley number is reported in figure 5.16a-d (corresponding to run exp-7-s1, $\alpha = 11.74$, $a = 20°$).

The path followed by the vortex, relative to the container, however, is strongly dependent on α. At low values of α (figure 5.15 a-d) the flow is symmetrical with respect to the indentation, no wake effect is observed in the rear part of the moving lens and, in this case, the vortex migrates roughly along a straight pattern from the apex of the lens towards the center of the globe. As the Womersley number of the flow increases, differences between the flow in the regions upstream and downstream of the lens become visible (see figure 5.16 a-d) and such an asymmetry of flow conditions is responsible for a deviation of the vortex path. In this case the vortex does not move towards the center nor the trajectory it follows is straight. Consequently, at each cycle, the dissipation of the vortex occurs alternatively in a region under or below

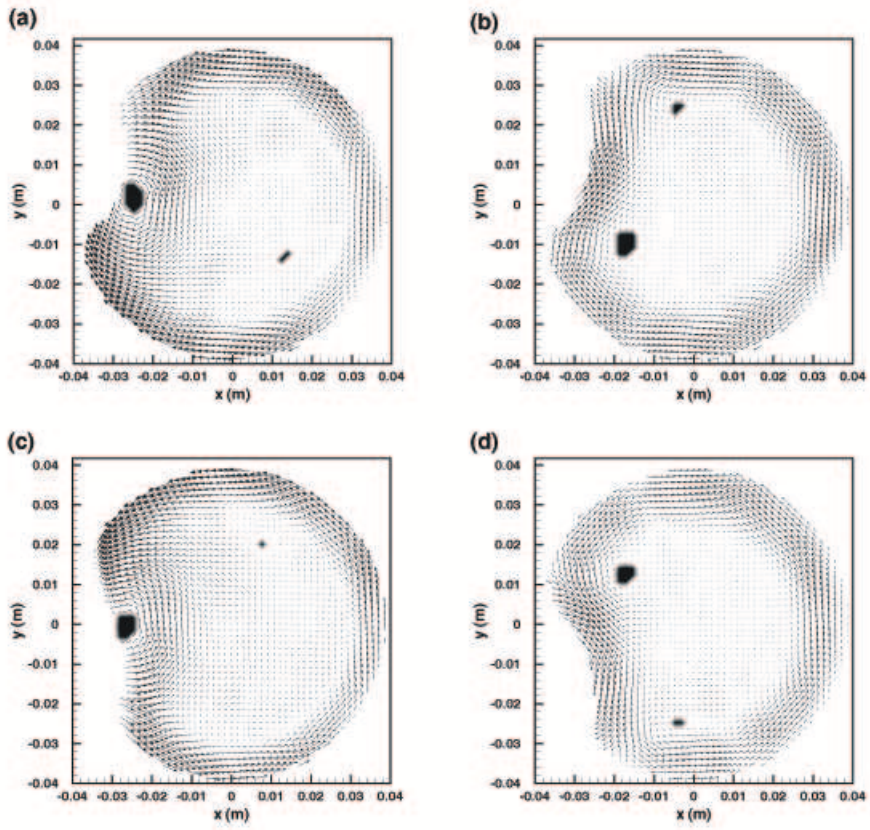

Figure 5.16: Flow fields on the equatorial plane at 4 different times ($\alpha = 11.74$, $a = 20°$). Contour lines refer to the scalar function Γ defined by equation (5.3.1). (a) $t \simeq 0$ (maximum container angular velocity), (b) $t \simeq 1/4T$, (c) $t \simeq 1/2T$, (d) $t \simeq 3/4T$.

the antero-posterior axis.

The differences between vortex trajectories at low and high values of the Womersley number are clearly shown in figure 5.17 (left). In the figure the path relative to the container covered by the vortex, identified by the position of the vortex core, is shown for different experiments varying α and keeping A constant. It appears that increasing the value of α the deviation of the vortex trajectory from the straight line connecting the apex of the lens to the center of the container become stronger. Moreover, for given values of α and A, the size of the indentation does not strongly influence the evolution of the vortical structures, as shown in figure 5.17 (right).

The vortex structure described above originates at time $t = 0$ (corresponding to the maximum container angular velocity)when there is a region, fairly

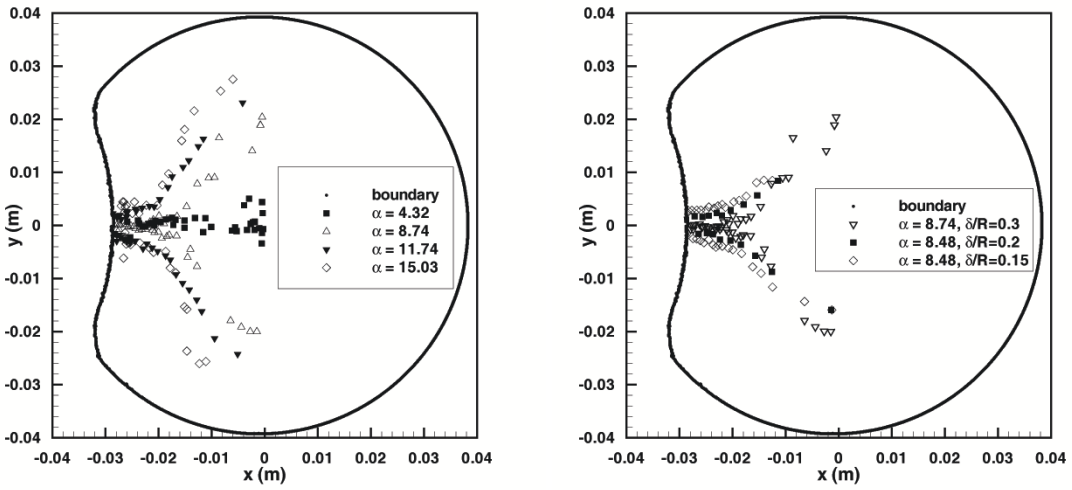

Figure 5.17: Trajectories relative to the container wall covered by the vortex centre during a cycle. Left: different values of the Womersley number α ($a = 20°$, $\delta/R = 0.3$). Right: different values of δ/R ($a = 20°$, $f = 4$ Hz, $\alpha \sim 8.5$).

close to the lens, where velocity vectors are directed in the opposite direction with respect to the wall motion and, consequently, also with respect to the velocity of the fluid located in the immediate vicinity of the wall. A similar flow pattern close to the lens can be inferred from the theoretical predictions of Repetto (2006) [18], obtained in the limit of low values of the fluid viscosity. The irrotational flow within the weakly deformed rotating sphere considered by the author is such that, close to the lens, the velocity is in the opposite direction with respect to the motion of the wall. If the flow within the boundary layer adjacent to the wall is accounted for this situation leads to the generation of a vortex structure similar to that observed in our experiments.

Particles trajectories

Particle trajectories can be obtained from the PIV measurements by integrating in time the measured velocity fields. The time integration is performed adopting an explicit fourth-order Runge-Kutta scheme and spatial velocity vectors are interpolated using a cubic polynomial interpolation algorithm. The above numerical integration is a delicate operation and might be affected by errors due to several reasons: first, we integrate an experimental velocity field that is affected by some error; second, the spatial resolution might not be sufficient to describe small scale effects and third, the temporal interval is fixed by the PIV acquisition rate. Nevertheless the computed trajectories represent,

with a good approximation, the paths followed by a fluid particle inside the domain. A better evaluation of the particle trajectories can be 0bviously achieved by a direct measurement of the trajectories by means of Particle Tracking Velocimetry technique. In figure 5.18 a particle paths relative to the container, sampled with a time step equal to the cycle period T, are plotted starting from different initial locations within the domain. In the case shown in the figure the Womersley number is equal to 11.74 and the rotations amplitude is 20°. A quite interesting pattern appears. Starting from different positions on the equatorial plane the particles tend to move towards a region located close to the anterior segment of the eye model. Clearly, for continuity to be satisfied, in such a region an ejection of fluid must occur in the direction orthogonal to the measurement plane according to the secondary flow observed for the spherical chamber and forecasted by theoretical results. Notice that quite a few periods of rotation (of the order of a hundred cycles on the average) are required for a full trajectory to be completed by a particle. The existence of a region of particle concentration is nicely shown in figure 5.18b, which is an instantaneous picture shot after having rotated the container for some tens of cycles. The light curves which can be observed are due to an initially higher concentrations of particle tracers in certain regions of the domain. Notice that these curves represent streaklines rather than particle trajectories. However, since we are dealing with a periodic motion and the trajectories are built sampling every of rotation period, streaklines and trajectories coincide in this case, see for instance Ottino (1989) [16].

In figures 5.19 particle paths relative to different values of the Womersley number are reported. For whatever value of α, particles tend to concentrate in the anterior part of the eye globe, regardless of their initial position. However, at low values of the Womersley number, i.e. when the flow field is remarkably symmetrical with respect to the axis $y = 0$, particles tend to concentrate close to the apex of the lens. On the other hand, as the value of the Womersley number increases, the generation of two distinct cells clearly appears, probably generated by the flow incipient asymmetry. The pattern of the cells is symmetrical with respect to the axis $y = 0$, and the cells seem to (slightly) move apart as α increases. This is possibly related to the different paths followed by the vortical structures during the container rotation, as discussed above.

The above observations show that the flow field generated by the container rotations has a very complex three-dimensional structure. A three-dimensional flow field was also observed in the case of a perfectly spherical rotating container, consisting in two slowly-rotating toroidal vortices whose formation is essentially induced by the spherical shape of the container. However, the vertical velocity components associated with such three-dimensional structures

(a) (b)

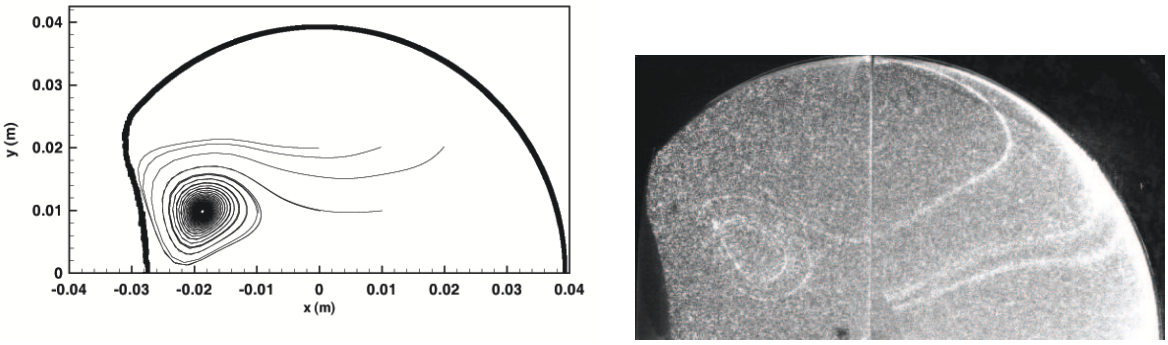

Figure 5.18: (a)Relative particle paths sampled every cycle; (b) Flow field structure visualisation obtained by injecting a higher particle concentration at a certain position. ($\alpha = 11.74$, $a = 20°$).

are found to be various orders of magnitude (~ 4) smaller than the velocity components on planes perpendicular to the axis of rotation. In the present case, the vertical fluid ejections described above, which are entirely related to the presence of the indentation in the anterior part of the cavity, generate a much more intense overall mixing of the fluid and are likely to be responsible of strong convection processes within the vitreous cavity.

This confirms the hypothesis that advective transport within the vitreous cavity is probably much more efficient than pure diffusion when the vitreous dynamics is intense enough, i.e. when the vitreous has lost its elastic properties (as a consequence of vitreous liquefaction or vitreous substitution with purely viscous tamponade fluids).

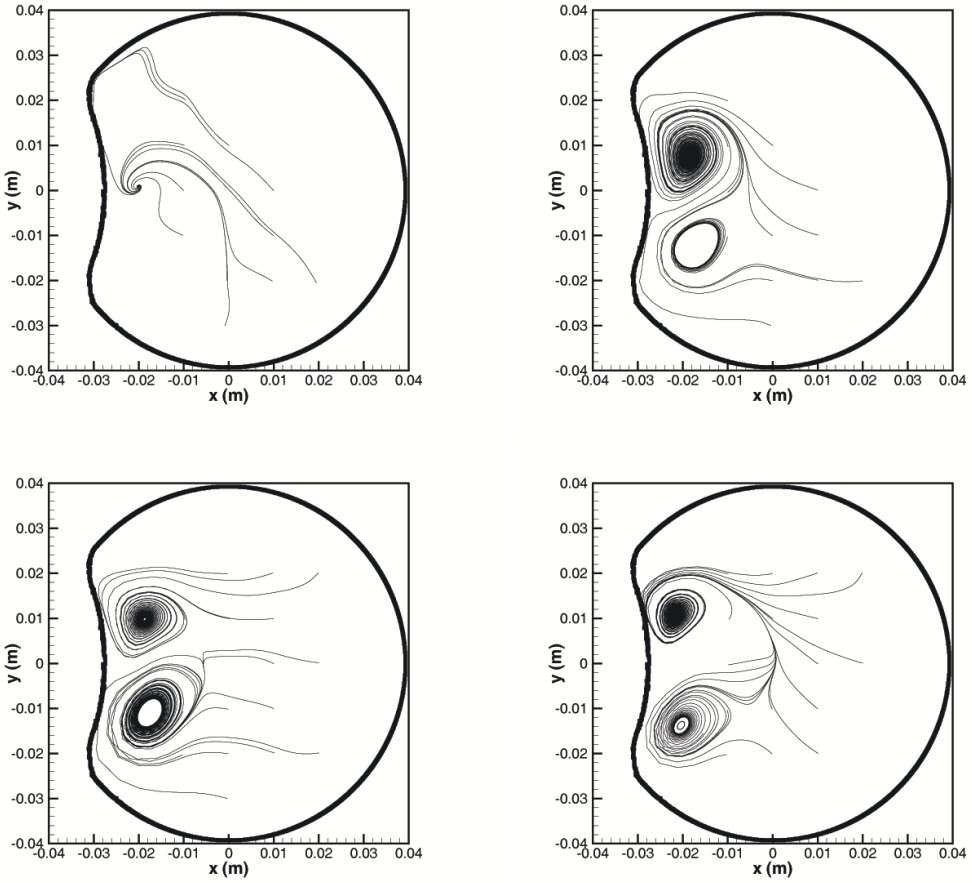

Figure 5.19: Relative particle paths sampled every cycle for different values of the Womersley number and constant amplitude $a = 20°$. From top left to bottom right $\alpha = 4.32$, 7.60, 9.79 and 11.74.

Shear stress at the wall

The distribution of the tangential stress τ_t at the wall on the equatorial plane has been computed considering only the deviatoric part of the stress tensor $\mathbf{T}_d$, namely $\mathbf{T}_d = 2\mu\mathbf{D}$ for a Newtonian fluid, with μ being the dynamic viscosity of the fluid and $\mathbf{D}$ the rate of strain tensor defined as the symmetric part of the velocity gradient tensor $\nabla\mathbf{u}$. The evaluation of $\mathbf{T}_d$ requires the determination of first spatial derivatives of the measured velocity field. It is well known that this is a quite delicate operation since derivation enhances measurement noise. Derivatives of the generic velocity component are computed through the second order scheme known as least square. Starting from the discretized

boundary $\mathfrak{F}$ of the equatorial plane, obtained with the edge detection algorithm described above, we have computed numerically the normal inward unit vector $\mathbf{n} = -\nabla\mathfrak{F}/|\nabla\mathfrak{F}|$ and the tangential unit vector $\mathbf{t} = (n_y, -n_x)$ (positive when rotating in the counter clockwise direction with respect to the center). The tangential stress τ_t is then obtained as $(\mathbf{n} \cdot \mathbf{T}_d) \cdot \mathbf{t}$.

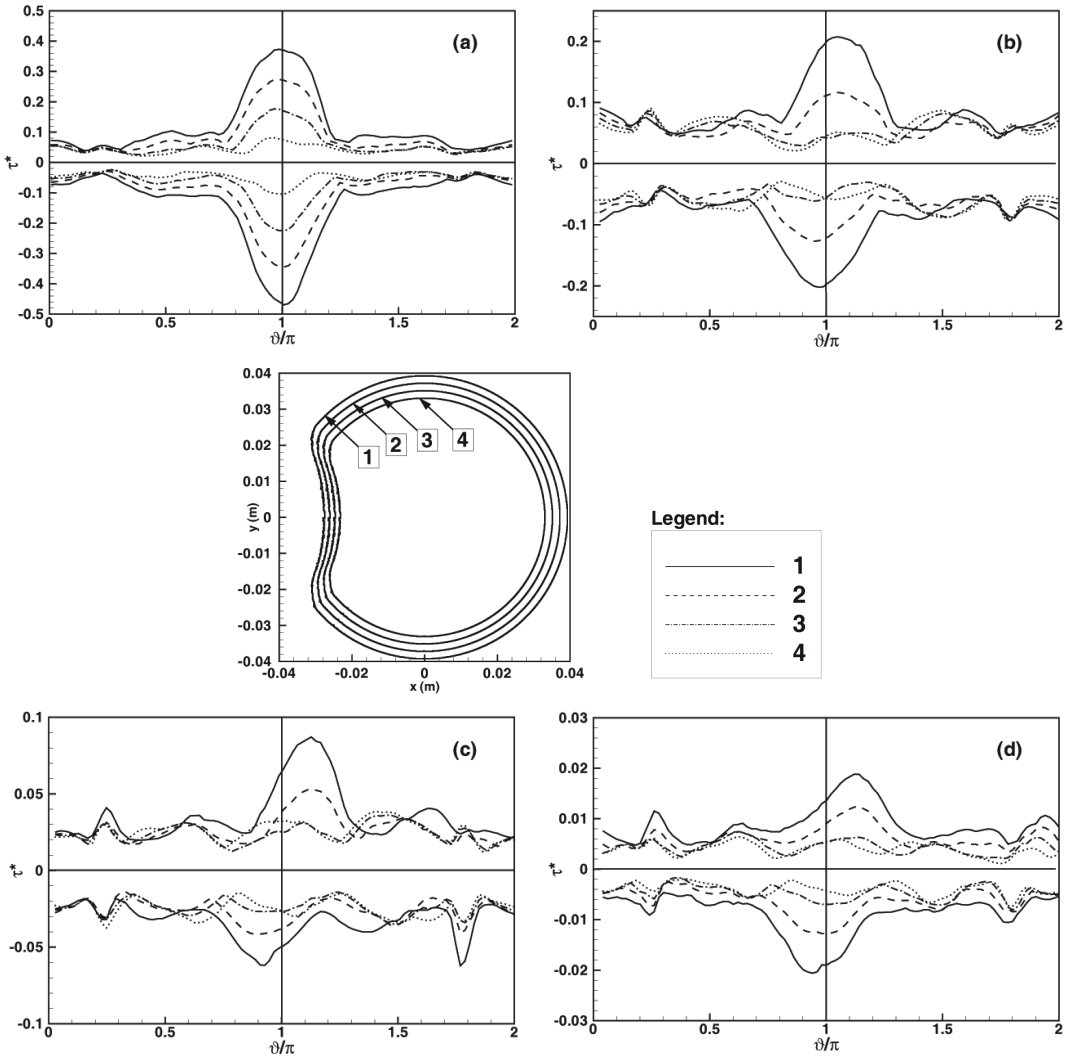

Figure 5.20: Distribution of the dimensionless tangential stress $\tau_t/(\rho\omega^2 R^2)$ as a function of the angle ϑ/π along the four different concentric boundaries shown in the central panel. $a = 20°$; (a) $\alpha = 4.32$, (b) $\alpha = 8.74$, (c) $\alpha = 11.74$, (d) $\alpha = 15.03$. The correspondence between shear stress profiles and the boundary is indicated in the central legend.

The stress distribution within the equatorial plane orthogonal to the axis

of rotation have been evaluated along concentric curves depicted in figure 5.20, labeled 1, 2, 3 and 4 moving toward the center of the eye. These curves are obtained by progressively contracting the wall profile to the axis of rotation. All such curves are spanned by the angular coordinate ϑ, the origin of which is assumed such that $\vartheta = \pi$ at the apex of the lens. The figures 5.20a-d show the maximum and minimum tangential stresses reached during a cycle along such curves. Notice that, from the clinical point of view, the most interesting quantities to be evaluated are the extremal values of the shear stress on the retina, i.e. on the cavity walls. The labels in the legend of figures 5.20a-d allow to associate each shear stress distribution with the curve on which the stress is computed. In all plots the shear stress is made dimensionless with the quantity $\rho\omega^2 R^2$. Each plot is relative to a different value of the Womersley number and the amplitude of the rotations A is kept constant. It clearly appears that the presence of the lens has a great influence on the shear stress distribution at the wall, especially at low values of the Womersley number. In particular, a clear stress peak is invariably observed close to the lens, where the maximum shear stress may well be up to five times larger than in most of the container wall. On the other hand, within the range $0.7 < \phi/\pi < 1.3$ the extreme values of the shear stress at the wall reached during each cycle remain fairly constant. As we move towards the center of the eye model the influence of the lens weakens and the shear stress distribution tends to become constant. Notice that the shear stress peaks at the wall have different locations for different values of the Womersley number of the flow. In particular, at low α, the maximum tangential stress is located almost exactly at the lens ($\vartheta = \pi$) as shown in figure 5.20a. As α increases (figures 5.20b-d) the location of the maximum shear stress progressively moves away from the lens in the upstream direction. This is a further effect of the asymmetry of the flow field with respect to the lens which is observed at high values of the Womersley number.

Finally, in figures 5.21 the maximum dimensional shear stress at the wall τ_{max}, reached during the entire period is shown as a function of the Womersley number (left) and of the amplitude of the container rotations (right). As one might expect it appears that τ_{max} increases, almost linearly, with both α and A.

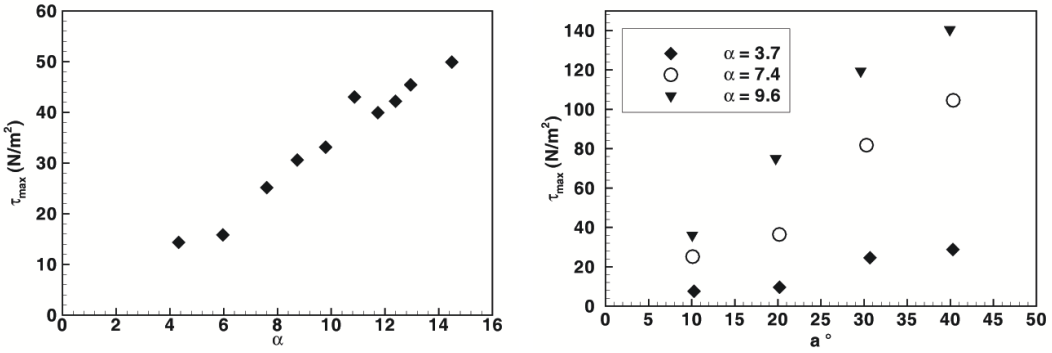

Figure 5.21: Maximum value of the dimensional tangential stress τ_{max} on the boundary of the eye globe model as a function of the Womersley number α (left panel, $a = 20°$) and of the amplitude of rotations a (right).

Chapter 6

Conclusions and future development

This first part of PhD Thesis deals with the study of the vitreous humour dynamic during eye rotations through experimental measurements of the flow field on a model of the vitreous chamber. Mechanical problems related to vitreous dynamics are extremely complicated, but their understanding is fundamental since it seems ascertained they are in connection with the occurrence of retinal detachment. As it often happens in attempting to model biological systems, the first approach to the problem is fairly idealized. One time the basic features of the phenomenon have been clarified we can eventually proceed to progressively complicate the scenario. The present work is based on the following main simplifying assumptions:

- the vitreous chamber has been considered a rigid cavity;
- the vitreous body has been modeled as a purely viscous homogeneous fluid.

The vitreous chamber model consists in a cavity carved in a perspex cylinder rotating around its axis with a prescribed time law. Preliminary flow visualizations permitted to evaluate the consistency of a secondary flow inside the chamber and justify the practical choice to take measurements only on the equatorial plane of the model orthogonal to the rotation axis. Velocity flow field have been measured with the PIV technique.

The experiments have been performed with two different models of the vitreous chamber:

- a **spherical model**;
- a **deformed sphere model** with an indentation simulating the lens presence.

With the **spherical model** two set of experiments were performed, imposing to the container a sine rotation and a polynomial time law reproducing the features of the real saccade. The *first set of experiments* was aimed to prove the feasibility of the theoretical predictions provided by David et al.(1998) [7] in the case of purely viscous fluid filling a sphere rotating with a periodic time law. Experimental results and theoretical predictions are found to be in good agreement: in particular we found that:

- a significant phase lag characterizes the motion of fluid in the most internal regions respect to the motion close to the wall;
- for larger values of the Womersley number, the oscillating boundary layer at the wall is thick and the fluid particles in the most inner region are subject to small amplitude oscillations.
- As far as the Womersley number decreases the envelope of the velocity profiles along the radius progressively tends to a linear shape.

In the *second set of experiments* real saccadic movements were reproduced. Such experiments show that:

- by analogy with the periodic rotations the motion of fluid in the most internal regions is out of phase respect to the motion close to the wall;
- fluid motion is generated at the wall, it spreads towards the center of the domain and, finally,it is dissipated;
- the maximum shear stress at the wall is not strongly dependent on the saccade amplitude, while it seems to increase significantly with the fluid viscosity .

The latter result implies that small amplitude eye rotations, being by far more frequent than large ones, are mainly responsible for the generation of stresses on the retina. Moreover, the dependence of maximum shear stress at the wall on the fluid viscosity, leads to the conclusion that, from a purely mechanical point of view, low viscosity tamponade fluids should be adopted after vitrectomy, in order to minimize the stresses on the retina. Thus, even in the quite simple conditions considered here, information which are potentially relevant from the clinical point of view have been obtained.

The second model employed during the experimentation allow us to focus our attention on the effect of the vitreous cavity shape on flow characteristics. Theoretical results provided by Repetto (2006) [18] suggest that the irregular shape of the vitreous chamber due to the presence of the lens in the

anterior segment of the eye, is responsible for the generation of a complex three-dimensional flow field. The lens presence was found to be more influent on the flow field than the different length of the three eye axis. A set of chamber models have been thus specifically designed for this purpose from a careful analysis of several images of normal human eyes, taken from ultrasound and magnetic resonance scans. The **deformed sphere model** consists in a spherical cavity with an indentation with its apex facing inward. Such a model has been moved only with periodic rotations. The experimental measurements come to the following results:

- in all experiments, the generation and migration of a vortex structure within the eye has been observed;

- the vortex path has been tracked in time, showing its strong dependence on the Womersley number of the flow: at low values of a such parameter the flow is symmetrical with respect to the indentation, during the motion the vortex migrates roughly along a straight pattern from the apex of the lens towards the center of the globe; as the Womersley number increases the flow in the regions upstream and downstream of the lens becomes different and a deviation of the vortex path is observed.

- Particle trajectories computed from the measured flow fields show that particles on the equatorial plane orthogonal to the axis of rotation tend to concentrate within narrow regions located close to the lens;

- a strong stress concentration was found in the region close to the lens, which is essentially related to the larger speeds reached in that region due to the concavity of the wall.

The generation of a vortex during each half period of rotation has a strong influence both on the shear stress distribution on the vitreous chamber wall and on the mixing processes taking place within the vitreous. The presence of regions of particles concentration suggests that in such areas the fluid is vertically ejected and this produces a highly three-dimensional circulation within the vitreous cavity. This clearly indicates the existence of an efficient mixing process which leads to the conclusion that advection should be accounted for in order to study transport phenomena within the vitreous cavity (in medical literature mixing processes have invariantly been interpreted in terms of a diffusion phenomena). As a matter of fact advection induced by such a complex flow field will invariably produce a highly non-isotropic transport, which can be hardly described by means of a single diffusion coefficient as done in all existing studies on the subject.

The present work provides a detailed description of the flow field within the vitreous cavity on the equatorial plane and demonstrates the existence of interesting hydrodynamic phenomena which will deserve future attention. Three-dimensional measurements (not very easy to conduct) and numerical simulations will be needed to complete the picture. Moreover, a next step in experiments will consist in the visco-elastic fluid employment, in order to better model the real physiological conditions.

Part II

A numerical model

Introduction to part II

As shown in the first part, experimental observations, though limited to the equatorial plane, provide a deep insight into the flow inside the posterior chamber of the eye. Although some clinical information have been already derived from the experimental models of the vitreous chamber, the PIV experimental set-up used to perform the experiments was obviously not able to appreciate three-dimensional flow inside the cavity. Besides, a more detailed knowledge of this class of flows is necessary to understand, for example, the actions of the fluid on the vitreous chamber wall or the mixing processes dynamics taking place inside the chamber.

Some analytical tools have been developed by Repetto (2006) [18], that allow for a successful description of the 3D oscillating flow inside the eye chamber, which are however limited to small amplitude oscillations.

The second part of the present work is dedicated to the formulation of a numerical model of the viscous flow inside an oscillating sphere. Spherical geometries in fluid dynamics have been deeply investigated for external flows. On the contrary, quite a few studies have been carried out for the case of internal flows.

The numerical model is based upon the work of Quartapelle et al. (1995) [41], who used a spectral technique to solve the Navier Stokes equations in primitive velocity-pressure form. The incompressibility constraint in the governing equation is replaced by the Poisson equation for the pressure. This approach allows the derivation of an uncoupled formulation by joining the Poisson equation with an independent condition of integral character for the pressure. The latter requires the evaluation of a set of metaharmonic functions, which, in turn, are solutions of the metaharmonic (Helmoltz) problem.

The equations are discretized in time by means of a non-fractional-step scheme. Non-linear terms are evaluated in an explicit manner using the three-level Adams-Bashforth method while the linear terms are accounted implicitly using the Crank-Nicolson method. A spectral technique is used to transform the original problem in a sequence of stationary problem for the coefficients of some suitable expansion. In this regard, the shape of the investigated domain

suggest the expansion of the unknown in series of spherical harmonics (see Mohlenkamp (1995) [37] and [38]). A suitable transform of variables allows to calculate, once and for all, an analytic solution for the metaharmonic problem, while the problem for the pressure and the velocity are solved by means of a tau projection spectral method having expanded the coefficients in terms of Chebyshev polynomials.

The second part of the present work is organized as follows: *chapter 7* contains a description of the numerical methods used to solve PDEs and provides an overview of the necessary mathematical tools with special focus on spectral methods; in *chapter 8* the problem is formulated, describing the time discretization adopted and introducing the integral condition for the pressure, modal equations for the Navier-Stokes problem in the specific case of a spherical gap are then introduced; in *chapter 9*the structure of the numerical code is detailing described, finally in *chapter 10* the solution of the metaharmonic problem is presented as a test case.

Chapter 7

Numerical methods for PDEs

A wide class of physical processes such as heat conduction, fluid flow, and sound propagation is described by suitable PDEs. Unfortunately, analytical solutions of PDEs are rarely found, as the available analytical techniques (such as variables separation) are limited in use. Moreover, even if the general integral is known, a particular integral depending on boundary conditions is typically required to solve the problem. Numerical methods allows to compute an approximation f_N of the exact solution f and to estimate the error $(f - f_N)$, obtained by substituting f with f_N .

In the present chapter the most common numerical technique to solve PDEs are described and compared. For further information on this topic the reader is referred to Boyd (2000) [28] and Quarteroni (2008) [44] .

7.1 Numerical methods for PDEs

Partial differential equations are differential equations containing derivatives of unknown function with respect to more than one variable. If we denote with $f(\mathbf{x}, t)$ the unknown function depending on $(d+1)$ independent variables $\mathbf{x} = (x_1, \ldots, x_d)$ and t, a generic PDE reads:

$$\mathcal{P}(f, g) = F\left(\mathbf{x}, t, f, \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d}, \ldots, \frac{\partial^{p_1+\ldots+p_d+p_t} f}{\partial x_1^{p_1} \ldots \partial x_d^{p_d} \partial t^{p_t}}, t\right) = 0 \quad (7.1.1)$$

being g the set of data from which the PDE depends on, and $p_1, \ldots, p_d, p_t \in \mathbf{N}$. The order of the PDE corresponds to the maximum order of the partial derivatives appearing in 7.1.1, therefore it is equal to the maximum value assumed by $p_1 + \ldots + p_d + p_t$. When the equation 7.1.1 depends only linearly on the unknown f and its derivatives, the equation is called *linear*. In this case:

$$\mathcal{P}(\alpha f + \beta v, g) = \alpha \mathcal{P}(f, g) + \beta \mathcal{P}(v, g) \quad (7.1.2)$$

A function $f = f(x_1, \ldots, x_d, t)$ is a solution or *particular integral* of the 7.1.1 if an identity is found when substituting it and its derivatives in 7.1.1. The whole set of solutions of 7.1.1 is called the *general integral*.

Numerical techniques to solve PDEs have massively developed together with the evolution of computer technology and can be considered as the most fundamental tool to solve the differential problems pertaining all the fields of science and engineering. Among the families of numerical methods for PDE we recognize the following:

- **finite difference method**, in which functions are represented by their values at certain grid points and derivatives are approximated through differences of these values;
- **finite element method**, where functions are represented in terms of basis functions and the PDE is solved in its integral (weak) form;
- **finite volume method**, which divides space into regions or volumes and computes the change within each volume by considering the flux (flow rate) across the surfaces of the volume.
- **spectral method**, which represents functions as a sum of particular basis functions, usually represented by orthogonal functions.

7.1.1 Finite Difference Methods

The **finite difference method** (FD) is often regarded as the simplest numerical method to solve differential equations. This technique is based on the transformation of continuous domain of the unknowns in a grid of points (usually evenly spaced) and so in a discrete domain .

Let h and k be some fixed space and time step, respectively. Set $x_j = jh$ and $t_n = nk$, for any integers j and n, a spatial and time discretization defining a regular grid or mesh of the integration domain. The continuous function $f(x, t)$ can be approximated by grid functions $f_j^n = f(x_j, t_n)$. The original problem can be thus expressed in term of the f_j^n.

The simplest kind of finite procedure is based on relationship that express f_j^{n+1} as a function of neighboring grid values at time steps n and $n-1$ (explicit case) or $n+1$ (implicit case).

In finite difference method the derivatives appearing in the PDE are replaced with approximately equivalent difference quotients. That is, because the first derivative of a function f is, by definition, the limit of the incremental ratio:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{7.1.3}$$

then a reasonable approximation for the first derivative, for some (small) value of h, would be:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} \quad (forward\ scheme) \tag{7.1.4}$$

which can be easily shown to correspond to a Taylor expansion truncated at first order in h. The error committed substituting 7.1.4 to the exact derivatives is therefore of order $O(h^2)$. First derivative can also be expressed by:

$$f'(x) \approx \frac{f(x) - f(x-h)}{h} \quad (backward\ scheme) \tag{7.1.5}$$

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} \quad (centered\ scheme) \tag{7.1.6}$$

Similarly, the second derivative:

$$f''(x) = \lim_{h \to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \tag{7.1.7}$$

can be replaced by:

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \tag{7.1.8}$$

where the error involved is now of $O(h^3)$.

Replacing all of the derivatives in a differential equation with their discrete counterparts reduces it to a recurrence expression for all the spatial and temporal grid nodes involving the value of the unknowns at the grid point and their value in the neighboring points. This means that the solution of the differential equation can be found solving an algebraic system of equations in the mesh nodes. The problem can be solved in a direct or indirect manner according to the assumed discretization scheme using appropriated boundary conditions.

The Finite Different Methods are usually classified as:

1. **Explicit**: when the PDE can be solved directly using the appropriated boundary conditions and proceeding forward in time through small intervals. Convergence is assured for specifics size of increments interval length.

2. **Implicit**: when the PDE can be solved indirectly by solving a system of simultaneous linear equations. Convergence is always assured.

The error between the finite difference approximation and the exact solution arises essentially from two sources: the *round-off error* (produced by computer rounding of decimal quantities) and the *truncation error* (due to the difference between the exact solution of the finite difference equation and the exact quantity), which is proportional to some power of the step size. Several discretization schemes of different accuracy exist, the choice of the best numerical scheme depends on the differential structure of the PDE.

7.1.2 Finite Element Methods

Finite element techniques are a good choice for solving PDE when the geometry of the domain is complicated, the domain changes in time, the desired accuracy varies over the domain, or when the solution lacks smoothness. The **finite element method** (FEM) consists in finding approximate solutions of partial differential equations (PDE) as well as of integral equations. In order to apply the finite element methods one has first to reformulate the original boundary value problem in its weak, or *variational form.* A little number of computations is usually required for this step. Secondly, one has to *discretize* the weak form of the problem in a finite dimensional space. After this second step, one usually ends up with a large but finite dimensional linear problem whose solution will approximately solve also the original BVP. This finite dimensional problem is then implemented on a computer.

Let us take for example the following simple one-dimensional problem:

$$\text{P1} \ : \begin{cases} f''(x) = g(x) \text{ in } (0,1), \\ f(0) = f(1) = 0, \end{cases} \tag{7.1.9}$$

where g is given, f is an unknown function of x, and f'' is the second derivative of f with respect to x. To convert this problem in its variational, or weak form we observe that if f is a solution of 7.1.9, then for any smooth function v that satisfies the displacement boundary conditions, i.e. $v = 0$ at $x = 0$ and $x = 1$, we have:

$$\begin{aligned} \int_0^1 g(x)v(x)\,dx &= \int_0^1 f''(x)v(x)\,dx \\ &= -\int_0^1 u'(x)v'(x)\,dx = -\phi(u,v) \end{aligned} \tag{7.1.10}$$

This is the variational form of the problem, equivalent to 7.1.9. In fact, if f with $f(0) = f(1) = 0$ satisfies 7.1.10 for every smooth function $v(x)$ then one may show that this f will also solve 7.1.9.

Multi-dimensional problems can be reduced to their weak formulation using Green's theorem. The function $v(x)$ belongs to a finite dimensional subspace V of H_0^1, i.e. the Hilbert space of the solution. For the finite element method we take V to be a space of piecewise linear functions. For example, for the problem 7.1.9, we take the interval $(0,1)$, choose n values of x with $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = 1$ and we define V by:

$$V = \{v : [0,1] \to \mathbb{R} \ : v \text{ is continuous, } v|_{[x_k, x_{k+1}]} \text{ is linear for } k = 0, ..., n, \text{ and } v(0) = v(1) = 0\} \tag{7.1.11}$$

Notice that functions in V are not differentiable according to the elementary definition of calculus. Indeed, if $v \in V$ then the derivative is typically not defined at any $x = x_k$, $k = 1, \ldots, n$. However, the derivative exists at every other value of x and one can use this derivative for the purpose of integration by parts.

In the one-dimensional case the basis, for each control point x_k is the piecewise linear function $v_k \in V$ whose value is 1 at x_k and zero at every x_j, $j \neq k$, i.e.,

$$v_k(x) = \begin{cases} \frac{x - x_{k-1}}{x_k - x_{k-1}} & \text{if } x \in [x_{k-1}, x_k], \\ \frac{x_{k+1} - x}{x_{k+1} - x_k} & \text{if } x \in [x_k, x_{k+1}], \\ 0 & \text{otherwise,} \end{cases} \tag{7.1.12}$$

for $k = 1, \ldots, n$.

If we expand the functions g and f on v_k basis:

$$f(x) = \sum_{k=1}^{n} f_k v_k(x) \quad g(x) = \sum_{k=1}^{n} g_k v_k(x) \tag{7.1.13}$$

then problem 7.1.9 becomes:

$$-\sum_{k=1}^{n} u_k \phi(v_k, v_j) = \sum_{k=1}^{n} f_k \int v_k v_j \ \ for \ j = 1, \ldots, n \tag{7.1.14}$$

If we denote by $\mathbf{f}$ and $\mathbf{g}$ the column vectors $(f_1, \ldots, f_n)^T$ and $(g_1, \ldots, g_n)^T$, and if we let $L = (L_{ij})$ and $M = (M_{ij})$ be matrices whose entries are $L_{ij} = \varphi(vi, vj)$ and $M_{ij} = \int v_i v_j$ then we may rewrite 7.1.14 as:

$$-L\mathbf{u} = M\mathbf{f} \tag{7.1.15}$$

Most of the entries of L and M are zero because the basis functions v_k are non-zero only over a small number of nodes. So we have to solve a linear

system in the unknown $\mathbf{f}$ where most of the entries of the matrix L, which we need to invert, are zero. Such matrices are known as sparse matrices, and there are efficient solvers for such problems. In addition, L is symmetric and positive definite, so techniques like the *conjugate gradient method* are favored. For problems that are not too large, sparse LU *decompositions* and *Cholesky decompositions* still work quite well.

7.1.3 Spectral method

Spectral methods for partial differential equations were originally developed by meteorologists. By the present time, thank to the progress of computer technology and the development of methods to treat nonlinearities, spectral methods have become the prevailing numerical technique to solve certain problem of CFD (such as three-dimensional direct and large eddy simulations of turbulent flows, laminar turbulent transition, etc.) and constitute a valid alternative of other numerical techniques for many other applications. The basic idea of **spectral methods** is to assume that the unknown $f(x)$ can be expanded in terms of some global and, usually orthogonal test functions $P(x)$.

$$f(x) \approx \sum_{i=0}^{N} f_i P_i(x) \tag{7.1.16}$$

that are required to ensure that the differential equation for the original function is satisfied as closely as possible by its the truncated series expansion. This is achieved by minimizing the residual (i.e. the error committed replacing the solution with its approximation) with respect to a suitable norm. The original partial differential problem reduces then, to the research of expansion coefficients values through the solution of algebraic or ordinary differential equations. The choice of the test functions distinguishes between the three most commonly used spectral schemes, namely, the collocation, Galerkin and tau approaches.

- In the **collocation** approach the test functions are translated Dirac delta functions centered at special points called collocations nodes. Collocation approach is also denoted as *interpolating* or *pseudospectral*. Being:

  $$f(x) \approx p_N(x) = \sum_{i=0}^{N} a_i \delta(x - x_i) \tag{7.1.17}$$

  an approximation of the function $f(x)$, solution of a differential equation, the coefficients a_n are found by requiring that the residual function is

identically equal to zero at the collocation nodes:

$$R(x_i, a_0, a_1, \ldots, a_N) = 0 \quad i = 0, 1, \ldots, N \tag{7.1.18}$$

This means that this technique requires the differential equation being exactly satisfied at the *collocation* or *interpolation* points. Increasing the number of points, where $R(x; a_n)$ is forced to vanish, the residual will become smaller and smaller in the gaps between neighboring collocation points and the approximation $p_N(x)$ will smoothly converge to $f(x)$.

- In the spectral **Galerkin** approach, the test functions are orthogonal basis functions which are infinitely smooth and individually satisfy the boundary conditions. Galerkin's methods is also denoted as *non-interpolating* or *orthogonal collocation*. There is no grid of interpolation points. The exact solution of the differential equation $f(x)$ is then approximated by:

$$f(x) \approx p_N(x) = \sum_{i=0}^{N} a_i P_i(x) \tag{7.1.19}$$

where $P_i(x)$ are orthogonal functions. The expansion coefficients a_i are computed by multiplying 7.1.19 by the basis functions and integrating, making use of the orthogonality properties of $P_i(x)$.

- Spectral **tau methods** are similar to Galerkin's. The exact solution of the differential equation $f(x)$ is approximated by:

$$f(x) \approx p_N(x) = \sum_{i=0}^{N} a_i P_i(x) \tag{7.1.20}$$

where $P_i(x)$ are orthogonal functions. The expansion coefficients a_i are computed by integration. The difference between Galerkin and tau methods is that none of the test functions needs to satisfy the boundary conditions. Hence a supplementary set of equations is used to apply the boundary conditions.

Spectral methods are distinguished non only by the choice of the test functions used but also by the particular choice of the basis on which the solution is to be expanded. The most frequently used basis functions are Fourier series, Chebyshev polynomials, and Legendre polynomials. Nevertheless, the choice of the best basis functions depends to the problem conditions. A more detailed discussion on this topic is presented in section 7.3.1. In the following, with the aim of understanding the essence of spectral methods, some classical numerical analysis on topics like polynomial interpolation, function approximation, and Gaussian integration will be reviewed in sections 7.2, 7.3 and 7.4, respectively. A more complete treatment on these topics can be found in [28].

7.1.4 Comparison between Spectral Methods and Finite Difference Methods

Finite difference methods approximate the unknown $f(x)$ by a sequence of overlapping polynomials which interpolate $f(x)$ at a set of grid points. The derivative of the local interpolant is used to approximate the derivative of $f(x)$. The result takes the form of a weighted sum of the values of $f(x)$ at the interpolation points.

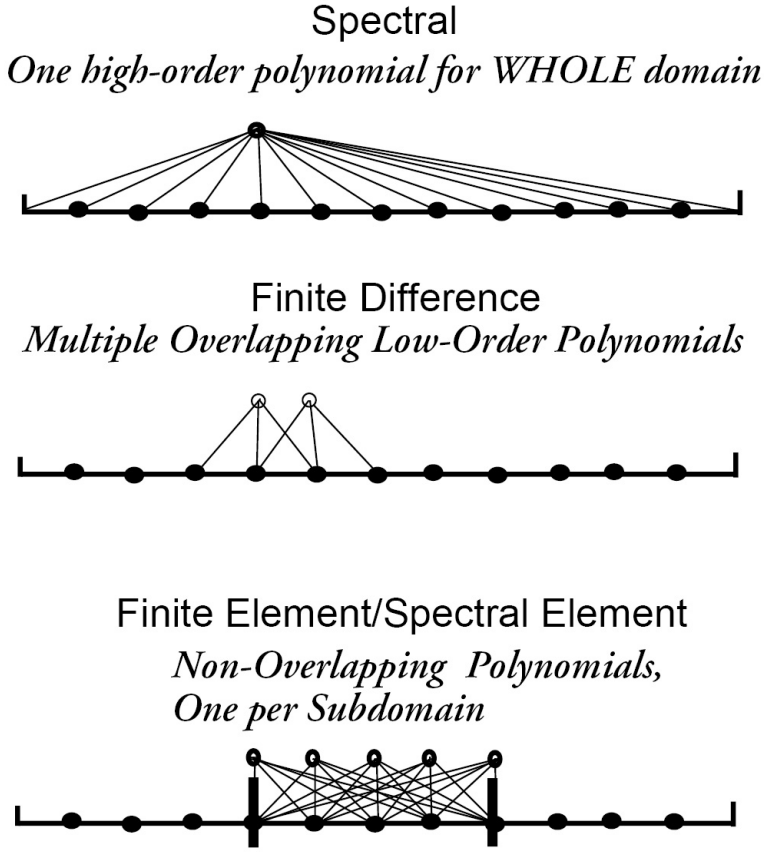

Figure 7.1: Three types of numerical algorithms. The thin, slanting lines illustrate all the grid points (black circles) that directly affect the estimates of derivatives at the points shown above the lines by open circles. The thick black vertical lines in the bottom grid are the subdomain walls.

The most accurate schemes are tipycally centered ones. For example quadratic,

three-point interpolation and quartic, five-point interpolations give:

$$\frac{df}{dx} = \frac{[f(x+h) - f(x-h)]}{2h} + O(h^2) \tag{7.1.21}$$

$$\frac{df}{dx} = \frac{[-f(x+2h) + 8f(x+h) + f(x-2h)]}{12h} + O(h^4) \tag{7.1.22}$$

Notice that the order of magnitude of the error committed replacing these expressions to the exact derivatives is proportional to h^2 and h^4, respectively.

Similarly, the finite element and spectral collocation algorithms approximate derivatives as a weighted sum of grid point values. However, in finite difference and finite element methods only those points which lie within a given subdomain contribute directly to the derivative approximations in that subdomain[1].

For a given number N of grid points pseudospectral differentiation has an accuracy of order $O(h^N)$, much greater than finite difference 3-point formulas, or even 5-point formulas. The high accuracy of spectral method depends on the fact that the basis functions can be differentiated analytically and since each spectral coefficient a_n is determined by all the grid point values of $f(x)$, to equal the accuracy of the pseudospectral procedure for $N = 10$, one would need a tenth-order finite difference or finite element method. As the number of nodes increases, the pseudospectral method benefits in two ways. First, the interval h between grid points becomes smaller: this would cause the error to rapidly decrease even if the order of the method were fixed. Unlike finite difference and finite element methods, however, the order is not fixed. When, for example, N increases from 10 to 20, the error becomes $O(h^{20})$ in terms of the new, smaller h. Since h is $O(\frac{1}{N})$, we have:

$$pseudospectral\, error \approx O\left[\left(\frac{1}{N}\right)^N\right] \tag{7.1.23}$$

The error is decreasing faster than any finite power of N because the power in the error formula is always increasing, too. This is the so-called *infinite order* or *exponential* convergence. for comparison, we may recall that the finite difference counterpart of the 7.1.23 is:

$$finite\, difference\, quadratic\, error \approx O\left[\left(\frac{1}{N}\right)^2\right] \tag{7.1.24}$$

[1]Because the solution in one subdomain is matched to that in the other subdomain, there is an indirect connection between derivatives at a point and the whole solution.

7.1.5 Comparison between Spectral Methods and Finite Element Methods

Spectral methods and Finite Element Methods are closely related. The main difference between them is that in Spectral Methods the solution is approximated as a linear combination of continuous functions that are generally defined and infinitely differentiable over the whole domain of solution (for example sinusoids or Chebyshev polynomials), while in Finite Element method the solution is approximated as a linear combination of piecewise functions that are nonzero only over subdomains. Because of this, it is said that Spectral methods take a global approach while, on the contrary, Finite Element Methods are local.

Finite Elements Methods have the advantages to convert differential equations into matrix equations that are sparse[2], because only a handful of basis functions are non-zero in a given sub-interval. Moreover the local character of trial functions make them well suited for handling complex geometries, especially in multi-dimensional problems where the little sub-intervals become little triangles or tetrahedra which can be fitted to irregularly-shaped bodies. The principal disadvantage of finite element methods is the low accuracy because each basis function is a polynomial of low degree. When accuracy is an issue three different strategies are available for FEMs. The first is to subdivide each element so as to improve resolution uniformly over the whole domain. This is usually called *h-refinement* because h is the common symbol for the size or average size of a subdomain. The second alternative is to refine only in regions of steep gradients where high resolution is needed. This is the called *r-refinement*. The third option is to keep the subdomains fixed while increasing the degree of the polynomials in each subdomain. This strategy is called *p-refinement*, where p is the degree of the interpolating polynomials, is in some sense precisely that employed by Spectral Methods.

Spectral Methods generate algebraic equations with full matrices, but in compensation, the high order of the basis functions ensures the highest accuracy for a given N. In fact, if there are $N + 1$ grid points, the derivatives of the functions are calculated from a polynomial of degree N and the same polynomials are used in all the nodes. Local methods produce derivatives with an accuracy order α equal to the order of local polynomial used to approximate the solution minus one, with an error decreasing as $1/n^{\alpha}$, whereas the error from the spectral methods decreases exponentially. As spectral method use global information they are most useful when the geometry of the problem is

[2]Sparse matrix equations can be solved in a fraction of the computational cost of problems of similar size with full matrices even if computational accuracy can become an issue.

fairly smooth and regular. Nowadays, the development of fast iterative matrix-solvers is making Spectral Methods much more efficient than Finite Element Methods in many applications.

7.2 Interpolation of functions

In engineering or science applications, data collected from sampling, experimentation or numerical processes are usually discrete and their physical meanings are not always well known. To estimate the outcomes and, eventually, to have a better insight of the physical phenomenon, an analytical functional form that fits the field data is desirable. Interpolation techniques mind to this task. The process through which one constructs a function which closely fits a collection of discrete data is called curve fitting or regression. Interpolation is a specific case of curve fitting, in which the interpolating function must go exactly through the data points. Interpolation provides a means of estimating the function at intermediate points and deriving analytically further information on physical quantity related to the original one. There are many different interpolation techniques. Most common methods use polynomials for fitting the data. This choice is justified by the fact that polynomials are easy to handle and are the only functions that a computer can evaluate exactly. Trigonometric functions, sines and cosines, give rise to trigonometric interpolation and related Fourier methods, useful with data characterized by a periodic behaviour. Before choosing an appropriate interpolation algorithm one has to take in account the accuracy of the method, its computational cost and the interpolating function features. Consider for example that, local interpolation based on a small number of neighbor points gives interpolated values f(x) that in general do not have continuous derivatives. So, in situation where continuity of derivatives is a concern, one must use an interpolation technique whose coefficients are determined non-locally in order to guarantee the smoothness of the interpolated function. The number of points (minus one) used in an interpolation scheme is called *order* of the interpolation. Notice that increasing the order does not necessarily increase the accuracy of the formula, especially in polynomial interpolation. In such interpolated functions high order of interpolation give rise to Runge's phenomenon which consists in a oscillation of the function between the tabulated values especially close to the endpoints.

In the following we discuss the main feature of some common interpolation methods.

7.2.1 Linear interpolation

Linear interpolation is the simplest method pertaining to polynomial interpolation. Given a collection of point (x_i, y_i) the interpolating function consists in a concatenation of linear interpolants between each pair of data points. Given two data points, say (x_A, y_A) and (x_B, y_B) the linear interpolant function is:

$$y(x) = y_A - \frac{(y_B - y_A)}{(x_B - x_A)}(x - x_A) \tag{7.2.1}$$

Facing an easy implementation, linear interpolation is not is not very accurate unless the tabulated points are very, very close together and presents the disadvantage that the interpolated functions are not differentiable in nodes.

7.2.2 Polynomial interpolation

Polynomial interpolation is a generalization of linear interpolation referred to interpolant polynomials with order equal to the number of the abscissas minus one. Generally, if we have n+1 data points, we look for a polynomial p of degree at most n going through all the data points:

$$p(x_i) = y_i \quad i = 0, ..., n \tag{7.2.2}$$

If the interpolation polynomial is in the form:

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0. \tag{7.2.3}$$

Substituting the constraint 7.2.2 in 7.2.3, we obtain a system of linear equations in the coefficients a_k. The system in matrix-vector form reads:

$$\begin{bmatrix} x_0^n & x_0^{n-1} & x_0^{n-2} & \dots & x_0 & 1 \\ x_1^n & x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_n^n & x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{bmatrix} \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_0 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix} \tag{7.2.4}$$

The matrix of coefficients of this system of equations is usually referred to as a Vandermonde matrix. Its determinant is nonzero, which proves the interpolant polynomial 7.2.3 exists and it is unique.

The solution of the linear system 7.2.4 can be a costly operation. In order to avoid this drawback the interpolation polynomial can be expressed using

the Lagrange form:

$$p(x) = \frac{(x-x_1)(x-x_2)...(x-x_n)}{(x_0-x_1)(x_0-x_2)...(x_0-x_n)}y_0 + \frac{(x-x_2)(x-x_2)...(x-x_n)}{(x_1-x_0)(x_1-x_2)...(x_1-x_n)}y_1 + \\ +... + \frac{(x-x_0)(x-x_1)...(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)...(x_n-x_{n-1})}y_n = L_0y_0 + L_1y_1 + ... + L_ny_n \tag{7.2.5}$$

where L_i are the Lagrange polynomial defined such that

$$L_i(x) = \prod_{l=1, l\neq i}^{j} \frac{x-x_l}{x_i-x_j}. \tag{7.2.6}$$

The interpolating polynomial 7.2.5 consists in n+1 terms each a polynomial of degree n costructed to be zero at all of the x_i except one, at which it assume the value y_i. Choosing Lagrange form of the interpolating polynomial you don't have to compute expansion coefficients by solving 7.2.4 anymore, but you have to compute all the $n+1$ Lagrange polynomials.

With polynomial interpolation the interpolant is a polynomial defined over the entire interval $x_0, ..., x_n$ and thus infinitely differentiable. Thus, the convergence of the interpolating polynomial to the interpolated function is not assured and does not improve increasing the order of the polynomial. In the case of equidistant nodes the uniform convergence is not even guaranteed for infinitely differentiable functions. One classical example, due to Carl Runge, is the function $f(x) = \frac{1}{1+x^2}$ considered on the interval $[-5, 5]$. As we can see in figure 7.2.2 the interpolation error $\|f - p_n\|$ grows without bound as n tends to infinity especially near the endpoints.

Possible solutions to this problem could be:

- use high order polynomial defined over interpolation intervals much greater than the desired one, from which extract the selected interval and cut the worse region of approximation;
- use an optimal distribution of interpolation nodes: choosing the points of intersection of the Chebyshev polynomial of order n as interpolation nodes we obtain the interpolating polynomial coinciding with the best approximation polynomial;
- use piece defined polynomials of low order to interpolate the data.

The latter choice leads to the spline interpolation.

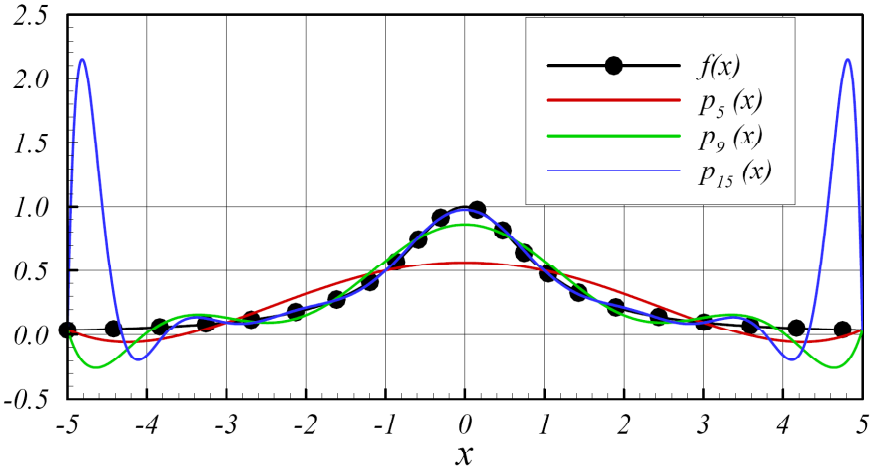

Figure 7.2: Runge function f(x) and its interpolant polynomials of order 5, 9 and 15

7.2.3 Spline interpolation

Spline interpolation uses low-degree polynomials to interpolate a collection of data by splitting the domain in small intervals, and chooses the polynomial pieces such that they fit smoothly together. The most useful kind of spline is the cubic one. Let $x_i i = 0, ..., n$ be a set of n+1 points, we can construct a cubic spline with n piecewise cubic polynomials between the data points. The spline function:

$$S(x) = \begin{cases} S_0(x), & x \in [x_0, x_1] \\ S_1(x), & x \in [x_1, x_2] \\ & \dots \\ S_{n-1}(x), & x \in [x_{n-1}, x_n] \end{cases} \tag{7.2.7}$$

interpolating the function f, has to respect the following constraint:

- the spline has to go exactly trough the data points, $S(x_i) = y_i$
- the splines have to join up, $S_{i-1}(x_i) = S_i(x_i)$, $i = 1, ..., n$
- twice continuous differentiation must be allowed and the derivatives have to be continuous $S'_{i-1}(x_i) = S'_i(x_i)$ and $S''_{i-1}(x_i) = S''_i(x_i)$

Since for one cubic polynomial of degree three, there are four conditions, using n data points we need of $4n$ conditions. However, the interpolating property gives us $n+1$ conditions, and the conditions on the interior data points give us $n+1-2 = n-1$ data points each, summing to $4n-2$ conditions. We require two other conditions. Several choice are possible depending on the problem nature. You can impose first derivative values at the boundary or second derivatives to be zero on boundary (giving rise to natural splines), or chose periodic conditions.

Like polynomial interpolation, spline interpolation incurs a smaller error than linear interpolation and the interpolant is smoother. However, the interpolant is easier to evaluate than the high-degree polynomials used in polynomial interpolation. It also does not suffer from Runge's phenomenon.

7.3 Approximation of functions

A problem closely related with function interpolation is the approximation of a complicated function with simpler functions.

Suppose we know the function but it is too complex to evaluate efficiently. Then we could pick a few known data points from the complicated function, by sampling it in particular points called nodes, and try to interpolate those data points to construct a simpler function. Notice that when approximating a function, the constraint that the interpolant has to go exactly through the data points can be relaxed, introducing other constraints such as to approach the data points as closely as possible (least squares approximation).

In the previous section 7.2 it has been shown that a collection of data points can be interpolated by polynomials according to different techniques. Another possibility consists in approximating the functions with generalized Fourier series, which are, approximations based upon summation of a series of orthogonal polynomials.

$$f(x) = \sum_{n=0}^{N} a_n p_n(x) \tag{7.3.1}$$

where the zeros of the polynomials selected as functional basis define the nodes of interpolation.

Orthogonal polynomials are classes of polynomials $p_n(x)$ defined over a range $[a, b]$ that obey an orthogonality relation:

$$\int_a^b w(x) p_m(x) p_n(x) dx = \delta_{mn} h_n \tag{7.3.2}$$

where $w(x)$ is a weighting function and δ_{mn} is the Kronecker delta. If $h_n = 1$, then the polynomials are not only orthogonal, but orthonormal. Orthogonal polynomials provide a natural way to expand functions on a complete orthogonal basis and have attractive applications in solving differential equations. A table of common orthogonal polynomials is given below, where w(x) is the weighting function and

$$h_n = \int_a^b w(x)[p_n(x)]^2 dx \tag{7.3.3}$$

Polynomials	Interval	$w(x)$	h_n
Legendre Polynomials	[-1,1]	1	$\frac{2}{2n+1}$
Chebyshev Polynomials of the first kind	(-1,1)	$\frac{1}{\sqrt{1-x^2}}$	$\begin{cases} \pi & for\, n = 0 \\ \frac{\pi}{2} & otherwise \end{cases}$
Laguerre Polynomials	$[0, \infty)$	e^{-x}	1
Hermite Polynomials	$(-\infty, \infty)$	e^{-x^2}	$\sqrt{\pi} 2^n n!$

Table 7.1: Orthogonal polynomials commonly used like basis of function expansion (from [25], pp. 774-775)

7.3.1 Choice of basis functions

A good functional basis ϕ_i used to expand the function $f(x)$ must have a number of properties:

- independence of ϕ_i basis functions;
- completeness, which means that any solution can be represented to arbitrarily high accuracy by taking the truncation N to be sufficiently large;
- easiness to compute;
- rapid convergence (uniform convergence) to the original function $f(x)$

Fourier series and orthogonal polynomials satisfy all the previous conditions. The choice of the best basis mainly depends on the geometry of the problem and on boundary conditions. In the table 7.2 the most suitable choice of orthogonal basis according to problem conditions is reported:

Notice that normally, boundary and initial conditions are not a major complication for spectral methods. For example, when the boundary conditions require the solution to be spatially periodic, the sines and cosines of a Fourier

Condition	Basis Set
f(x) is periodic	*Fourier series*
f(x) is periodic and symmetric about $x = 0$	*Fourier cosine*
f(x) is periodic and antisymmetric about $x = 0$	*Fourier sine*
$x \in [a, b]$ and f(x) is non periodic	$\begin{cases} \text{Chebyshev polynomial} \\ \text{Legendre polynomials} \end{cases}$
$y \in [0, \infty]$ and f(x) decays exponentially as $y \to \infty$	Laguerre functions
$y \in [-\infty, \infty]$ and f(x) decays exponentially as $y \to \infty$	*sinc functions* *Hermite functions*
f(x) is defined over a sphere	*Spherical Harmonics*

Table 7.2: Best choice of orthogonal basis for different problem conditions

series (which are the natural basis functions for all periodic problems) automatically and individually satisfy the boundary conditions. For non-periodic problems, if boundary conditions are explicit they have to be expanded in term of the orthogonal basis.

7.4 Numerical integration

Numerical integration constitutes a family of algorithms for calculating an integral using numerical techniques, and by extension, the term describes also the numerical solution of differential equations. The numerical computation of an integral is sometimes called quadrature. There are a wide range of methods available for numerical integration. A good source of information for such techniques is Press et al.(1992)[40]. There are several reasons that may lead to perform a quadrature. The resort to numerical integration is necessary if the integrand is known only at certain points. Otherwise, the analytical formula of the integrand may be known, but it may be difficult or impossible to find an antiderivative that is an elementary function. Moreover, it may be possible to find an antiderivative symbolically, but it may be easier to compute a numerical approximation than to compute the antiderivative, for example when the antiderivative is given as an infinite series or product, or its evaluation requires a special function which is not available. The usual approach in numerical integration consists in converting the integral into a sum, involving the integrand evaluated in certain points multiplied for suitable

weights:

$$\int_a^b f(x)dx = \sum_{i=1}^{N} w_i f(x_i) \tag{7.4.1}$$

A large class of quadrature rules is provided by substituting the integrands with interpolating functions (typically polynomials) which are easy to integrate. Interpolation with polynomials evaluated at equally-spaced points in interval $[a, b]$ yields the Newton-Cotes formulas. The 2- and 3-point formulas are called the trapezoidal rule and Simpson's rule, respectively. The 5-point formula is called Boole's rule. A generalization of the trapezoidal rule is Romberg integration, which can yield accurate results for many fewer function evaluations.

If we allow the intervals between interpolation points to vary, we find another group of quadrature formulas, called Gaussian quadrature formulas. A Gaussian quadrature rule is typically more accurate than any Newton-Cotes rule which requires the same number of function evaluations, if the integrand is smooth.

7.4.1 Newton-Cotes formulas

Newton-Cotes formulas computed 7.4.1, using sampling values of $f(x)$ evaluated at N nodes x_i, which are evenly spaced. The weights w_i are variable and chosen so that the quadrature is exact for all polynomials of order up to $N-1$. A quadrature formula that uses the value of the function $f(x)$ at the endpoint a and b of the integration interval is called *closed*, while the formulas which estimate 7.4.1 using only internal nodes (strictly between a and b) are called *open* formulas.

Let h be the constant step between the abscissas $x_0, x_1, ..., x_N$ so that:

$$x_i = x_0 + ih \quad i = 0, 1, ..., N \tag{7.4.2}$$

In the following some of the most common closed formulas are listed.

Closed formulas

Trapezoidal rule

This rule allow to compute the definite integral of a function $f(x)$ between two points x_1 and x_2 by approximating the function with the straight line passing

trough $f(x_1)$ and $f(x_2)$.

$$\int_{x_1}^{x_2} f(x)dx = h\left[\frac{1}{2}f_1 + \frac{1}{2}f_2\right] \tag{7.4.3}$$

Equation 7.4.3 is a **two points formula**. It is exact for polynomials up to and including degree 1. The result of this rule is, in fact, the area of the trapezoid under the straight line between the abscissas x_1 and x_2.

Simpson's rule

This is a **three point formula**, and evaluate exactly integrals of polynomials up to degree 2:

$$\int_{x_1}^{x_3} f(x)dx = h\left[\frac{1}{3}f_1 + \frac{4}{3}f_2 + \frac{1}{3}f_3\right] \tag{7.4.4}$$

Note that the formula gives the integral over an interval of size 2h, so that the coefficients add up to 2.

The **four and five points formula** are reported below:

Simpson's $\frac{3}{8}$ rule

$$\int_{x_1}^{x_4} f(x)dx = h\left[\frac{3}{8}f_1 + \frac{9}{8}f_2 + \frac{9}{8}f_3 + \frac{3}{8}f_4\right] \tag{7.4.5}$$

Bode's rule

$$\int_{x_1}^{x_5} f(x)dx = h\left[\frac{14}{65}f_1 + \frac{64}{45}f_2 + \frac{24}{45}f_3 + \frac{64}{45}f_4 + \frac{14}{65}f_5\right] \tag{7.4.6}$$

We can use equations 7.4.3, 7.4.4, 7.4.5 and 7.4.6 many times to do the integration over the interval (x_1, x_N) using from time to time the necessary number of nodes, non overlapping, according to the selected rule and then add the results, obtaining composite or extended formulas:

Extended trapezoidal rule

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{1}{2}f_1 + f_2 + f_3 + ... + f_{N-1} + \frac{1}{2}f_N\right] \tag{7.4.7}$$

Extended Simpson's rule

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{1}{3}f_1 + \frac{4}{3}f_2 + \frac{2}{3}f_3 + \frac{4}{3}f_4 + ... + \frac{2}{3}f_{N-2} + \frac{4}{3}f_{N-1} + \frac{1}{3}f_N\right] \tag{7.4.8}$$

7.4.2 Gaussian Quadrature

When the functions are known analytically instead of being tabulated at equally spaced intervals, the best numerical method of integration is called Gaussian quadrature (see Gatteschi [31] and [32]). The Gaussian quadrature is a rule constructed to yield an exact result for polynomials of degree up to $(2N-1)$[3] by a suitable choice of the abscissas x_i and weights w_i for $i = 1, ..., N$. This result is achieved if the function $f(x)$ is well approximated by a polynomial function within the range $[a, b]$ that is to say the function have to be smooth, without singularities in $[a, b]$. However, if the integrated function can be written as $f(x) = W(x)g(x)$, where $g(x)$ is a polynomial or close to polynomial, and $W(x)$ is a weighting function which can be singular, then:

$$\int_a^b f(x)dx = \int_a^b W(x)g(x)dx = \sum_{i=1}^{N} w_i g(x_i) \tag{7.4.9}$$

The fundamental theorem of Gaussian quadrature states that **the optimal abscissas of the N-point Gaussian quadrature formulas are precisely the roots of the orthogonal polynomial for the same interval and weighting function**. Some possible choices of integrating gaussian rules are listed in the table 7.3 below.

Interval	W(x)	Quadrature rule
[-1,1]	1	Gauss-Legendre
(-1,1)	$\frac{1}{\sqrt{1-x^2}}$	Gauss-Chebyshev
$[0, \infty)$	e^{-x}	Gauss-Laguerre
$(-\infty, \infty)$	e^{-x^2}	Gauss-Hermite

Table 7.3: Intervals and weighting function in typical gaussian quadrature rules

[3]The correspondent Newton-Cotes rule on N equally-spaced points integrates precisely the $N-1$ degree polynomial.

Note that the nodes vary in location for different orders of the quadrature rule, so that function evaluations cannot be reused as the order increases.

The rough approach to determine the weights and nodes of a Gaussian quadrature rule for a particular weight function consists of solving $2N$ simultaneous nonlinear equations:

$$\begin{aligned} \sum_{k=1}^{N} w_k f_1(x_k) &= \int_a^b f_1(x)W(x)dx \\ \sum_{k=1}^{N} w_k f_2(x_k) &= \int_a^b f_2(x)W(x)dx \\ &\vdots \\ \sum_{k=1}^{N} w_k f_{2N}(x_k) &= \int_a^b f_{2N}(x)W(x)dx \end{aligned} \tag{7.4.10}$$

where $f_i(x)$ are monomials of order $i-1$. These equations are nonlinear because the unknown nodes appear inside the arguments of the functions $f_i(x)$. Solving such a nonlinear system in general is difficult, so mathematicians developed more sophisticated methods for finding the weights and nodes. The classical approach is based on the theory of orthogonal polynomials.

7.4.3 Orthogonal polynomials and Gaussian quadrature

Using the weight function $W(x)$, the following inner product is defined:

$$\langle f, g\rangle = \int_a^b f(x)g(x)W(x)dx. \tag{7.4.11}$$

A set of orthonormal polynomials $p_n(x)$ satisfies

$$\langle p_m, p_n\rangle = \begin{cases} 1 & m = n \\ 0 & \text{otherwise.} \end{cases} \tag{7.4.12}$$

Such polynomials can be constructed by Gram-Schmidt orthogonalization of the monomials $1, x, x^2, ...$ The goal is to find nodes and weights such that

$$\int_a^b f(x)W(x)dx = \sum_{k=1}^{N} w_k f(x_k) + \epsilon \tag{7.4.13}$$

is exact ($\epsilon = 0$) if $f(x)$ is a polynomial of degree $2N-1$ or less. Let $f(x)$ be a polynomial of degree $2N-1$. Dividing by the orthogonal polynomial $p_N(x)$ gives

$$f(x) = p_N(x)Q_{N-1}(x) + R(x), \tag{7.4.14}$$

where R is a remainder polynomial of degree $N-1$. Multiplying by the weight function and integrating both sides, the following equation is obtained:

$$\int_a^b f(x)W(x)dx = \int_a^b p_N(x)Q_{N-1}(x)W(x)dx + \int_a^b R(x)W(x)dx. \quad (7.4.15)$$

The second term is zero, since Q_{N-1} can be expressed as a linear combination of the orthogonal polynomials $p_0, p_1, ..., p_{N-1}$, and so must be orthogonal to p_N. Hence, it is:

$$\int_a^b f(x)W(x)dx = \int_a^b R(x)W(x)dx. \quad (7.4.16)$$

Now, let the nodes x_k be the N roots of $p_N(x)$; from equation (7.4.14) it can be obtained that $f(x_k) = R(x_k)$. Now a special set of polynomials is introduced, the Lagrange polynomials

$$L_k(x) = \prod_{l=1, l\neq k}^{N} \frac{x - x_l}{x_k - x_l}. \quad (7.4.17)$$

These functions are equal to 1 at $x = x_k$ and 0 at all the other points x_l for $l \neq k$. Because of this property, they are called interpolatory functions. Since $R(x)$ is of order $N-1$, it can be written as a sum of Lagrange polynomials, so that

$$R(x) = \sum_{k=1}^{N} f(x_k)L_k(x). \quad (7.4.18)$$

Combining this expression with equations (7.4.13) and (7.4.16) leads to

$$\int_a^b f(x)W(x)dx = \sum_{k=1}^{N} f(x_k) \underbrace{\left[\int_a^b L_k(x)W(x)dx\right]}_{\text{weights } w_k}, \quad (7.4.19)$$

where the terms in square brackets can be identified as the weights w_k. To simplify the computation of the weights, we can express the Lagrange polynomials in terms of orthogonal polynomials as

$$L_k(x) = \frac{p_N(x)}{(x - x_k)p'_N(x_k)}. \quad (7.4.20)$$

It is easy to see from the way the nodes are chosen that this polynomial is zero at all of the nodes except for x_k, so that the polynomial on the right hand

side has the same $N-1$ zeros as L_k. It is also easy to see that it is equal to one at $x = x_k$. Because a polynomial of order $N-1$ is completely determined by N specified values, the function on the right must be equal to L_k. This leads to the following prescription for the weights and nodes of the Gaussian quadrature rule:

1. The p_n are orthonormal polynomials with respect to the weight function W on the interval $[a, b]$;

2. The nodes are given by the zeros of the $N-$th order orthogonal polynomial: $p_N(x_k) = 0, k = 1, 2, ..., N$;

3. the weights are given by the relation

$$w_k = \int_a^b \frac{p_N(x)}{(x - x_k)p'_N(x_k)} W(x)dx, \tag{7.4.21}$$

or, in a more convenient form

$$w_k = -\frac{A_N}{A_{N-1}} \frac{\langle p_{N-1}, p_{N-1}\rangle}{p_{N-1}(x_k)p'_N(x_k)}, \tag{7.4.22}$$

where A_n is the coefficient of x^n in $p_n(x)$.

A convenient tool to compute the nodes x_i and weights w_i of Gaussian quadrature rules, is the three-term recurrence relation satisfied by the set of orthogonal polynomials associated to the corresponding weight function. If, for instance, p_n is the monic orthogonal polynomial of degree n (the orthogonal polynomial of degree n with the highest degree coefficient equal to one), one can show that such orthogonal polynomials are related through the recurrence relation:

$$p_{n+1}(x) + (B_n - x)p_n(x) + A_n p_{n-1}(x) = 0, \qquad n = 1, 2, \ldots \tag{7.4.23}$$

Nodes and weights can be computed from the eigenvalues and eigenvectors of an associated linear algebra problem. if x_i is a root of the orthogonal polynomial p_n then, using the previous recurrence formula for $k = 0, 1, \ldots, n-1$ and because $p_n(x_i) = 0$, we have:

$$J\tilde{P} = x_i\tilde{P} \tag{7.4.24}$$

where $\tilde{P} = [p_0(x_i), p_1(x_i), ..., p_{n-1}(x_i)]^T$ and J is the so-called Jacobi matrix:

$$J = \begin{pmatrix} B_0 & 1 & 0 & \ldots & \ldots & \ldots \\ A_1 & B_1 & 1 & 0 & \ldots & \ldots \\ 0 & A_2 & B_2 & 1 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & A_{N-2} & B_{N-2} & 1 \\ \ldots & \ldots & \ldots & \ldots & A_{N-1} & B_{N-1} \end{pmatrix} \tag{7.4.25}$$

The nodes of gaussian quadrature can therefore be computed as the eigenvalues of a tridiagonal matrix.

The error of the Gaussian quadrature rule 7.4.9 can be stated as follows:

$$\int_a^b \omega(x)\, f(x)\, dx - \sum_{i=1}^{N} w_i\, f(x_i) = \frac{f^{(2N)}(\xi)}{(2N)!}\,(p_N, p_N) \tag{7.4.26}$$

or some ξ in (a, b), where p_N is the orthogonal polynomial of order N.

7.4.4 Gauss-Legendre quadrature

If $W(x) = 1$ and $[a, b] = [-1, 1]$, the orthogonal polynomials are Legendre polynomials (see Appendix D):

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ &\vdots \\ P_n(x) &= \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2 - 1)^n. \end{aligned} \tag{7.4.27}$$

The nodes can be found using a zero-finding algorithm such as the Newton' s method or the Newton Rapson method. The weights as given by equation (7.4.22) can be simplified using the identities:

$$\int_{-1}^{1} |P_n(x)|^2 dx = \frac{2}{2n+1}, \tag{7.4.28}$$

$$(x^2 - 1)P_n'(x) = nxP_n(x) - nP_{n-1}(x), \tag{7.4.29}$$

$$\frac{A_N}{A_{N-1}} = 2 - \frac{1}{N}. \tag{7.4.30}$$

Hence, the weights become

$$w_k = \frac{2}{(1 - x_k^2)[(P_N'(x_k)]^2}, \tag{7.4.31}$$

which is convenient for numerical evaluation.

A helpful fact in implementing a root finder for the Legendre polynomials is that the zeros of P_{n+1} are in between those of $P_n(x)$, so the zeros of one

value of n can be used as starting guesses for the zeros of P_{n+1}. Alternately, the approximation

$$x_k \simeq cos[\pi(k - 1/4)/(N + 1/2)] \tag{7.4.32}$$

can be used. To compute the derivative of $P_n(x)$, a central difference approximation could be used, but it is better to use the recursion relation (7.4.23).

Note that Gauss-Legendre quadrature can be extended to any integration interval [a,b], by introducing the new variable ξ, such that

$$x = \frac{b-a}{2}\xi + \frac{a+b}{2}. \tag{7.4.33}$$

Hence, we can write

$$\int_a^b f(x)dx = \frac{b-a}{2}\int_{-1}^{1} f(\xi)d\xi \tag{7.4.34}$$

and compute the nodes and the weights according to the classical Gauss-Legendre procedure. The nodes and the weights of the integration over the interval $[a, b]$ can then be derived by applying the following relations:

$$x_k = x(\xi_k) \quad \text{and} \quad \hat{w}_k = \frac{b-a}{2}w_k. \tag{7.4.35}$$

7.4.5 Gauss-Chebyshev quadrature

If $W(x) = \frac{1}{\sqrt{1-x^2}}$ and $[a, b] = [-1, 1]$, the orthogonal polynomials are Chebyshev polynomials (see Appendix C):

$$\begin{array}{ll} T_0(x) & = 1 \\ T_1(x) & = x \\ T_2(x) & = (2x^2 - 1) \\ T_n(x) & = arccos(ncosx) \end{array} \tag{7.4.36}$$

Nodes x_i and weights w_i for such polynomials are analytic:

$$x_i = \cos\left(\frac{\pi}{2}\frac{2i-1}{n}\right), \quad i = 1, \ldots, n. \tag{7.4.37}$$

and

$$w_i = \frac{\pi}{n} \tag{7.4.38}$$

Gauss-Chebyshev quadrature can be applied to any integration interval [a,b], by introducing the new variable ξ, such that

$$x = \frac{b-a}{2}\xi + \frac{a+b}{2}. \tag{7.4.39}$$

Hence, we can write:

$$\int_a^b f(x)dx = \frac{b-a}{2}\int_{-1}^{1} f(\xi)d\xi \tag{7.4.40}$$

and compute the nodes and the weights according to the classical Gauss-Chebyshev procedure.

7.4.6 Gaussian Quadrature over an unit sphere

The quadrature rule presented above refers to one-dimensional integration. In order to apply the Gaussian integration for surface integrals on a unit sphere, we need to separate the two-dimensional integration of the angular flux into two one-dimensional integrations (see Atkinson (1982) [26], Keiner et al.(2008) [35] and Graf et al.(2009) [33]).

Let $f(\theta, \varphi)$ be a real function defined on a unit sphere, where θ, is the polar angle, $0 \leq \theta \leq \pi$ and φ, is the azimutal angle, $0 \leq \varphi \leq 2\pi$. We want to estimate:

$$I(f) = \int\limits_{\Omega} f(\theta, \varphi)\, d\Omega = \int\limits_0^{\pi}\int\limits_0^{2\pi} f(\theta, \varphi)\sin\theta\, d\theta d\varphi \tag{7.4.41}$$

Since the weighting function for 7.4.41 is $W(x) = 1$ the integral is approximated by:

$$I_N(f) = \frac{\pi}{N}\sum_{j=1}^{2N}\sum_{i=1}^{N} w_i f(\theta_i, \varphi_j) \tag{7.4.42}$$

where the nodes θ_i are chosen so that $x_i = cos(\theta_i)$ are the Gauss-Legendre nodes and w_i are the Gauss-Legendre weights on $[-1, 1]$. The points φ_j are evenly spaced in $[0, 2\pi]$ with spacing $\frac{\pi}{N}$.

With this choice of nodes points and weights, $I_m(f)$ integrates exactly any polynomial $f(\theta, \varphi)$ up to degree $2N$.

Integrations in θ and φ directions can be calculated in the meantime or in sequence. Particularly, if function $f(\theta, \varphi)$ involve spherical Harmonics (see appendix B), in θ direction we use a Gauss-Legendre quadrature, while in φ direction we use a discrete Fourier transform.

7.5 Discrete Transforms

Numerical solution of PDEs requires the functions appearing in the equations to be discretized with different techniques according to the method chosen to solve the equations. In particular, spectral techniques substitute the original problem for the functions unknown with a sequence of problem for the coefficients of the expansion. Such coefficients can be found by suitable transforms, which utilize the orthogonality properties of the basis expansion to extract an expression for the coefficients. The transform can be also inverted and used to reconstruct the functions starting from the coefficients in terms of discrete values .

7.5.1 Discrete Fourier Transform

The discrete Fourier transform (DFT) is a kind of Fourier transform that transforms one function into another, which is called the frequency domain representation, or simply the DFT, of the original function. The DFT requires an input function that is discrete and whose non-zero values have a limited (finite) duration.

Consider a complex continuous function $f(x)$ sampled at $N+1$ evenly spaced time steps $x_k = 0, 1, \ldots, N$. Denoting with Δ the interval between two consecutive samples, its reciprocal $\frac{1}{\Delta}$ is called *sampling rate*.

The sequence of the $N+1$ values assumed by the function $f(x)$ at the steps $x_k = 0, 1, \ldots, N$ can be expressed by means of the relation:

$$f(x_k) = \frac{1}{N+1} \sum_{n=0}^{N} F_n e^{-\frac{2\pi i}{(N+1)} kn} \tag{7.5.1}$$

in which the coefficients F_n are the amplitudes of the modes in the frequency domain. Equation 7.5.1 is called *Inverse Discrete Fourier Transform*. F_n coefficients can be evaluated by means of the *Discrete Fourier Transform*, according to the formula:

$$F_n = \sum_{k=0}^{N} f(x_k) e^{\frac{2\pi i}{(N+1)} kn} \tag{7.5.2}$$

F_n coefficients are derived from the 7.5.1 using the discrete orthogonality properties of Fourier series. This transform, does not depend on any dimensional parameter, such as Δ and maps $N+1$ complex number (the $f(x_k)$) into $N+1$ complex numbers (the F_n). If the function $f(x)$ is real, the imaginary parts of 7.5.1 are identically equal to zero, producing an output which is half

redundant. In fact, the resulting transform F_n with $n = 0, 1, \ldots, N$ satisfies $F_{N-n} = F_n^*$, being F_n^* the complex conjugate. In this case the real part of the coefficients can be stored in a complex array of half dimension, alternatively it is possible to reformulate the Discrete Fourier Transform and its Inverse Transform as a sum of sine and cosine functions with real coefficients using Euler's formula:

$$f(x_k) = \sum_{n=1}^{N} F_n^c \cos\left[\frac{\pi}{N+1} nk\right] + \sum_{n=1}^{N} F_n^s \sin\left[\frac{\pi}{N+1} nk\right] + F_0^c \tag{7.5.3}$$

$$F_n^c = \frac{2}{N+1} \sum_{k=1}^{N} f(x_k) \cos\left[\frac{\pi}{N+1} nk\right] \tag{7.5.4}$$

$$F_n^s = \frac{2}{N+1} \sum_{n=1}^{N} f(x_k) \sin\left[\frac{\pi}{N+1} nk\right] \tag{7.5.5}$$

$$F_0^c = \frac{1}{N+1} \sum_{k=1}^{N} f(x_k) \tag{7.5.6}$$

7.5.2 Discrete Legendre Transform

A real continuous function $f(x)$ defined on the interval $[-1, 1]$ can be expanded in series of *Legendre polynomials*:

$$f(x) = \sum_{i=0}^{\infty} l_i P_i(x) \tag{7.5.7}$$

The expansion coefficients l_i of 7.5.7 can be calculated using the orthogonality properties of Legendre polynomials obtaining:

$$l_i = \frac{2N+1}{2} \int_{-1}^{1} f(x) P_i(x) dx \tag{7.5.8}$$

In practical application the series 7.5.7 reduces to a summation on a finite number of element:

$$f_N(x_k) = \sum_{i=0}^{N} l_i P_i(x_k) \tag{7.5.9}$$

and it is called *Inverse Discrete Legendre Transform*. The coefficients l_i are computed by means of a Gaussian quadrature of 7.5.8:

$$l_i = \frac{2N+1}{2} \sum_{k=0}^{N} f(x_k) P_i(x_k) w_k \tag{7.5.10}$$

in which x_k are the Legendre nodes, which are the zeros of P_{N+1}, contained inside the interval $(-1, 1)$, and the w_k are the Legendre weights. 7.5.10 is the *Discrete Legendre transform* DLT and allows to transfer $N + 1$ values in the real domain to $N + 1$ amplitudes of the modes of expansion in the Legendre basis.

A similar procedure have to be followed if the function is expanded in terms of *Associated Legendre Functions* $P_n^m(x)$. In this case, fixing the value of the m index:

$$f_m(x) = \sum_{i=0}^{N} l_i P_i^m(x) \tag{7.5.11}$$

the coefficients l_i are easily calculated using the orthogonality relation for the Associated Legendre Functions leading to:

$$l_i = \frac{2N+1}{2} \frac{(n-m)!}{(n+m)!} \sum_{k=0}^{N} f(x_k) P_i^m(x_k) w_k \tag{7.5.12}$$

where x_k and w_k are the Legendre nodes and weights respectively.

7.5.3 Discrete Chebyshev Transformation

Let us consider a Chebyshev expansion:

$$f(x) = \sum_{i=0}^{\infty} c_i T_i(x) \tag{7.5.13}$$

and its truncated series $f_N(x)$:

$$f_N(x) = \sum_{i=0}^{N} c_i T_i(x) \tag{7.5.14}$$

approximation of the continuous function $f(x)$, defined on the interval $[-1, 1]$. $f_n(x)$ is the truncated series of Chebyshev of $f(x)$ and it coincides with $f(x)$ at the Chebyshev nodes, defined as:

$$x_k = cos(\theta_k) = \cos\left(\frac{(2k+1)\pi}{2(N+1)}\right) \quad k = 0, 1, \ldots, N \tag{7.5.15}$$

The coefficients c_i of 7.5.14 expansion are the amplitude of the modes of the function $f(x)$ in the Chebyshev polynomials basis. They can be easily computed adopting the discrete orthogonality procedure leading to:

$$c_i = \frac{2}{N} \sum_{k=0}^{N} f(x_k) T_i(x_k) = \frac{2}{N} \sum_{k=0}^{N} f(\cos(\theta_k)) \cos(i\theta_k) \tag{7.5.16}$$

which is the *Discrete Chebyshev Transform* DCT. Thanks to its trigonometric structure DCT is similar to DFT and both can be evaluated by means of FFT algorithm. The 7.5.14 is the *Inverse Discrete Chebyshev Transform* allowing to reconstruct the original function starting from the coefficients c_i . See [30] for a wide treatment of Chebyshev polynomials in numerical analysis.

7.5.4 Spherical Harmonics Transform

A generic real scalar function $f(\theta, \varphi)$ defined over a unit sphere can be expanded in spherical harmonics (see appendix B):

$$f(\theta,\varphi) = \sum_{l=0}^{nl} \sum_{m=-l}^{l} f_{lm} Y_{lm}(\theta,\varphi) = \tag{7.5.17}$$
$$= \sum_{l=0}^{nl} \sum_{m=-l}^{l} f_{lm} \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\varphi}$$

Notice that expanding $f(\theta, \varphi)$ on SH basis, its θ and ϕ dependence has been separated. The expansion coefficients f_{lm} of 7.5.17, can be calculated using the orthogonality properties of spherical harmonics, which involves a product Gaussian integration according to 7.4.42.

The structure of spherical harmonics suggest to treat the θ and ϕ terms separately and in sequence. A scheme of the transform procedure is reported below:

$$\begin{array}{c} \text{DFT on } f(\theta,\varphi) \\ \Downarrow \\ f_m^c(\theta) f_m^s(\theta) f_0^c(\theta) \\ \text{DLT on } f_m(\theta) \\ \Downarrow \\ f_{lm} \end{array}$$

Firstly, let us isolate the φ dependence:

$$
\begin{aligned}
f(\theta,\varphi) &= \sum_{l=0}^{nl}\sum_{m=-l}^{l} f_{lm}Y_{lm}(\theta,\varphi) = \qquad (7.5.18)\\
&= \sum_{l=0}^{nl}\sum_{m=-l}^{l} \underbrace{f_{lm}\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\varphi}}_{f_m(\theta)} = \\
&= \sum_{m=-nl}^{nl} f_m(\theta)e^{im\varphi} = \\
&= \sum_{m=1}^{nl} f_m^c(\theta)\cos(m\varphi) + \sum_{m=1}^{nl} f_m^s(\theta)\sin(m\varphi) + f_0^c(\theta)
\end{aligned}
$$

Since the function $f(\theta,\varphi)$ is real, and for spherical harmonics properties $f_{-m} = f*_m$ (the star indicates the complex conjugate), negative values of m index are redundant. Thus, using Euler's formula we replace the summation of imaginary exponential function with real sine and cosine functions sums extended up to half terms. Now, the coefficients $f_m^c(\theta)$, $f_m^s(\theta)$ and $f_0^c(\theta)$ can be evaluated by means of a Discrete Fourier Transform.

Coefficients f_{lm} are then calculated using a Discrete Legendre Transform in θ direction, for $f_m^c(\theta)$, $f_m^s(\theta)$ and $f_0^c(\theta)$ functions.

Function $f(\theta,\varphi)$ can be reconstructed using 7.5.18, reminding that:

$$f_m^c(\theta) = \sum_{l=0}^{nl} f_{lm}\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta) \qquad (7.5.19)$$

$$f_m^s(\theta) = \sum_{l=0}^{nl} f_{lm}\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta) \qquad (7.5.20)$$

$$f_0^c(\theta) = f_{l0}\sqrt{\frac{(2l+1)}{4\pi}}P_l^0(\cos\theta) \qquad (7.5.21)$$

Chapter 8

Problem formulation

This chapter provides an overview of the governing equations for the problem of the flow of a viscous fluid inside a bounded surface. Several equivalent formulations are presented. Among these, the time discretized primitive variables formulation when the incompressibility constraint is replaced by the Poisson equation is critically examined following Quartapelle (1993) [43]. In order to obtain a system of split-equations, condition of integral character are derived for the pressure according to Quartapelle et al. (1986) [42].

8.1 Incompressible Navier-Stokes equations

The motion of a viscous incompressible fluid is governed by the Navier-Stokes equations:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{\nabla \mathrm{p}}{\rho} + \nu \nabla^2 \mathbf{u} \tag{8.1.1}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{8.1.2}$$

in which $\mathbf{u}$ is the velocity, p the pressure and ν the kinematic viscosity. This "classical" formulation made in the primitive variable $\mathbf{u}$ and p can be suitably replaced by an equivalent expression in term of the two scalar unknowns vorticity ζ and stream function ψ, where:

$$\zeta = \nabla \times \mathbf{u} \tag{8.1.3}$$

Using non-primitive variables several equivalent formulations are possible, for example:

$$\frac{\partial \zeta}{\partial t} + \nabla \times (\zeta \times \nabla \times \psi) = \nu \nabla^2 \zeta \tag{8.1.4}$$

$$-\nabla^2\psi = \zeta \tag{8.1.5}$$

The choice of the formulation (in primitive/non-primitive variables, intermediate character formulation) depends on the particular numerical approach chosen to solve the problem, together with the space and time discretization method.

Working with primitive variables, a convenient formulation of the problem involves the use of Poisson equation 8.1.6 in substitution of the continuity equation:

$$\nabla^2 p = f \tag{8.1.6}$$

where f is a real function.

The convenience of this approach consists in the possibility to derive uncoupled formulation of the incompressible equations as later better explained in section 8.2.

To make the formulation of the problem complete then, you have to specify boundary and initial conditions. *Boundary conditions* are often of a no-slip kind. These conditions consist in prescribing the value of the velocity at the boundary:

$$\mathbf{u}\,|_S = \mathbf{v}_S(\mathbf{x}_S, t) \tag{8.1.7}$$

where S is the boundary of the volume V occupied by the fluid and $\mathbf{x}_S \in S$. If the boundary of the fluid domain is a solid wall, $\mathbf{v}_S$ is the velocity of the wall. Notice that no boundary condition for the pressure is prescribed and it would be incorrect to impose one together with the velocity boundary condition for the reasons explained later. In some applications, because of the particular shape of the fluid domain inflow or outflow boundaries, velocity boundary condition different from 8.1.7 are convenient. In these situations the pressure can be supplemented by boundary conditions of Neumann or Dirichlet type.

Finally, the initial condition consists in the specification of the velocity field $\mathbf{v}_o$ at the initial time $t = 0$

$$\mathbf{u}\,|_{t=o} = \mathbf{v}_o(x) \tag{8.1.8}$$

Moreover, the incompressibility constraint implies that the initial condition and the boundary conditions cannot be assigned in an arbitrary and independent way. The boundary velocity $\mathbf{v}_S$ must satisfy for all $t > 0$ the global condition:

$$\oint \mathbf{n} \cdot \mathbf{v}_S dS = 0 \tag{8.1.9}$$

following from the integration of the continuity equation 8.1.2 over V using the divergence theorem[1].

[1]The symbol $\oint$ denotes the integral over the closed surface S

In 8.1.9 and in the following, $\mathbf{n}$ denotes the outward unit normal to the boundary S. The initial velocity field $\mathbf{v}_o$ is assumed to be solenoidal like any instantaneous velocity field:

$$\nabla \cdot \mathbf{v}_o = 0 \tag{8.1.10}$$

Finally, the following compatibility conditions are stated for the boundary and the initial conditions $\mathbf{v}_S$ and $\mathbf{v}_o$:

$$n \cdot v_S \,|_{t=o} = n \cdot v_o \,|_S \tag{8.1.11}$$

The set of equations 8.1.1,8.1.2, with the initial and boundary conditions 8.1.8, 8.1.7, 8.1.9, 8.1.10 and the compatibility condition 8.1.11, admits solutions $(\mathbf{u}(\mathbf{x}, t), p(\mathbf{x}, t))$ with the pressure field determined up to an arbitrary additive function of time.

8.2 Time discretized pressure-velocity equations

Introducing the characteristic length L and velocity $\mathbf{U}$, we consider the problem in a dimensionless form with variables scaled such those[2]:

$$\mathbf{x} = \frac{\mathbf{x}^*}{L} \tag{8.2.1}$$

$$\mathbf{u} = \frac{\mathbf{u}^*}{\mathbf{U}} \tag{8.2.2}$$

$$t = t^* \frac{\mathbf{U}}{L} \tag{8.2.3}$$

$$p = \frac{p^*}{\rho \mathbf{U}^2} \tag{8.2.4}$$

Reminding that the ratio $\frac{\nu}{L\mathbf{U}}$ indicates the Reinolds number, the dimensionless equations take the following form:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\,\mathbf{u} = -\nabla p + \frac{1}{\mathrm{Re}} \nabla^2 \mathbf{u} \tag{8.2.5}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{8.2.6}$$

The governing equations 8.2.5 are discretized in time by mean of a non-fractional-step scheme. Let us indicate the unknown velocity and pressure fields at the new time level with $\mathbf{u} = \mathbf{u}^n + 1$ and $p = p^n + 1$. The non-linear

[2]Here the superscript * indicates the dimensional variables.

term in 8.1.1 is evaluated in an explicit manner with the three-level Adams-Bashforth method:

$$(\mathbf{u}\cdot\nabla)\,\mathbf{u} = \frac{3}{2}\,(\mathbf{u}^n\cdot\nabla)\,\mathbf{u}^n - \frac{1}{2}\left(\mathbf{u}^{n-1}\cdot\nabla\right)\mathbf{u}^{n-1} \tag{8.2.7}$$

The linear terms are taken in account in a fully implicit manner using the Crank-Nicolson method:

$$\nabla p = \frac{1}{2}\nabla\left(p^{n+1}+p^n\right) \tag{8.2.8}$$

$$\nabla^2\mathbf{u} = \frac{1}{2}\nabla^2\left(\mathbf{u}^{n+1}+\mathbf{u}^n\right) \tag{8.2.9}$$

$$\nabla\cdot\mathbf{u} = \frac{1}{2}\nabla\cdot\left(\mathbf{u}^{n+1}+\mathbf{u}^n\right) \tag{8.2.10}$$

where the superscript n , $n-1$ and $n+1$ in the 8.2.7 and 8.2.8-10 are the values of the variables $\mathbf{u}$ and p at the corresponding time steps. Adopting 8.2.7 and 8.2.8-10 discretizing schemes we have:

$$\left(-\nabla^2+\gamma\right)\mathbf{u} = -\nabla p + \mathbf{f} \tag{8.2.11}$$

$$\nabla\cdot\mathbf{u} = 0 \tag{8.2.12}$$

$$\mathbf{u}\,|_S = \mathbf{v}_S(\mathbf{x}_S,t) \tag{8.2.13}$$

in which, $\gamma = 1/\nu\Delta t$, while the source term $\mathbf{f}$ is given by:

$$\mathbf{f} = \left(\nabla^2+\gamma\right)\mathbf{u}^n - \mathrm{Re}\left[3\,(\mathbf{u}^n\cdot\nabla)\,\mathbf{u}^n - \left(\mathbf{u}^{n-1}\cdot\nabla\right)\mathbf{u}^{n-1}\right] \tag{8.2.14}$$

This formulation of the problem is usually referred to as generalized or time-discretized unsteady Stokes problem.

Taking the divergence of the momentum equation 8.2.11 and using the continuity equation 8.1.2 we derive the Poisson equation for the pressure:

$$\nabla^2 p = \nabla\cdot\mathbf{f} \tag{8.2.15}$$

This equation can be introduced in order to eliminate the continuity equation but unfortunately, due to the lack of proper boundary conditions for p, is not completely equivalent with the incompressibility condition. In fact, the fulfillment of the Poisson equation is only a *necessary condition* for $\mathbf{u}$ to be solenoidal. In order to ensure the mass conservation a constraint imposing the equivalence of the primitive variable system containing the Poisson equation for pressure and the original Navier-Stokes problem is required. However,

Kleiser and Schumann (1980) demonstrated that, if we impose the boundary condition $\nabla \cdot \mathbf{u}\,|_S = 0$, then the harmonicity of $\nabla \cdot \mathbf{u}$ is guaranteed in the whole domain. The resulting time-discretized Navier-Stokes problem consists of the elliptic boundary value problem below:

$$\left(-\nabla^2 + \gamma\right)\mathbf{u} = -\nabla p + \mathbf{f} \tag{8.2.16}$$

$$\nabla^2 p = \nabla \cdot \mathbf{f} \tag{8.2.17}$$

$$u\,|_S = v_S \qquad \nabla \cdot u\,|_S = 0 \tag{8.2.18}$$

Besides these equations also the following compatibility constraints have to be satisfied:

$$\oint \mathbf{n} \cdot \mathbf{v}_S dS = 0 \tag{8.2.19}$$

$$n \cdot v_S\,|_{t=o} = n \cdot v_o\,|_S \tag{8.2.20}$$

According to the above scheme the solution of the time-dependent Navier-Stokes problem is reduced to the solution of a sequence of unsteady Stokes problem. The main difficulty encountered in the solution of the above system is due to the absence of boundary conditions on the pressure. In fact, in the form written above, because of the lack of independent boundary condition for the pressure, problem 8.2.16 is a coupled system of equations, that is to say that pressure and velocity have to be solved simultaneously.

8.3 Pressure integral conditions

The incompressibility constraint impose that the velocity field must be solenoidal everywhere and at every time step. Nevertheless, the motion of a fluid is governed by the complete Navier-Stokes equations, which, in general, don't ensure the velocity field to be divergence free. The assumption of the incompressibility of the flow is in fact an external constraint imposed on the governing equation. It can be accomplished only by considering the pressure as a free variable necessary to satisfy this constraint. This implies that the pressure is deprived of its thermodynamic meaning. From this point of view the incompressibility constraint constitutes a violation of the thermodynamic principles, nevertheless, the incompressible Navier -Stokes equations can be suitably used to describe a wide class of fluid dynamic phenomena when thermodynamic properties of the fluid are negligible.

Ultimately, lack of a boundary condition for the pressure prevents a direct solution of the problem as a system of split elliptic equations, to be integrated

in sequence. However, proper condition for p are obtained by considering the vector Green identity for the metaharmonic (Helmholtz) elliptic operator $\nabla^2_\gamma = (\nabla^2 - \gamma)$:

$$\begin{aligned} &\int \big(\mathbf{u}\cdot\nabla^2_\gamma\mathbf{v} - \mathbf{v}\cdot\nabla^2_\gamma\mathbf{u}\big)dV = \\ &= \int (\mathbf{n}\cdot\mathbf{u}\nabla\cdot\mathbf{v} - \mathbf{n}\cdot\mathbf{v}\nabla\cdot\mathbf{u} + \mathbf{n}\times\mathbf{u}\cdot\nabla\times\mathbf{v} - \mathbf{n}\times\mathbf{v}\cdot\nabla\times\mathbf{u})dS \end{aligned} \tag{8.3.1}$$

in which $\mathbf{n}$ is the outward normal unit vector on S. Now, since $-\nabla p + \mathbf{f} = (-\nabla^2 + \gamma)\,\mathbf{u} = -\nabla^2_\gamma u$ the Green identity 8.3.1 implies that the pressure field satisfies the following integral condition:

$$\int (-\nabla p + \mathbf{f})\, b\, dV = \int (\mathbf{n}\cdot\mathbf{v}_S\nabla\cdot\mathbf{b} + \mathbf{n}\times\mathbf{v}_S\cdot\nabla\times\mathbf{b})dS \tag{8.3.2}$$

for any vector $\mathbf{b}$ solution of the metaharmonic (Helmholtz) problem:

$$\big(-\nabla^2 + \gamma\big)\, b = 0 \qquad \int n\cdot b dS = 0 \qquad n\times b\,|_S = 0 \tag{8.3.3}$$

Notice that the explicit treatment of the advection term in Navier-Stokes produces equations with an independent boundary condition for the pressure not involving the unknown velocity field u^{n+1}. The number of independent fields $\mathbf{b}$ that are solutions of 8.3.3 is equal to that of the boundary points. These conditions allow to reformulate the problem 8.2.16 in split form:

$$\nabla^2 p^{n+1} = -\nabla\cdot\mathbf{f}^n \tag{8.3.4}$$

$$-\int \nabla p^{n+1}\cdot\mathbf{b}\, dV = -\int \mathbf{f}^n\cdot\mathbf{b}\, dV + \int \big(\mathbf{n}\cdot\mathbf{v}_S^{n+1}\nabla\cdot\mathbf{b} + \mathbf{n}\times\mathbf{v}_S^{n+1}\cdot\nabla\times\mathbf{b}\big)dS \tag{8.3.5}$$

$$\big(\nabla^2 - \gamma\big)\,\mathbf{u}^{n+1} = -\nabla p^{n+1} + \mathbf{f}^n \tag{8.3.6}$$

$$\mathbf{u}^{n+1}\,|_S = \mathbf{v}_S^{n+1} \tag{8.3.7}$$

Equations 8.3.4 constitute a general factorized form of the Navier-Stokes equations for the primitive variables and can be solved in sequence. The solution proceeds as follows:

1. Solution of the metaharmonic problem in order to find vector field $\mathbf{b}$ values to introduce in the integral condition for the pressure;

2. Solution of the Poisson equation with the integral condition for the pressure;

3. Solution of the momentum equation supplemented by the boundary condition for the velocity of ordinary Dirichlet type.

The imposition of the integral condition for the pressure assures on one hand that the pressure field is compatible with the no-slip velocity condition given on the solid boundary and, on the other hand, that the velocity field at the time step $n+1$ is exactly solenoidal, as required by the incompressibility condition.

8.4 Modal formulation of the unsteady Stokes problem

Let us consider the problem of a fluid moving in the region between two concentric spheres of radius r_1 and r_2. The geometric shape of the fluid domain suggests as natural, the choice of a spherical system of coordinates (r, θ, φ)[3]. Pressure and Velocity variables are expanded in series of scalar and vector spherical harmonics respectively, which constitute a convenient orthogonal bases in this reference system:

$$p(r,\theta,\varphi) = \sum_{l=0}^{L} \sum_{m=-l}^{l} p_{lm}(r)\, Y_{lm}(\theta,\varphi) \tag{8.4.1}$$

$$\mathbf{u}(r,\theta,\varphi) = \sum_{l=0}^{L} \sum_{m=-l}^{l} \left[u^r_{lm}(r)\, \mathbf{P}_{lm}(\theta,\varphi) + u^\theta_{lm}(r)\, \mathbf{B}_{lm}(\theta,\varphi) + u^\varphi_{lm}(r)\, \mathbf{C}_{lm}(\theta,\varphi) \right] \tag{8.4.2}$$

The functions $Y_{lm}(\theta,\varphi)$ appearing in 8.4.1 are the scalar orthonormal spherical harmonics, while the vectorial functions $(\mathbf{P}_{lm}(\theta,\varphi), \mathbf{B}_{lm}(\theta,\varphi), \mathbf{C}_{lm}(\theta,\varphi))$ are the vector spherical harmonics[4]. Notice that this expansion permits to separate the radial dependence (present only in the expansion coefficients) from the angular one (in the expanding basis functions. Moreover, this operation allows the transformation of the original three-dimensional partial differential equations into a set of ordinary differential equations for the expansion coefficients, which are function only of the radial variable. In a similar way, the integral conditions for the pressure transforms into one-dimensional definite

[3]For formulation of Navier-Stokes equations in spherical coordinates the reader is referred to the chapter 2 in the first part of this Thesis, while a collection of differential operator written in spherical coordinate can be consulted in appendix A

[4]Scalar and vector spherical harmonics definitions, properties and applications are discussed in appendixB

integrals on the radial variable. According to the expressions of the differential operators for scalar and vector fields expanded in spherical harmonics contained in appendix B the time-discretized Stokes problem 8.3.4 then transforms into the following system of equations:

$$D_l p_{lm} = \left(\frac{d}{dr} + \frac{2}{r}\right) f^r_{lm} - \frac{s_l}{r} f^\theta_{lm} \tag{8.4.3}$$

$$-\int_{r1}^{r2} \left(\frac{dp_{lm}}{dr} b^r + \frac{s_l}{r} p_{lm} b^\theta\right) r^2 dr = -\int_{r1}^{r2} \left(f^r_{lm} b^r + f^\theta_{lm} b^\theta\right) r^2 dr + \left[v^r_{Slm} r^2 \left(\frac{d}{dr} + \frac{2}{r}\right) b^r - v^\theta_{Slm} r^2 \left(\frac{s_l}{r} b^r - \frac{d}{dr} b^\theta\right)\right]\Bigg|_{r1}^{r2} \tag{8.4.4}$$

$$\left(-\mathbf{D}_l + \gamma \mathbf{I}\right) \mathbf{u}_{lm} = -\nabla p_{lm} + \mathbf{f}_{lm} \tag{8.4.5}$$

$$\mathbf{u}_{lm}(r_1) = \mathbf{v}_{Slm}(r_1) \qquad \mathbf{u}_{lm}(r_2) = \mathbf{v}_{Slm}(r_2) \tag{8.4.6}$$

where:

$$D_l = \frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} \quad l = 0, 1, 2, \ldots \tag{8.4.7}$$

$$\mathbf{D}_l = \begin{pmatrix} D_l - \frac{2}{r^2} & 2\frac{S_l}{r^2} & 0 \\ 2\frac{S_l}{r^2} & D_l & 0 \\ 0 & 0 & D_l \end{pmatrix} \tag{8.4.8}$$

In the equations above the subscript lm indicate the coefficients of 8.4.1 and 8.4.2 expansions; moreover, the terms $f_{lm}(r)$ and $v_{Slm}(r)$ are the expansion coefficients of the source term $\mathbf{f}(r, \theta, \phi)$ and of the boundary velocity $\mathbf{v}_S(r, \theta, \phi)$ respectively.

8.4.1 Solution of the modal equations

The solution of the modal incompressible Navier-Stokes problem consists in the research of the expansion coefficients of pressure and velocity. The scalar unknown $p_{lm}(r)$, can be determined independently and before the velocity $\mathbf{u}_{lm}(r)$, as solution of the modal Poisson problem in addition with its integral boundary condition. In the general case there are two of such conditions provided by the solutions of the metaharmonic problem as better explained in the following. The problem defining $\mathbf{u}_{lm}(r)$ is instead a standard two-points boundary value problem of a vector type. The first two components u^r_{lm} and u^θ_{lm} are coupled together because of the structute of the differential operator $\mathbf{D}_l$, whereas the u^φ_{lm} component is uncoupled. This latter component is independent from the pressure and has only the role to ensure the fulfillment of

the incompressibility condition. As far the unexpanded differential problem, the solution of the modal equations proceed as follows:

1. Solution of the metaharmonic problem in order to find vector field **b** values to introduce in the integral condition for the pressure;

2. Solution of the Poisson equation with the integral condition for the pressure;

3. Solution of the momentum equation supplemented by the boundary condition for the velocity of ordinary Dirichlet type.

8.4.2 Solution of the metaharmonic problem

In the general case, the metaharmonic differential problem has two linearly independent solutions according with the number of solid boundaries physically existing in the problem. This will lead to the formulation of two different integral conditions for the pressure problem. The differential operator $\mathbf{D}_l$ appearing in the metaharmonic problem depends only on the index l, and the right hand side of the equation is homogeneous. It follow that the problem is independent on the m index and we obtain a distinct differential problem for each value of l, irrespective of the value of m. The modal equation of the metharmonic problem in the vector form, for $l > 0$ is the following:

$$(\mathbf{D}_l + \gamma I)\, \mathbf{b}_l(r) = 0 \tag{8.4.9}$$

to solve alternatively with the following set of linearly independent boundary conditions:

$$1)\quad b_l^r(r_1) = 1 \quad b_l^r(r_2) = 0 \quad b_l^\theta(r_j) = 0 \quad b_l^\varphi(r_j) = 0 \quad j = 1,2 \tag{8.4.10}$$

$$2)\quad b_l^r(r_1) = 0 \quad b_l^r(r_2) = 1 \quad b_l^\theta(r_j) = 0 \quad b_l^\varphi(r_j) = 0 \quad j = 1,2 \tag{8.4.11}$$

The problem 8.4.9 can be written in its components:

$$\left(D_l - \frac{2}{r^2} - \gamma\right) b_l^r(r) + 2\frac{S_l}{r^2} b_l^\theta(r) = 0 \tag{8.4.12}$$

$$2\frac{S_l}{r^2} b_l^r(r) + (D_l - \gamma)\, b_l^\theta(r) = 0 \tag{8.4.13}$$

$$(D_l - \gamma)\, b_l^\varphi(r) = 0 \tag{8.4.14}$$

Notice that the problem for the φ component is homogeneous with homogeneous boundary condition, so that $b_l^\varphi(r)$ vanishes. The problems for $b_l^r(r)$ and

$b_l^\theta(r)$ are coupled together. Introducing the similarity transformation:

$$\begin{pmatrix} \beta_l^r \\ \beta_l^\theta \end{pmatrix} = S_l \begin{pmatrix} b_l^r \\ b_l^\theta \end{pmatrix} \quad \text{where } S_l = \begin{pmatrix} \left(\frac{l}{2l+1}\right)^{1/2} & \left(\frac{l+1}{2l+1}\right)^{1/2} \\ \left(\frac{l+1}{2l+1}\right)^{1/2} & -\left(\frac{l}{2l+1}\right)^{1/2} \end{pmatrix} \tag{8.4.15}$$

the equations in r and θ components of 8.4.12 can be written in uncoupled form for the unknowns β_l^r and β_l^θ:

$$(D_{l-1} - \gamma)\,\beta_l^r(r) = 0 \tag{8.4.16}$$

$$(D_{l+1} - \gamma)\,\beta_l^\theta(r) = 0 \tag{8.4.17}$$

to solve with the following set of linearly independent boundary conditions:

$$\text{first set} \;\rightarrow\; \beta_l^r(r_1) = \left(\frac{l}{2l+1}\right)^{1/2} \quad \beta_l^r(r_2) = 0 \quad \beta_l^\theta(r_1) = \left(\frac{l+1}{2l+1}\right)^{1/2} \quad \beta_l^\theta(r_2) = 0 \tag{8.4.18}$$

$$\text{second set} \;\rightarrow\; \beta_l^r(r_1) = 0 \quad \beta_l^r(r_2) = \left(\frac{l}{2l+1}\right)^{1/2} \quad \beta_l^\theta(r_1) = 0 \quad \beta_l^\theta(r_2) = \left(\frac{l+1}{2l+1}\right)^{1/2} \tag{8.4.19}$$

An analytic solution of 8.4.16, with the boundary conditions 8.4.18 in term of Bessel functions can be found considering the change of variable:

$$r \rightarrow x = \gamma^{1/2} r \quad \beta(r) \rightarrow B(r) = r^{1/2}\beta(r) = \gamma^{-1/4} x^{1/2} \beta(\gamma^{-1/2} x) \tag{8.4.20}$$

With the above transformation the equations 8.4.16 become:

$$\mathbf{B}_{l-1/2} B_l^r(x) = 0 \tag{8.4.21}$$

$$\mathbf{B}_{l+3/2} B_l^\theta(x) = 0 \tag{8.4.22}$$

where $\mathbf{B}$ is the modified Bessel operator of order $l + 1/2$ for the variable x:

$$\mathbf{B}_{l+1/2} = x^2 \frac{d^2}{dx^2} + x \frac{d}{dx} - \left[x^2 + (l + 1/2)^2\right] \tag{8.4.23}$$

Being $I_\alpha(x)$ and $K_\alpha(x)$ the modified Bessel functions of the first and second kind respectively, solution of the modified Bessel function of order α, the solution of 8.4.21 problem are the following:

$$B_l^r(x) = a_l I_{l-1/2}(x) + b_l K_{l-1/2}(x) \tag{8.4.24}$$

$$B_l^\theta(x) = c_l I_{l+3/2}(x) + d_l K_{l+3/2}(x) \tag{8.4.25}$$

where the coefficients a_l, b_l, c_l and d_l are found imposing the boundary conditions:

$$\text{first set} \;\rightarrow\; B_l^r(x_1) = r_1^{1/2}\left(\frac{l}{2l+1}\right)^{1/2} \quad B_l^r(x_2) = 0 \tag{8.4.26}$$

$$B_l^\theta(x_1) = r_1^{1/2}\left(\frac{l+1}{2l+1}\right)^{1/2} \quad B_l^\theta(x_2) = 0$$

$$\text{second set} \;\rightarrow\; B_l^r(x_1) = 0 \quad B_l^r(x_2) = r_2^{1/2}\left(\frac{l}{2l+1}\right)^{1/2} \tag{8.4.27}$$

$$B_l^\theta(x_1) = 0 \quad B_l^\theta(x_2) = r_2^{1/2}\left(\frac{l+1}{2l+1}\right)^{1/2}$$

8.4.3 Solution of the problem for the pressure

The problem for the pressure coefficients of expansion 8.4.1 is solved by means of a tau spectral method. The second order differential operator D_l is firstly transformed into a constant coefficients operator using the following radial variable transformation:

$$x = x(r) = \frac{\ln\left(\frac{r^2}{r_1 r_2}\right)}{\ln\left(\frac{r_2}{r_1}\right)} \quad \rightarrow \quad r = r(x) = r_1 e^{(x+1)/\alpha} \text{ where } \alpha = \frac{2}{\ln\left(\frac{r_2}{r_1}\right)} \tag{8.4.28}$$

which leads to the following relationships:

$$r\frac{d}{dr} \rightarrow x\frac{d}{dx} \quad r^2\frac{d^2}{dr^2} \rightarrow \alpha^2\frac{d^2}{dx^2} - \alpha\frac{d}{dx} \quad D_l^{(x)} = r^2 D_l \rightarrow \alpha^2\frac{d^2}{dx^2} - \alpha\frac{d}{dx} - l(l+1) \tag{8.4.29}$$

The modal equations 8.4.3 for the pressure with the above transformation reduce to the following:

$$\left[\alpha^2\frac{d^2}{dx^2} - \alpha\frac{d}{dx} - l(l+1)\right] p_{lm}(x) = r(x)\left[\left(\alpha\frac{d}{dx} + 2\right) f_{lm}^r(x) - S_l f_{lm}^\theta(x)\right] \tag{8.4.30}$$

$$\begin{aligned} &-\int_{-1}^{+1} r(x)^2\left[\alpha\frac{dp_{lm}(x)}{dx} b_l^r(x) + S_l p_{lm}(x) b_l^\theta(x)\right] dx = \qquad (8.4.31)\\ &-\int_{-1}^{+1} r(x)^3\left[f_{lm}^r(x) b_l^r(x) + f_{lm}^\theta(x) b_l^\theta(x)\right] dx + \\ &+\alpha\left\{v_{s\,lm}^r r(x)\left(\alpha\frac{d}{dx} + 2\right) b_l^r(x) - v_{s\,lm}^\theta r(x)\left[S_l b_l^r(x) - \alpha\frac{db_l^\theta(x)}{dx}\right]\right\}\Bigg|_{-1}^{+1} \end{aligned}$$

Notice that the radial streching 8.4.28 allows to transform the second-order differential operator D_l into a constant coefficients operator but it implies also the introduction of variables coefficients in the right-hand side of the equation 8.4.30. In order to solve the equation 8.4.30 with the boundary condition 8.4.31, a tau projection method is used. Firstly all the functions appearing in the 8.4.28 are expanded in series of Chebyshev polynomials $T_n(x) = cos(n\, arccos(x))$ (see Appendix C for a wide exposure of Chebyshev

polynomials properties):

$$f_{lm}(x) = \sum_{n=0}^{N} f_{lmn} T_n(x) \tag{8.4.32}$$

$$p_{lm}(x) = \sum_{n=0}^{N} p_{lmn} T_n(x) \tag{8.4.33}$$

$$\tag{8.4.34}$$

After introducing the expansion 9.2.6 in the 8.4.30, the final form of the equations for the pressure Chebyshev coefficients can be obtained substituting the Chebyshev polynomials derivatives with their expansion in term of Chebyshev polinomials. Multiplying the equations for the term $\frac{T_q(x)}{\sqrt{1-x^2}}$ and integrating between $[-1;1]$ we can use the orthogonality properties of such polynomials and the equations for the pressure coefficients become:

$$\alpha^2 \sum_{k=0}^{N} p_{lmk} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} 2k \sum_{j=1}^{\frac{k-1}{2}} \sum_{i=1}^{j} 4j T_{2i-1}(x)\, dx & k \text{ odd} \\ 2k \sum_{j=0}^{\frac{k-2}{2}} (2j+1) \left[1 + 2\sum_{i=1}^{j} T_{2i}(x)\right] dx & k \text{ even} \end{cases}}_{II} + \tag{8.4.35}$$

$$+ \alpha \sum_{k=0}^{N} p_{lmk} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} k \left[1 + 2 \sum_{j=1}^{\frac{k-1}{2}} T_{2j}(x)\right] dx & k \text{ odd} \\ 2k \sum_{j=1}^{\frac{k}{2}} T_{2j-1}(x) dx & k \text{ even} \end{cases}}_{I} +$$

$$- \left[l(l+1)\right] \sum_{k=0}^{N} p_{lmk} \underbrace{\int_{-1}^{1} \frac{T_q(x) T_k(x)}{\sqrt{1-x^2}} dx}_{O} =$$

$$\sum_{j=0}^{N} \left[\alpha f^r_{lmj} \underbrace{\int_{-1}^{1} \frac{T_q(x) T'_j(x) r(x)}{\sqrt{1-x^2}} dx}_{D_{qj}} + \left(2 f^r_{lmj} - S_l f^{\theta}_{lmj}\right) \underbrace{\int_{-1}^{1} \frac{T_q(x) T_j(x) r(x)}{\sqrt{1-x^2}} dx}_{C_{qj}} \right]$$

with the boundary integral condition:

$$-\operatorname{Re}\alpha\sum_{k=0}^{N}p_{lmk}\underbrace{\int_{-1}^{1}T_k'(x)b_l^r(x)r^2(x)dx}_{H_k'}-\operatorname{Re}S_l\sum_{k=0}^{N}p_{lmk}\underbrace{\int_{-1}^{1}T_k(x)b_l^\theta(x)r^2(x)dx}_{H_k}= \tag{8.4.36}$$

$$=-\sum_{k=0}^{N}f_{lmk}^r\underbrace{\int_{-1}^{1}T_k(x)b_{lm}^r(x)r^3(x)dx}_{K_k}-\sum_{k=0}^{N}f_{lmk}^\theta\underbrace{\int_{-1}^{1}T_k(x)b_{lm}^\theta(x)r^3(x)dx}_{K_k}+ \tag{8.4.37}$$

$$+\alpha\left\{v_{Slm}^r r(x)\left(\alpha\frac{d}{dx}+2\right)b_{lm}^r-v_{Slm}^\theta r(x)\left(S_l b_{lm}^r-\alpha\frac{d}{dx}b_{lm}^\theta\right)\right\}\Big|_{-1}^{1} \tag{8.4.38}$$

where:

$$\begin{aligned}
II &= \begin{cases} k\ odd \begin{cases} q\ odd \rightarrow \pi k\delta_{(2i-1),q}\sum\limits_{l=\frac{q+1}{2}}^{\frac{k-1}{2}}4l \\ q\,even \rightarrow 0 \end{cases} \\ k\ even \begin{cases} q\ odd \rightarrow 0 \\ q\,even \rightarrow 2\pi k\delta_{2i,q}\sum\limits_{l=\frac{q}{2}}^{\frac{k}{2}}(2l+1) \end{cases} \end{cases} \\
I &= \begin{cases} k\ odd \begin{cases} q\ odd \rightarrow 0 \\ q\,even \rightarrow k\pi\delta_{2j,q} \end{cases} \\ k\ even \begin{cases} q\ odd \rightarrow k\pi\delta_{(2j-1),q} \\ q\,even \rightarrow 0 \end{cases} \end{cases} \\
O &= \begin{cases} \frac{\pi}{2}\delta_{q,k} \\ \pi\delta_{q,0} \end{cases}
\end{aligned} \tag{8.4.39}$$

The integrals D_{qj}, C_{qj} can be evaluated by means of a Gauss-Chebyshev quadrature formula:

$$D_{qj}=\int_{-1}^{1}\frac{r(x)T_q(x)T_j'(x)}{\sqrt{1-x^2}}dx=\int_{-1}^{1}\frac{u(x)}{\sqrt{1-x^2}}dx=\frac{\pi}{N+1}\sum_{k=1}^{N+1}u(x_k) \tag{8.4.40}$$

$$C_{qj}=\int_{-1}^{1}\frac{r(x)T_q(x)T_j(x)}{\sqrt{1-x^2}}dx=\int_{-1}^{1}\frac{v(x)}{\sqrt{1-x^2}}dx=\frac{\pi}{N+1}\sum_{k=1}^{N+1}v(x_k) \tag{8.4.41}$$

and the points x_k are the Chebyshev nodes.

The integrals H'_k, H_k and K_k are instead evaluated by means of a Gauss-Legendre quadrature formula:

$$H'_k = \int_{-1}^{1} T'_k(x) b_l^r(x) r^2(x) dx = \int_{-1}^{1} u(x) dx = \sum_{k=1}^{N+1} w_k u(x_k) \tag{8.4.42}$$

$$H_k = \int_{-1}^{1} T_k(x) b_l^\theta(x) r^2(x) dx = \int_{-1}^{1} v(x) dx = \sum_{k=1}^{N+1} w_k v(x_k) \tag{8.4.43}$$

$$K_k = \int_{-1}^{1} T_k(x) b_l^r(x) r^3(x) dx = \int_{-1}^{1} \omega(x) dx = \sum_{k=1}^{N+1} w_k \omega(x_k) \tag{8.4.44}$$

where w_k and x_k are the Legendre weights and nodes respectively.

The resulting linear system of equations for the pressure coefficients p_{lmk} with the addition of the associated integral conditions (one for each boundary and each lm mode) in the first two rows, has an upper triangular coefficients' matrix, with the coefficients distributed according to a regular pattern. This allows to transform the matrix according to a quasi-pentadiagonal profile by means of suitable linear combinations. Finally, the problem can be solved by factorizing the matrix with a UL decomposition starting from the bottom and proceeding to the top of the matrix.

8.4.4 Solution of the problem for the velocity

The velocity problem can be solved independently, after having solved the pressure problem, by means of a tau projection method, similar to the above described technique used for the pressure problem. The application of the 8.4.28 radial stretching to the velocity equations leads to the following modal equations:

$$\left(-\mathbf{D}_l^{(x)} + \gamma r^2(x)\mathbf{I}\right)\mathbf{u}_{lm}(x) = -r(x)\nabla_l^{(x)} p_{lm}(x) + r^2(x)\mathbf{f}_{lm}(x) \tag{8.4.45}$$

$$\mathbf{u}_{lm}(-1) = v_{Slm}^1 \qquad\qquad \mathbf{u}_{lm}(1) = v_{Slm}^2 \tag{8.4.46}$$

where:

$$\mathbf{D}_l^{(x)} = \begin{pmatrix} D_l^{(x)} - 2 & 2S_l & 0 \\ 2S_l & D_l^{(x)} & 0 \\ 0 & 0 & D_l^{(x)} \end{pmatrix} \quad and \quad \nabla_l^{(x)} = \begin{pmatrix} \alpha\frac{d}{dx} \\ S_l \\ 0 \end{pmatrix} \tag{8.4.47}$$

Notice that the problem for the components u_{lm}^r and u_{lm}^θ is coupled. The original formulation can be split in the r and θ uncoupled components using the similarity transformation:

$$\begin{pmatrix} U_{lm}^r \\ U_{lm}^\theta \end{pmatrix} = S_l \begin{pmatrix} u_{lm}^r \\ u_{lm}^\theta \end{pmatrix} \quad \text{where } S_l = \begin{pmatrix} \left(\frac{l}{2l+1}\right)^{1/2} & \left(\frac{l+1}{2l+1}\right)^{1/2} \\ \left(\frac{l+1}{2l+1}\right)^{1/2} & -\left(\frac{l}{2l+1}\right)^{1/2} \end{pmatrix} \tag{8.4.48}$$

Applying the transformation 8.4.48 we obtain:

$$\begin{aligned} \left(-D_l^{(x)} + 2 + \gamma r^2(x)\right)\left[\left(\frac{l}{2l+1}\right)^{1/2} U_{lm}^r(x) + \left(\frac{l+1}{2l+1}\right)^{1/2} U_{lm}^\theta(x)\right] \\ -2S_l\left[\left(\frac{l+1}{2l+1}\right)^{1/2} U_{lm}^r(x) - \left(\frac{l}{2l+1}\right)^{1/2} U_{lm}^\theta(x)\right] = \\ -r(x)\alpha\frac{dp_{lm}(x)}{dx} + r^2(x) f_{lm}^r(x) \end{aligned} \tag{8.4.49}$$

$$\begin{aligned} -2S_l\left[\left(\frac{l}{2l+1}\right)^{1/2} U_{lm}^r(x) + \left(\frac{l+1}{2l+1}\right)^{1/2} U_{lm}^\theta(x)\right] \\ +\left(-D_l^{(x)} + \gamma r^2(x)\right)\left[\left(\frac{l+1}{2l+1}\right)^{1/2} U_{lm}^r(x) - \left(\frac{l}{2l+1}\right)^{1/2} U_{lm}^\theta(x)\right] = \\ -r(x)S_l p_{lm}(x) + r^2(x) f_{lm}^\theta(x) \end{aligned} \tag{8.4.50}$$

Adding or subtracting the 8.4.49 to the 8.4.50 after suitable scaling, the problem can be finally reformulated in the split form :

$$\left(D_l^{(x)} - \gamma r^2(x) - 2l\right)(2l+1)^{1/2} U_{lm}^r(x) = \tag{8.4.51}$$
$$r(x)\left[\alpha l^{1/2}\frac{dp_{lm}(x)}{dx} + S_l\,(l+1)^{1/2}\,p_{lm}(x)\right] - r^2(x)\left[l^{1/2} f_{lm}^r(x) - (l+1)^{1/2}\,f_{lm}^\theta(x)\right]$$

$$\left(D_l^{(x)} - \gamma r^2(x) - 2\,(l+1)\right)(2l+1)^{1/2}\,U_{lm}^\theta(x) = \tag{8.4.52}$$
$$r(x)\left[\alpha\,(l+1)^{1/2}\,\frac{dp_{lm}(x)}{dx} - S_l l^{1/2} p_{lm}(x)\right] - r^2(x)\left[(l+1)^{1/2}\,f_{lm}^r(x) - l^{1/2} f_{lm}^\theta(x)\right]$$

The velocity equation in the φ direction is originally already uncoupled:

$$\left(-D_l^{(x)} + \gamma r^2(x)\right) u_{lm}^\varphi(x) = r^2(x) f_{lm}^\varphi(x) \tag{8.4.53}$$

In this manner, the three-dimensional vector problem is split into a sequence of three uncoupled scalar one-dimensional problem for the velocity components U_{lm}^r, U_{lm}^θ and u_{lm}^φ.

Now, introducing a truncated expansion of the functions appearing in the 8.4.51, 8.4.52 and 8.4.53 in terms of Chebyshev polynomials, such as:

$$f_{lm}(x) = \sum_{n=0}^{N} f_{lmn} T_n(x) \tag{8.4.54}$$

$$p_{lm}(x) = \sum_{n=0}^{N} p_{lmn} T_n(x) \tag{8.4.55}$$

$$\mathbf{u}_{lm}(x) = \sum_{n=0}^{N} \mathbf{u}_{lmn} T_n(x) \tag{8.4.56}$$

the final form of the equations for the velocity Chebyshev coefficients can be obtained substituting the Chebyshev polynomials derivatives with their expansion in term of Chebyshev polinomials. Multiplying the equations for the term $\frac{T_q(x)}{\sqrt{1-x^2}}$ and integrating between $[-1;1]$ we can use the orthogonality properties of such polynomials. In the following the equations for the three components of the velocity are reported with an indication of their structure and matrix of coefficients aspect.

Equation for U^r_{lm}

$$(2l+1)^{1/2}\alpha^2\sum_{k=0}^{N}U^r_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)}{1-x^2}\begin{cases}2k\sum\limits_{j=1}^{\frac{k-1}{2}}\sum\limits_{i=1}^{j}4jT_{2i-1}(x)\,dx & k\ odd\\ 2k\sum\limits_{j=0}^{\frac{k-2}{2}}(2j+1)\left[1+2\sum\limits_{i=1}^{j}T_{2i}(x)\right]dx & k\ even\end{cases}}_{II}+$$

(8.4.57)

$$+(2l+1)^{1/2}\alpha\sum_{k=0}^{N}U^r_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)}{1-x^2}\begin{cases}k\left[1+2\sum\limits_{j=1}^{\frac{k-1}{2}}T_{2j}(x)\right]dx & k\ odd\\ 2k\sum\limits_{j=1}^{\frac{k}{2}}T_{2j-1}(x)dx & k\ even\end{cases}}_{I}+$$

$$-(2l+1)^{1/2}\gamma\sum_{k=0}^{N}U^r_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_k(x)r^2(x)}{\sqrt{1-x^2}}dx}_{E_{qk}}-(2l+1)^{1/2}\left[l(l-1)\right]\sum_{k=0}^{N}U^r_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_k(x)}{\sqrt{1-x^2}}dx}_{O}=$$

$$l^{1/2}\alpha\sum_{j=0}^{N}p_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T'_j(x)r(x)}{\sqrt{1-x^2}}dx}_{D_{qj}}+S_l(l+1)^{1/2}\sum_{j=0}^{N}p_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_j(x)r(x)}{\sqrt{1-x^2}}dx}_{C_{qj}}$$

$$l^{1/2}\sum_{j=0}^{N}f^r_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_j(x)r^2(x)}{\sqrt{1-x^2}}dx}_{E_{qj}}+(l+1)^{1/2}\sum_{j=0}^{N}f^\theta_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_j(x)r^2(x)}{\sqrt{1-x^2}}dx}_{E_{qj}}$$

where:

$$II=\begin{cases}k\ odd\begin{cases}q\ odd\ \rightarrow \pi k\delta_{(2i-1),q}\sum\limits_{l=\frac{q+1}{2}}^{\frac{k-1}{2}}4l\\ q\,even\ \rightarrow\ 0\end{cases}\\ k\ even\begin{cases}q\ odd\ \rightarrow 0\\ q\,even\ \rightarrow\ 2\pi k\delta_{2i,q}\sum\limits_{l=\frac{q}{2}}^{\frac{k}{2}}(2l+1)\end{cases}\end{cases}\qquad(8.4.58)$$

$$
\begin{aligned}
I \;\; &= \begin{cases} k\ odd \begin{cases} q\ odd \;\rightarrow 0 \\ q\, even \;\rightarrow\; k\pi\delta_{2j,q} \end{cases} \\ k\ even \begin{cases} q\ odd \;\rightarrow k\pi\delta_{(2j-1),q} \\ q\, even \;\rightarrow\; 0 \end{cases} \end{cases} \\
O &= \begin{cases} \frac{\pi}{2}\delta_{q,k} \\ \pi\delta_{q,0} \end{cases}
\end{aligned}
\tag{8.4.59}
$$

The integrals C_{qj}, D_{qj} and E_{qj} can be evaluated by means of a Gauss-Chebyshev quadrature formula:

$$C_{qj} = \int_{-1}^{1} \frac{r(x)T_q(x)T_j(x)}{\sqrt{1-x^2}}dx = \int_{-1}^{1} \frac{u(x)}{\sqrt{1-x^2}}dx = \frac{\pi}{N+1}\sum_{j=1}^{N+1} u(x_j) \tag{8.4.60}$$

$$D_{qj} = \int_{-1}^{1} \frac{r(x)T_q(x)T_j'(x)}{\sqrt{1-x^2}}dx = \int_{-1}^{1} \frac{v(x)}{\sqrt{1-x^2}}dx = \frac{\pi}{N+1}\sum_{j=1}^{N+1} v(x_j) \tag{8.4.61}$$

$$E_{qj} = \int_{-1}^{1} \frac{r^2(x)T_q(x)T_j(x)}{\sqrt{1-x^2}}dx = \int_{-1}^{1} \frac{\omega(x)}{\sqrt{1-x^2}}dx = \frac{\pi}{N+1}\sum_{j=1}^{N+1} \omega(x_j) \tag{8.4.62}$$

$$\tag{8.4.63}$$

and the points x_k are the Chebyshev nodes.

Equation for U^{θ}_{lm}

$$(2l+1)^{1/2}\alpha^2\sum_{k=0}^{N}U^{\theta}_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)}{1-x^2}\begin{cases}2k\sum\limits_{j=1}^{\frac{k-1}{2}}\sum\limits_{i=1}^{j}4jT_{2i-1}(x)\,dx & k\ odd\\ 2k\sum\limits_{j=0}^{\frac{k-2}{2}}(2j+1)\left[1+2\sum\limits_{i=1}^{j}T_{2i}(x)\right]dx & k\ even\end{cases}}_{II}+ \tag{8.4.64}$$

$$+(2l+1)^{1/2}\alpha\sum_{k=0}^{N}U^{\theta}_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)}{1-x^2}\begin{cases}k\left[1+2\sum\limits_{j=1}^{\frac{k-1}{2}}T_{2j}(x)\right]dx & k\ odd\\ 2k\sum\limits_{j=1}^{\frac{k}{2}}T_{2j-1}(x)dx & k\ even\end{cases}}_{I}+$$

$$-(2l+1)^{1/2}\gamma\sum_{k=0}^{N}U^{\theta}_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_k(x)r^2(x)}{\sqrt{1-x^2}}dx}_{E_{qk}}+$$

$$-(2l+1)^{1/2}\left[(l+1)(l+2)\right]\sum_{k=0}^{N}U^{\theta}_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_k(x)}{\sqrt{1-x^2}}dx}_{O}=$$

$$(l+1)^{1/2}\alpha\sum_{j=0}^{N}p_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T'_j(x)r(x)}{\sqrt{1-x^2}}dx}_{D_{qj}}-S_l l^{1/2}\sum_{j=0}^{N}p_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_j(x)r(x)}{\sqrt{1-x^2}}dx}_{C_{qj}}$$

$$(l+1)^{1/2}\sum_{j=0}^{N}f^{r}_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_j(x)r^2(x)}{\sqrt{1-x^2}}dx}_{E_{qj}}+l^{1/2}\sum_{j=0}^{N}f^{\theta}_{lmk}\underbrace{\int_{-1}^{1}\frac{T_q(x)T_j(x)r^2(x)}{\sqrt{1-x^2}}dx}_{E_{qj}}$$

In which the underscript terms are the same in 8.4.57.

Equation for u_{lm}^{φ}

$$\alpha^2 \sum_{k=0}^{N} u_{lmk}^{\varphi} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} 2k \sum_{j=1}^{\frac{k-1}{2}} \sum_{i=1}^{j} 4jT_{2i-1}(x)\, dx & k\ odd \\ 2k \sum_{j=0}^{\frac{k-2}{2}} (2j+1) \left[1 + 2\sum_{i=1}^{j} T_{2i}(x)\right] dx & k\ even \end{cases}}_{II} +$$

(8.4.65)

$$+ \alpha \sum_{k=0}^{N} u_{lmk}^{\varphi} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} k \left[1 + 2 \sum_{j=1}^{\frac{k-1}{2}} T_{2j}(x)\right] dx & k\ odd \\ 2k \sum_{j=1}^{\frac{k}{2}} T_{2j-1}(x) dx & k\ even \end{cases}}_{I} +$$

$$- \gamma \sum_{k=0}^{N} u_{lmk}^{\varphi} \underbrace{\int_{-1}^{1} \frac{T_q(x) T_k(x) r^2(x)}{\sqrt{1-x^2}} dx}_{E_{qk}} - [l(l+1)] \sum_{k=0}^{N} u_{lmk}^{\varphi} \underbrace{\int_{-1}^{1} \frac{T_q(x) T_k(x)}{\sqrt{1-x^2}} dx}_{O} =$$

$$- \sum_{j=0}^{N} f_{lmk}^{\varphi} \underbrace{\int_{-1}^{1} \frac{T_q(x) T_j(x) r^2(x)}{\sqrt{1-x^2}} dx}_{E_{qj}}$$

In which the underscript terms are the same in 8.4.57 and 8.4.64.

The matrix of coefficients for the velocity problem is full for each component, for $\gamma > 0$ for the presence of E_{qj} integral. Except for this term the pattern of the matrix of coefficients is similar to the pressure problem one. Each components of the velocity problem associated with its boundary conditions, suitably expanded in Chebyshev series is calculated solving the system of equations by means of a standard LU decomposition.

8.4.5 First mode equations

For the first mode with $l = 0$ and $m = 0$ the modal equations reduce to the following:

$$D_0 p_{00} = \left(\frac{d}{dr} + \frac{2}{r}\right) f^r_{00} \tag{8.4.66}$$

$$-\int_{r1}^{r2} \left(\frac{dp_{00}}{dr} b^r_{l=0}\right) r^2 dr = -\int_{r1}^{r2} (f^r_{00} b^r_{l=0}) r^2 dr \left[v^r_{S00} r^2 \left(\frac{d}{dr} + \frac{2}{r}\right) b^r_{l=0}\right]\Bigg|_{r1}^{r2}$$

$$(-D_l + \gamma)\, u_{00} = -\frac{dp_{00}}{dr} + f^r_{00} \tag{8.4.67}$$

$$\mathbf{u}^r_{00}(r_1) = \mathbf{v}^r_{S00}(r_1) \quad \mathbf{u}^r_{00}(r_2) = \mathbf{v}^r_{S00}(r_2) \tag{8.4.68}$$

where only the radial component of the vector mode is non-zero. For this mode the metaharmonic problem assume the following scalar form:

$$(-D_l + \gamma)\, b^r_0 = 0 \tag{8.4.69}$$

$$r_1^2 b^r_0(r_1) = r_2^2 b^r_0(r_2) \neq 0 \tag{8.4.70}$$

which has only one linear independent solution.

To find an analytic solution of the **metaharmonic problem** 8.4.69 the following variable transformation is required:

$$r \to x = \gamma^{1/2} r \quad b^r_0(r) \to B0^r(x) = r^{1/2} b^r_0(r) = \gamma^{-1/4} x^{1/2} b^r_0(r)(\gamma^{-1/2} x) \tag{8.4.71}$$

which permit to reformulate the problem in the new variable such as:

$$\mathbf{B}_{3/2} B^r_0(x) = 0 \tag{8.4.72}$$

where $\mathbf{B}_{3/2}$ is the modified Bessel operator of order 3/2, whose solution is:

$$B^r_0(x) = a_0 I_{3/2}(x) + b_0 K_{3/2}(x) \tag{8.4.73}$$

where a_0 and b_0 must satisfy the boundary conditions:

$$x_1^{3/2} B^r_0(x_1) = x_2^{3/2} B^r_0(x_2) \neq 0 \tag{8.4.74}$$

The **problem** 8.4.66 **for the pressure** defines $p_{00}(r)$ up to an arbitrary additive constant. To determine the unknown $p_{00}(r)$ uniquely, the value at a single point, say r_0, $r_1 \leq r_0 \leq r_2$, must be fixed by imposing a supplementary condition of the type $p_{00}(r_0) = C$, with C a constant. Writing equation 8.4.66 in the factored form:

$$\left(\frac{d}{dr} + \frac{2}{r}\right) \frac{dp_{00}}{dr} = \left(\frac{d}{dr} + \frac{2}{r}\right) f^r_{00} \tag{8.4.75}$$

Assuming $L = \left(\frac{d}{dr} + \frac{2}{r}\right)$ a further quadrature of 8.4.75 leads to the following result:

$$\begin{aligned} &L\frac{dp_{00}}{dr} = Lf_{00}^r;\;\; L\left(\frac{dp_{00}}{dr} - f_{00}^r\right) = 0; \\ &L\phi = 0 \rightarrow \frac{d\phi}{dr} + \frac{2\phi}{r} = 0 \rightarrow \frac{d\phi}{\phi} = -\frac{2dr}{r} \rightarrow \ln\left(\frac{\phi}{\phi_0}\right) = -2\ln\left(\frac{r}{r_0}\right) \\ &\rightarrow \phi = \phi_0\left(\frac{r_0}{r}\right)^2 \rightarrow \left(\frac{dp_{00}}{dr} - f_{00}^r\right) = \frac{B}{r^2} \end{aligned} \tag{8.4.76}$$

where B is the constant of integration, whose value is determined by the integral condition for the pressure, which can be written as follow:

$$-\int_{r1}^{r2} Bb_{00}^r dr = \left[v_{S00}^r r^2 \left(\frac{d}{dr} + \frac{2}{r}\right) b_{l=0}^r\right]\Bigg|_{r1}^{r2} \tag{8.4.77}$$

Because of the existence of the constraint 8.1.9, $b_{00}^r(r_1)r_1^2 = b_{00}^r(r_2)r_2^2 = K$, where K is a problem datum (for example $K = 0$), so that:

$$-\int_{r1}^{r2} Bb_{00}^r dr = K\left[\left(\frac{d}{dr} + \frac{2}{r}\right) b_{l=0}^r\right]\Bigg|_{r1}^{r2} \tag{8.4.78}$$

using the fundamental theorem of calculus the 8.4.78 becomes:

$$\begin{aligned} -B\int_{r1}^{r2} b_{00}^r dr &= K\int_{r1}^{r2} \frac{d}{dr}\left(\frac{d}{dr} + \frac{2}{r}\right) b_{l=0}^r \\ -B\int_{r1}^{r2} b_{00}^r dr &= K\int_{r1}^{r2} \gamma b_{l=0}^r = K\gamma\int_{r1}^{r2} b_{l=0}^r \end{aligned} \tag{8.4.79}$$

Hence $B = K\gamma$ and the solution of 8.4.76 is obtained by integrating:

$$\left(\frac{dp_{00}}{dr}\right) = f_{00}^r - \frac{\gamma K}{r^2} \tag{8.4.80}$$

using the supplementary condition: $p_{00}(r_0) = C$. The **problem for** the first mode of **the velocity** $\mathbf{u}_{00}(r)$ does not differ from the problem for the generic mode lm and it is treated in the following.

Chapter 9

The numerical code

A schematic representation in form of block diagram of the main sections of the numerical code is shown in the figure below:

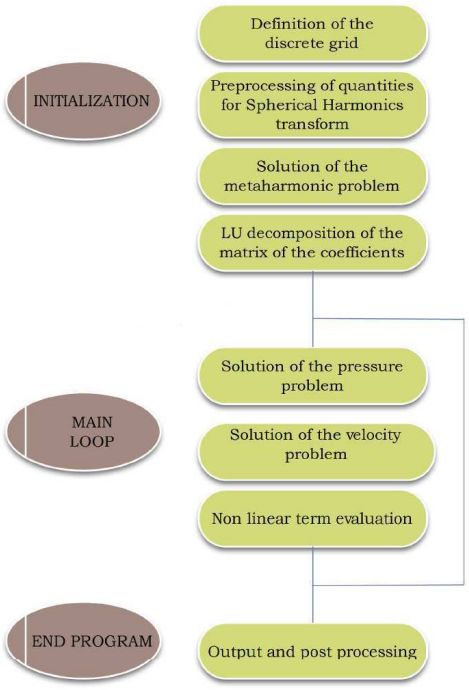

Figure 9.1: Block Diagram of the main sections of the numerical code

In the following sections each part of the code is described in detail. All the constant, variables and operators are defined in double precision. The basis of the code is constituted by suitable functions and subroutines that compute Chebyshev, Legendre polynomials and Spherical harmonics functions, transform and make the inverse transform of scalar and vector functions on these basis, solve linear system of equations. Many other routines have been used to make basic operations such as the factorial of an integer. These routines are trivial and will be not described in the following for simplicity. When possible NR routines have been conveniently adapted and used inside the code.

9.1 Initialization

9.1.1 Definition of the discrete grid

The numerical code requires the domain to be discretized in a regular grid. The flow field is represented in a spherical system of reference and have been discretized in r, θ and φ directions independently.

- In **r** direction (in which the r coordinate varies between r_1 and r_2) the grid has nr points, included the extremal. In this direction the independent variable r is firstly resized to the range $(-1, 1)$ by means of the linear transformation:

$$y_r = y_r(r) = \frac{\ln\left(\frac{r^2}{r_1 r_2}\right)}{\ln\left(\frac{r_2}{r_1}\right)} \quad \rightarrow \; r = r(y_r) = r_1 e^{(y_r+1)/\alpha} \text{ with } \alpha = \frac{2}{\ln\left(\frac{r_2}{r_1}\right)} \tag{9.1.1}$$

  The transformed radial variables y_r assumed the values -1 and $+1$ at the extremal r_1 and r_2, and in the internal points coincides with the $nc = (nr - 2)$ Chebyshev nodes:

$$y_{ri} = \cos\left(\frac{\pi}{2}\,\frac{2i-1}{nc}\right), \quad i = 1, \ldots, nc. \tag{9.1.2}$$

- θ coordinate is the polar angle and is comprised in the in range $(0, \pi)$, where the values 0 and π correspond to the poles of the sphere. In this direction the discrete grid has $n\theta$ points corresponding to the Legendre nodes internal to the range $(-1, 1)$, of the transformed variables y_t with:

$$y_t = y_t(\theta) = cos(\theta) \quad \rightarrow \; \theta = \theta(y_t) = arccos(-y_t) \tag{9.1.3}$$

The $n\theta$ Legendre nodes are calculated by means of the NR subroutine gauleg.f90. The extremal points corresponding to the values $y_t = -1$ and $y_t = 1$ are singular points and thus are excluded from the grid.

- φ coordinate is the azimutal angle and varies in the in range $(0, 2\pi)$. In this direction the discrete grid has $n\varphi = (2n\theta - 1)$ evenly spaced points with spacing $d\varphi = \frac{2\pi}{n\varphi}$ in the discrete variable y_p.

9.1.2 Preprocessing

In preprocessing section the code computes several quantities that are required for the transform and inverse transform operations. Such operations are necessary to evaluate the non linear term inside the main loop and to reconstruct the flow fields starting from the p_{lm} and $\mathbf{u}_{lm}$ modes of pressure and velocity expansions in scalar and vector spherical harmonics. Moreover, in this part the code computes some integrals that are required in pressure and velocity problem solver routine to evaluate the right hand side of the governing equations.

Spherical harmonics scalar Y_{lm} and vector $\mathbf{P}_{lm}$, $\mathbf{B}_{lm}$, $\mathbf{C}_{lm}$ basis functions depends on θ and φ coordinates and from l and m indices. In order to evaluate such functions in the y_t and y_p discrete nodes, we isolate the dependence from θ and φ. The dependence from φ in spherical harmonics basis is contained in complex exponential functions while the θ dependence is contained inside the Associated Legendre functions and their first derivatives.

Inside the code we do in sequence the following operations::

- Calculation of the Associated Legendre Functions (ALF) $P_l^m(y_t)$ for each combination of l - m indices using the recurrence relation for $m \geq 0$:

$$(l-m)P_{l+1}^m(y_t) = (2l-1)y_t P_l^{m+1}(y_t) - (l+m-1)P_l^m(y_t) \tag{9.1.4}$$

reminding that:

$$P_0^0(y_t) = 1 \tag{9.1.5}$$
$$P_1^0(y_t) = x \tag{9.1.6}$$

The Associated Legendre Functions with $m < 0$ have been calculated using the property:

$$P_l^{-m}(y_t) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(y_t) \tag{9.1.7}$$

- Calculation of the First Derivative of the Associated Legendre Functions (DALF) $P_l^{m'}(y_t)$ for each combination of l - m indices using the recurrence relation :

$$\sqrt{1-y_t^2}P_l^{m'}(y_t) = (l-m+1)P_{l+1}^m(y_t) - (l+1)y_t P_l^m(y_t) \tag{9.1.8}$$

reminding that:

$$P_0^{0'}(y_t) = 0 \tag{9.1.9}$$

$$P_1^{0'}(y_t) = 1 \tag{9.1.10}$$

- Calculation of the Scalar Spherical Harmonics $Y_{lm}(y_t, y_p)$ (defined as complex(DPC)) using its definition, reminding that $m \leq l$:

$$Y_{lm}(y_t, y_p) = N_{lm}\, P_l^m\,(y_t)\, e^{im\varphi} \tag{9.1.11}$$

- Calculation of the Vector Spherical Harmonics:
  - $\mathbf{P}_{lm}(y_t, y_p)$ (defined as complex(DPC)) using its definition, reminding that $m \leq l$:

$$\mathbf{P}_{lm}(y_t, y_p) = N_{lm}\, P_l^m\,(y_t)\, e^{im\varphi}\hat{r}(y_t, y_p) \tag{9.1.12}$$

  - $\mathbf{B}_{lm}(y_t, y_p)$ (defined as complex(DPC)) using its definition, reminding that $m \leq l$:

$$\begin{aligned}\mathbf{B}_{lm}(y_t, y_p) = &\frac{1}{S_l} N_{lm}\, P_l^{m'}(y_t) e^{im\varphi}\hat{\theta}(y_t, y_p) \\ &+ \frac{1}{S_l sen(\theta)} N_{lm}\, P_l^m(y_t)(-im) e^{im\varphi}\hat{\varphi}(y_t, y_p)\end{aligned} \tag{9.1.13}$$

  - $\mathbf{C}_{lm}(y_t, y_p)$ (defined as complex(DPC)) using its definition, reminding that $m \leq l$:

$$\begin{aligned}\mathbf{C}_{lm}(y_t, y_p) = &\frac{1}{S_l sen(\theta)} N_{lm}\, P_l^m(y_t)(-im) e^{im\varphi}\hat{\theta}(y_t, y_p) \\ &+ \frac{1}{S_l} N_{lm}\, P_l^{m'}(y_t) e^{im\varphi}\hat{\varphi}(y_t, y_p)\end{aligned} \tag{9.1.14}$$

Y_{lm} ,$\mathbf{P}_{lm}$, $\mathbf{B}_{lm}$, $\mathbf{C}_{lm}$ values for each y_t and y_p nodes and l and m indices are stored in matrix with four indices and are ready to be passed to the routines that do the transform and the inverse transform of scalar and vector quantities inside the main loop.

This section of the code ended with the evaluation of the integrals C_{qj}, D_{qj}, E_{qj}, H'_k, H_k and K_k appearing in the governing equations for the pressure and the velocity and inside the integrals condition for the pressure problem. Such integrals are evaluated by means of Gauss-Chebyshev quadrature formulas (C_{qk}, D_{qk} and E_{qk}) and Gauss-Legendre quadrature formulas (H'_k, H_k and K_k), according to what has been described in 8.4.60,8.4.42, 8.4.43 and 8.4.44. All these integrals contain a direct dependence from the radial variable, and have to be evaluated for each index q, k and for each radial node. In particular C_{qk}, D_{qk} and E_{qk} integrals involve the variable y_r and the Chebyshev polynomials or their first derivatives evaluated in Chebyshev nodes. H'_k, H_k and K_k integrals implies the integrand functions to be evaluated in Legendre nodes. This suggests the introduction of another discrete variable y_l varying into the range $(-1, 1)$ in the radial direction and assuming Legendre nodes values. In the extremal nodes -1 and 1 the variable y_l assumes the values:

$$\begin{aligned} y_l(-1) &= -1 \\ y_l(1) &= 1 \end{aligned} \tag{9.1.15}$$

The internal *nleg* nodes are evaluated using NR gauleg.f90 subroutine, that supplies also the *nleg* Legendre weights w_i required for the H'_k, H_k and K_k integrals evaluation.

9.1.3 Solution of the metaharmonic problem

The metaharmonic problem solution is necessary to built the integral condition for the pressure problem which allows to solve the problem for p_{lm} modes before and independently from the velocity problem. Metaharmonic functions b^r and b^θ (b^φ is identically equal to zero) depend only on the l index of Spherical Harmonics variables expansion, and are calculated by means of the analytic procedure presented in section 8.4.2, involving the modified Bessel Functions of half-odd order of the first and second kind $I_\alpha(x)$ and $K_\alpha(x)$ (these functions can be evaluated using NR bessij.f90 and bessik.f90 routines).

Notice that the metaharmonic function appear in the integral conditions inside the integrals H'_k and H_k, that are calculated by means of Gauss-Legendre quadrature formulas as shown in 8.4.42 and 8.4.43. This means that the functions b^r and b^θ have to be evaluated in $nleg = (nr-2)$ Legendre nodes along the radius, while the nodes in r direction, as previously remembered, are Chebyshev nodes.

9.1.4 LU decomposition of the coefficients matrices

To solve the pressure and the velocity problem, the solution of linear systems for the spherical harmonics coefficients of the unknowns are required. Such linear systems for the pressure modes p_{lm} and for the modes of each component of the velocity U^r_{lm}, U^θ_{lm} and u^φ_{lm} are in the form:

$$\mathbf{A}_{qk} \cdot \mathbf{x}_k = \mathbf{B}_q \tag{9.1.16}$$

where:

- $\mathbf{A}_{qk}$ is the coefficients' matrix with dimensions $(nc \times nc)$, being $nc = (nr - 2)$ the number of Chebyshev modes in which the variables p and $\mathbf{u}$ are expanded. The first two rows of $\mathbf{A}_{qk}$ coincide with integral/boundary conditions and the remaining rows contain the first $(nc - 2)$ left hand side of equations 8.4.35, 8.4.57, 8.4.64 and 8.4.65.

- $\mathbf{x}_k$ is the vector of unknowns, containing the nc pressure or velocity Chebyshev modes.

- $\mathbf{B}_q$ is a vector containing the right hand side of the integral/boundary conditions in the first two rows and of the first $(nc-2)$ equations 8.4.35, 8.4.57, 8.4.64 and 8.4.65 in the remaining rows.

The coefficients matrices $\mathbf{A}_{qk}$ vary only with l indices and are independent on the time step, therefore we calculate and invert them with an LU decomposition out of the solving loop once and for all (for this purpose NR LUdcmp.f90 routine have been used).

To construct matrix $\mathbf{A}_{qk}$ we do in sequence the following operations:

1. Calculation of the matrix of coefficients $\mathbf{AA}_{qk}$ for the nc equations 8.4.35, 8.4.57, 8.4.64 and 8.4.65.

2. Calculation of the left hand side of the integral/boundary conditions.

3. Forward displacement of the rows of $\mathbf{AA}_{qk}$ of two rows and replacement of the first rows with the integral/boundary conditions.

The matrix $\mathbf{AA}_{qk}$ has a slightly different form according to the problem that we are treating:

- For p_{lm} problem matrix $\mathbf{AA}_{qk}$ can be calculated such as:

$$\mathbf{AA}^p_{qk} = \alpha^2 \mathbf{M}_{qk} + \alpha \mathbf{N}_{qk} - l(l+1)\alpha^2 \mathbf{Q}_{qk}\delta_{qq} \tag{9.1.17}$$

in which the terms $\mathbf{M}_{qk}$, $\mathbf{N}_{qk}$ and $\mathbf{Q}_{qk}$ coincide with the terms II, I and O in the 8.4.35. The resulting matrix is upper triangular with the coefficients distributed with a regular pattern according with the parity of the rows and column q and k indices, as shown in 8.4.39.

- For U_{lm}^{r} problem matrix $\mathbf{AA}_{qk}$ can be calculated such as:

$$\mathbf{AA}_{qk}^{r} = \sqrt{(2l+1)}\left[\alpha^2\mathbf{M}_{qk} + \alpha\mathbf{N}_{qk} - \gamma\mathbf{E}_{qk} - l(l-1)\alpha^2\mathbf{Q}_{qk}\delta_{qq}\right] \tag{9.1.18}$$

in which the terms $\mathbf{M}_{qk}$, $\mathbf{N}_{qk}$ and $\mathbf{Q}_{qk}$ coincide with the terms II, I and O in the 8.4.57 and the term $\mathbf{E}_{qk}$ is calculated by means of a Gauss Chebyshev quadrature as shown in 8.4.60.

- For U_{lm}^{θ} problem matrix $\mathbf{AA}_{qk}$ can be calculated such as:

$$\mathbf{AA}_{qk}^{r} = \sqrt{(2l+1)}\left[\alpha^2\mathbf{M}_{qk} + \alpha\mathbf{N}_{qk} - \gamma\mathbf{E}_{qk} - (l+1)(l+2)\alpha^2\mathbf{Q}_{qk}\delta_{qq}\right] \tag{9.1.19}$$

- For u_{lm}^{φ} problem matrix $\mathbf{AA}_{qk}$ can be calculated such as:

$$\mathbf{AA}_{qk}^{r} = \alpha^2\mathbf{M}_{qk} + \alpha\mathbf{N}_{qk} - \gamma\mathbf{E}_{qk} - l(l-1)\alpha^2\mathbf{Q}_{qk}\delta_{qq} \tag{9.1.20}$$

The matrices for the three velocity components are full for the presence of the $\mathbf{E}_{qk}$ matrix.

As concerning the calculation of the left hand side of the integral/boundary conditions:

- In the case of pressure problem, the left hand side of the integral conditions involves summations of the integrals H_k' and H_k containing the metaharmonic functions calculated in the previous section, that can be calculated in this phase according to the 8.4.42 and 8.4.43.

- On the other hand the left hand side of the boundary conditions for the velocity consist in a sequence of Chebyshev polynomials of order coinciding with the column index k, evaluated in $y_r = -1$ and $y_r = 1$ respectively.

The last operation can be performed inside a loop on the q and k, rows and columns indices.

The resulting matrices of coefficients for the three velocity components can be decomposed by a standard LU decomposition and are ready to be passed to the velocity problem solver routine inside the main loop.

The matrix $\mathbf{A}_{qk}^{p}$ for the pressure is firstly reduced to a nearly penta-diagonal profile by means of suitable linear combinations as shown in [43], and then decomposed by a LU decomposition.

9.2 Main loop

The core of the numerical code consists in the time loop which calculates the spherical harmonics' coefficients of expansion for pressure and the velocity variables at each time step of the numerical simulation. The original time dependent Navier-Stokes problem is reduced to the solution of a sequence of unsteady Stokes problem. The dependence from the time is contained in the velocity boundary conditions and inside the definition of the non linear forcing term $\mathbf{f}$ defined in equation 8.2.14, whose value at time step $(n+1)$ depends on the value of p and $\mathbf{u}$ variables at the time n and $(n-1)$.

The time propagation of the flow inside a rotating spherical cap requires the definition of an initial condition for the variables assumed equal to the state of rest.

9.2.1 Solution of the pressure problem

The pressure problem results consist in the solution by means of a LU back substitution of the linear system:

$$\mathbf{A}^{p}_{qk} \cdot \mathbf{p}_{lmk} = \mathbf{B}^{p}_{q} \tag{9.2.1}$$

whose coefficients' matrix (independent from the time step) has already been decomposed in a LU form in the Initialization phase.

The pressure problem solver routine has to compute the known term $\mathbf{B}^{p}_{q}$ at the actual time step and to solve the system 9.2.1 resulting from the application of a tau Chebyshev spectral method to the Poisson equation.

The vector $\mathbf{B}^{p}_{q}$ has nc terms. It contains the right hand side of the integral conditions corresponding to the two independent solutions of the metaharmonic problem in the first two rows and the known term of the first $(nc-2)$ equations 8.4.35 in the remaining rows.

In this phase it must be noted that the radial stretching 9.1.1 simplify the coefficients' matrix but complicate the term $\mathbf{B}^{p}_{q}$.

In analogy with the sequence of operations followed for the construction of the coefficients' matrix:

1. We calculate the vector $\mathbf{BB}^{p}_{q}$ for the nc equations 8.4.35.

2. We have to determine the left hand side of the integral conditions.

3. The rows of $\mathbf{BB}^{p}_{q}$ are moved forward of two rows and the first two rows are filled with the integral conditions.

The vector $\mathbf{BB}_q^p$ for the nc equations 8.4.35 has the following expression:

$$\mathbf{BB}_q^p = \sum_{j=0}^{nc-1} [\alpha f_{lmj}^r D_{qj} + (2f_{lmj}^r - S_l f_{lmj}^\theta) C_{qj}] \tag{9.2.2}$$

where the matrix C_{qj} and D_{qj} have already been calculated in the preprocessing phase. The terms f_{lmj}^r and f_{lmj}^θ are the non linear source terms r and θ components, depending on pressure and velocity values at the time steps before the actual one. f_{lmj}^r and f_{lmj}^θ are evaluated at the end of each time step in the main loop.

With regard to the integral conditions right hand side terms, they involve summations of the integral K_k which has been evaluated in preprocessing phase, the f_{lmj}^r and f_{lmj}^θ terms and the term:

$$\alpha \left\{ v_{Slm}^r r(x) \left(\alpha \frac{d}{dx} + 2 \right) b_{lm}^r - v_{Slm}^\theta r(x) \left(S_l b_{lm}^r - \alpha \frac{d}{dx} b_{lm}^\theta \right) \right\} \Bigg|_{-1}^{1} \tag{9.2.3}$$

which has to be calculated in the extremal point -1 and 1. The evaluation of this term requires some care: it contains the r and θ components of the boundary velocity v_{Slm}^r and v_{Slm}^θ, the values of metaharmonic functions and their first derivatives at the boundary. In particular the first derivatives $\frac{db_{lm}^r}{dx}$ and $\frac{db_{lm}^\theta}{dx}$ can be easily calculated expanding the functions b_{lm}^r and b_{lm}^θ in Legendre polynomials and using the following properties:

$$\frac{dP_k(y_t)}{dy_t} = \begin{cases} 1 + \sum_{j=1}^{\frac{k-1}{2}} (4j+1) P_{2j}(y_t) & k \ odd \\ \sum_{j=1}^{\frac{k}{2}} (4j-1) P_{2j-1}(y_t) & k \ even \end{cases} \tag{9.2.4}$$

reminding that:

$$\begin{aligned} P_k(-1) &= \begin{cases} +1 & k \, even \\ -1 & k \, odd \end{cases} \\ P_k(+1) &= +1 \ \forall \, k \end{aligned} \tag{9.2.5}$$

Now we have all the elements to construct the vector $\mathbf{B}_q^p$.

Such vector together with the coefficients' matrix $\mathbf{A}_{qk}^p$ is then passed to the LU backsubstitution routine (we use NR LUbksb.f90 routine) that provides the p_{lmk} Chebyshev modes of each single spherical harmonics mode p_{lm} expansion.

Finally the spherical harmonics mode p_{lm} is calculated by means of the:

$$p_{lm}(x) = \sum_{k=0}^{nc-1} p_{lmk} T_k(x) \tag{9.2.6}$$

9.2.2 Solution of the velocity problem

The velocity problem solution is similar to pressure problem one. The coefficients' matrix (independent from the time step), for each one-dimensional problem for the $\mathbf{u}_{lm}$ components has already been calculated and decomposed in a LU form in the Initialization phase.

The velocity problem solver routine has to compute the known term $\mathbf{B}_q^{\mathbf{u}}$ at the actual time step and to solve the system:

$$\mathbf{A}_{qk}^{\mathbf{u}} \cdot \mathbf{u}_{lmk} = \mathbf{B}_q^{\mathbf{u}} \tag{9.2.7}$$

resulting from the application of a tau Chebyshev spectral method to the governing equation.

The vector $\mathbf{B}_q^{\mathbf{u}}$ has nc terms. It contains the right hand side of the boundary conditions in the first two rows and the known term of the first $(nc-2)$ equations 8.4.57, 8.4.64 and 8.4.65 in the remaining rows.

As usual to construct the vector $\mathbf{B}_q^{\mathbf{u}}$:

1. We calculate the vector $\mathbf{BB}_q^{\mathbf{u}}$ for the nc equations 8.4.57, 8.4.64 and 8.4.65.

2. We have to determine the left hand side of the bundary conditions.

3. The rows of $\mathbf{BB}_q^{\mathbf{u}}$ are moved forward of two rows and the first two rows are filled with the integral conditions.

The vector $\mathbf{BB}_q^{\mathbf{u}}$ for the three different velocity components U_{lm}^r, U_{lm}^θ and u_{lm}^φ has a slightly different form:

- For U_{lm}^r component, $\mathbf{BB}_q^{U_{lm}^r}$ have been calculated by means of the expression:

$$\begin{aligned}\mathbf{BB}_q^{U_{lm}^r} = l^{1/2}\alpha \sum_{j=0}^{nc-1} p_{lmk} D_{qj} + S_l (l+1)^{1/2} \sum_{j=0}^{nc-1} p_{lmk} C_{qj} \\ + l^{1/2} \sum_{j=0}^{nc-1} f_{lmk}^r E_{qj} + (l+1)^{1/2} \sum_{j=0}^{nc-1} f_{lmk}^\theta E_{qj}\end{aligned} \tag{9.2.8}$$

  in which the integrals C_{qj}, D_{qj} and E_{qj} have been calculated in preprocessing phase while p_{lmk}, f_{lmk}^r and f_{lmk}^θ are the pressure and non linear term $\mathbf{f}$ Chebyshev coefficients of expansion.

- For U^{θ}_{lm} component, $\mathbf{BB}^{U^r_{lm}}_q$ have been calculated with the relation:

$$\mathbf{BB}^{U^{\theta}_{lm}}_q = (l+1)^{1/2}\alpha \sum_{j=0}^{nc-1} p_{lmk} D_{qj} - S_l l^{1/2} \sum_{j=0}^{nc-1} p_{lmk} C_{qj} \qquad (9.2.9)$$
$$+ l + 1)^{1/2} \sum_{j=0}^{nc-1} f^r_{lmk} E_{qj} + l^{1/2} \sum_{j=0}^{nc-1} f^{\theta}_{lmk} E_{qj}$$

- Finally, u^{φ}_{lm} components $\mathbf{BB}^{u^{\varphi}_{lm}}_q$ have been computed according to:

$$\mathbf{BB}^{u^{\varphi}_{lm}}_q = -\sum_{j=0}^{nc-1} f^{\varphi}_{lmk} E_{qj} \qquad (9.2.10)$$

Boundary conditions right hand side terms are easy to compute in this case, because of the Dirichelet kind of these conditions. The values assumed of such conditions depends on the time law of motion imposed at the external boundary (the internal one corresponding to radial coordinate $r = r_1$ is supposed to be immobile). Such time law could be for example:

- Uniform rotation, in the form $\mathbf{v}_s = (0, 0, constant)$

- Periodic rotation, in the form $\mathbf{v}_s = (0, 0, Asen(\omega t))$

Once the boundary velocity time law has been decided, it has to be decomposed in its spherical harmonics modes (this operation has to be performed at each time step, updating each time the value of $\mathbf{v}_s$), by means of a call of the vector spherical harmonics transform routine. Lastly, the modes $\mathbf{v}_{slm}$ have to be expanded in Chebyshev polynomials obtaining nc $\mathbf{v}_{slmk}$ coefficients, using the Chebyshev transform routine.

With the elements calculated above the code is able to construct the vector $\mathbf{B}^{\mathbf{u}}_q$.

Such vector together with the coefficients' matrix $\mathbf{A}^{\mathbf{u}}_{qk}$ is then passed to the LU backsubstitution routine (we use NR LUbksb.f90 routine) that provides the $\mathbf{u}_{lmk}$ Chebyshev modes of each single spherical harmonics mode $\mathbf{u}_{lm}$ expansion.

Remind that the problem for the original u^r_{lm} and u^{θ}_{lm} components of the velocity is coupled and we have solved the velocity problem for the uncoupled variables U^r_{lm} and U^{θ}_{lm}, obtained using the similarity transform 8.4.48. Therefore, to obtain u^r_{lm} and u^{θ}_{lm} components, a call to the inverse similarity transform routine is necessary.

Finally the spherical harmonics modes $\mathbf{u}_{lm}$ are calculated by means of the:

$$\mathbf{u}_{lm}(x) = \sum_{k=0}^{nc-1} \mathbf{u}_{lmk} T_k(x) \tag{9.2.11}$$

9.2.3 Non linear term evaluation

In this section of the main loop is concentrated the time dependence of the variables from their previous history. Non linear term $\mathbf{f}$ evaluation is the thorny and demanding operation of the entire algorithm.

The evaluation of $\mathbf{f}$ requires several transform and inverse transform of p and $\mathbf{u}$ variables in spherical harmonics and for its complexity is performed in splitting $\mathbf{f}$ expression in three parts.

$$\mathbf{f}^{n+1} = \underbrace{(\nabla^2 + \gamma)\mathbf{u}^n}_{a^n} - 3\mathrm{Re}\underbrace{(\mathbf{u}^n \cdot \nabla)\mathbf{u}^n}_{b^n} + \mathrm{Re}\underbrace{(\mathbf{u}^{n-1} \cdot \nabla)\mathbf{u}^{n-1}}_{b^{n-1}} \tag{9.2.12}$$

where:

$$\mathbf{a}^n = (\nabla^2 + \gamma)\mathbf{u}^n \tag{9.2.13}$$
$$\mathbf{b}^n = (u^n \cdot \nabla)\mathbf{u}^n = \mathbf{c}^n - \mathbf{d}^n \tag{9.2.14}$$
$$\mathbf{c}^n = \nabla\left(\frac{\mathbf{u}^n \cdot \mathbf{u}^n}{2}\right) \tag{9.2.15}$$
$$\mathbf{d}^n = \mathbf{u}^n \times (\nabla \times \mathbf{u}^n) \tag{9.2.16}$$

In order to obtain the function $\mathbf{f}^{n+1}$, is then necessary to determine the vector function $\mathbf{a}^n$ in each grid's node, at the time n and the vector function $\mathbf{b}^n$, in each grid's node, at the time n and $(n-1)$.

In particular the amplitudes $\mathbf{f}_{lmk}^{n+1}$ have to be determined, proceeding, as usual, with a spherical harmonics expansion in the directions (θ, φ) and using a Chebyshev polynomials expansion for the radial coordinate dependence.

$\mathbf{f}^{n+1}$ are unambiguously determined once we have calculated $\mathbf{a}_{lmk}$, $\mathbf{c}_{lmk}$ and $\mathbf{d}_{lmk}$, omitting the time index n and $(n-1)$ for simplicity.

The terms $\mathbf{a}_{lmk}$, $\mathbf{c}_{lmk}$ and $\mathbf{d}_{lmk}$ are faced separately.

$\mathbf{d}_{lmk}$ evaluation

$$\mathbf{d} = \mathbf{u} \times (\nabla \times \mathbf{u}) = \mathbf{u} \times \omega = \begin{vmatrix} \hat{r} & \hat{\theta} & \hat{\varphi} \\ u_r & u_\theta & u_\varphi \\ \omega_r & \omega_\theta & \omega_\varphi \end{vmatrix} = \tag{9.2.17}$$

$$\begin{aligned} &= (u_\theta \omega_\varphi - u_\varphi \omega_\theta)\hat{r} + \\ &+ (u_\varphi \omega_r - u_r \omega_\varphi)\hat{\theta} + \\ &+ (u_r \omega_\varphi - u_\varphi \omega_r)\hat{\varphi} \end{aligned}$$

therefore:

$$\begin{cases} d_r = (u_\theta \omega_\varphi - u_\varphi \omega_\theta) \\ d_\theta = (u_\varphi \omega_r - u_r \omega_\varphi) \\ d_\varphi = (u_r \omega_\varphi - u_\varphi \omega_r) \end{cases} \tag{9.2.18}$$

The expressions 9.2.18 can be calculated using the values assumed by $\mathbf{u}$ and ω components on the grid's points (y_r, y_t, y_p), and then transforming the results in spherical harmonics and Chebyshev modes again.

The vector $\omega = \nabla \times \mathbf{u}$ is calculated by means of equation B.2.9 obtaining:

$$\begin{cases} \omega^r_{lm}(r) = \frac{S_l}{r} u^\varphi_{lm}(r) \\ \omega^\theta_{lm}(r) = \left(\frac{d}{dr} + \frac{1}{r}\right) u^\varphi_{lm}(r) \\ \omega^\varphi_{lm}(r) = \frac{S_l}{r} u^r_{lm}(r) - \left(\frac{d}{dr} + \frac{1}{r}\right) u^\theta_{lm}(r) \end{cases} \tag{9.2.19}$$

that can be computed expanding the $u^r_{lm}(r)$, $u^\theta_{lm}(r)$ and $u^\varphi_{lm}(r)$, in r direction by means of Chebyshev polynomials expansions.

In detail:

$$\sum_{k=0}^{nc-1} \omega^r_{lmk} T_k(x) = S_l \sum_{k=0}^{nc-1} \frac{u^\varphi_{lmk}(r) T_k(x)}{r(x)} \tag{9.2.20}$$

Multiplying the equations for the term $\frac{T_q(x)}{\sqrt{1-x^2}}$ and integrating between $[-1; 1]$ we obtain:

$$\omega^r_{lmq} = S_l \sum_{k=0}^{nc-1} u^\varphi_{lmk}(r) \int_{-1}^{1} \frac{T_k(x) T_q(x)}{r(x)\sqrt{1-x^2}} dx = S_l \sum_{k=0}^{nc-1} u^\varphi_{lmk}(r) I^1_{kq} \tag{9.2.21}$$

Analogously we have:

$$\omega^\theta_{lmq} = \sum_{k=0}^{nc-1} u^\varphi_{lmk}(r) \left[\int_{-1}^{1} \frac{T_k(x) T_q(x)}{\sqrt{1-x^2}} dx + I^1_{kq} \right] = \tag{9.2.22}$$

$$= \sum_{k=0}^{nc-1} u^\varphi_{lmk}(r) \left[I^0_{kq} + I^1_{kq} \right]$$

and

$$\omega^{\varphi}_{lmq} = \sum_{k=0}^{nc-1} \left[S_l u^r_{lmk}(r) I^1_{kq} - u^{\theta}_{lmk}(r)(I^1_{kq} + I^0_{kq}) \right] \tag{9.2.23}$$

$\mathbf{c}_{lmk}$ evaluation

$$\mathbf{c} = \nabla \left(\frac{\mathbf{u} \cdot \mathbf{u}}{2} \right) = \nabla \left(\frac{U}{2} \right) \tag{9.2.24}$$

Reminding that the gradient of a scalar function expanded in spherical harmonics can be calculated by means of B.2.2 obtaining:

$$\begin{cases} c^r_{lm}(r) = \frac{1}{2} \frac{dU_{lm}(r)}{dr} \\ c^{\theta}_{lm}(r) = \frac{1}{2} \frac{S_l}{r} U_{lm}(r) \\ c^{\varphi}_{lm}(r) = 0 \end{cases} \tag{9.2.25}$$

in which $U_{lm}(r)$ modes have to be expanded in Chebyshev polynomials leading to the following expressions: B.2.2 obtaining:

$$\begin{cases} c^r_{lmk}(r) = \frac{1}{2} \sum\limits_{k=0}^{nc-1} U_{lmk}(r) I^1_{kq} \\ c^{\theta}_{lmk}(r) = \frac{S_l}{2} \sum\limits_{k=0}^{nc-1} U_{lmk}(r) I^1_{kq} \\ c^{\varphi}_{lmk}(r) = 0 \end{cases} \tag{9.2.26}$$

U_{lmk} can be calculate making in each grid point (y_r, y_t, y_p) the scalar product :

$$U(y_r, y_t, y_p) = u^r u^r + u^{\theta} u^{\theta} + u^{\varphi} u^{\varphi} \tag{9.2.27}$$

and then expanding the variable U firstly in spherical harmonics and then in Chebyshev polynomials.

$\mathbf{a}_{lmk}$ evaluation

$$\mathbf{a} = (\nabla^2 + \gamma)\mathbf{u} \tag{9.2.28}$$

Expanding in spherical harmonics we obtain:

$$\begin{cases} a^r_{lm}(r) = \left(D^2_l - \frac{2}{r^2} + \gamma \right) u^r_{lm}(r) + \frac{2S_l}{r^2} u^{\theta}_{lm}(r) \\ a^{\theta}_{lm}(r) = \frac{2S_l}{r^2} u^r_{lm}(r) + \left(D^2_l + \gamma \right) u^{\theta}_{lm}(r) \\ a^{\varphi}_{lm}(r) = \left(D^2_l + \gamma \right) u^{\varphi}_{lm}(r) \end{cases} \tag{9.2.29}$$

in which $\mathbf{u}_{lm}(r)$ modes have to be expanded in Chebyshev polynomials leading to the following expressions:

$$a^r_{lmk}(r) = \sum_{k=0}^{nc-1} u^r_{lm}(r) \int_{-1}^{1} \frac{T''_k(x)T_q(x)}{\sqrt{1-x^2}}dx + 2\sum_{k=0}^{nc-1} u^r_{lm}(r) \int_{-1}^{1} \frac{T'_k(x)T_q(x)}{r(x)\sqrt{1-x^2}}dx + \tag{9.2.30}$$

$$+ (l(l+1)+2)\sum_{k=0}^{nc-1} u^r_{lm}(r) \int_{-1}^{1} \frac{T_k(x)T_q(x)}{r^2(x)\sqrt{1-x^2}}dx + \gamma u^r_{lm}(r) \int_{-1}^{1} \frac{T_k(x)T_q(x)}{\sqrt{1-x^2}} +$$

$$+ 2S_l \sum_{k=0}^{nc-1} u^\theta_{lm}(r) \int_{-1}^{1} \frac{T_k(x)T_q(x)}{r^2(x)\sqrt{1-x^2}}dx =$$

$$= \sum_{k=0}^{nc-1} u^r_{lm}(r) I^{2a}_{kq} + 2\sum_{k=0}^{nc-1} u^r_{lm}(r) I^{1a}_{kq} + (l(l+1)+2)\sum_{k=0}^{nc-1} u^r_{lm}(r) I^{0a}_{kq} + \gamma u^r_{lm}(r) I^0_{kq} +$$

$$+ 2S_l \sum_{k=0}^{nc-1} u^\theta_{lm}(r) I^{0a}_{kq}$$

$$a^\theta_{lm}(r) = 2S_l \sum_{k=0}^{nc-1} u^r_{lm}(r) I^{0a}_{kq} + \sum_{k=0}^{nc-1} u^\theta_{lm}(r) I^{2a}_{kq} + 2\sum_{k=0}^{nc-1} u^\theta_{lm}(r) I^{1a}_{kq} + \tag{9.2.31}$$

$$+ l(l+1)\sum_{k=0}^{nc-1} u^\theta_{lm}(r) I^{0a}_{kq} + \gamma u^\theta_{lm}(r) I^0_{kq}$$

$$a^\varphi_{lm}(r) = \sum_{k=0}^{nc-1} u^\varphi_{lm}(r) I^{2a}_{kq} + 2\sum_{k=0}^{nc-1} u^\varphi_{lm}(r) I^{1a}_{kq} + l(l+1)\sum_{k=0}^{nc-1} u_{lm}\varphi(r) I^{0a}_{kq} + \tag{9.2.32}$$

$$+ \gamma u^\varphi_{lm}(r) I^0_{kq}$$

9.2.4 Updating of variables

At the end of the time loop the variables p and $\mathbf{u}$ have to be updated so that the above described procedure to calculate $\mathbf{f}$ term can be revised at each time step. The updating procedure consists in throwing away the actual variables value at $(n-1)$ time step and doing the following allocations:

$$p^{n-1} = p^n p^n = p^{n+1}\mathbf{u}^{n-1} = \mathbf{u}^n\mathbf{u}^n = \mathbf{u}^{n+1} \tag{9.2.33}$$

9.3 End of program

9.3.1 Output and Postprocessing

The numerical code provide the solution of pressure and velocity problems using a tau spectral method, allowing to reconstruct the three-dimensional flow field inside a spherical cap rotating around its axis with a prescribed time law. The output of the code consists in the pressure and velocity fields at each time of the simulation disposed on a regular grid in the variable (y_r, y_t, y_p).

The data can be then post-processed to calculate derived quantities such as vorticity and shear stresses.

Chapter 10

A test case: solutions of the metaharmonic problem

The metaharmonic problem:

$$\left(-\nabla^2 + \gamma\right) b = 0 \qquad \int n \cdot b dS = 0 \qquad n \times b \,|_S = 0 \tag{10.0.1}$$

has been used as a test problem to verify the feasibility and the efficiency of the algorithms used for the solution of the complete problem. This choice is motivated by the fact that the differential operator of the metaharmonic problem is the same as the one for the velocity problem and that an analytic solution was indeed available. The **analytic solution**[1] of the problem has been compared with a **finite difference solution** , with first order derivatives, and two **spectral solution**, obtained by expanding the unknown function b in terms of Chebyshev polynomials and Legendre polynomials.

10.0.2 Finite difference solution

Let us consider the modal equations for the metaharmonic problem in their uncoupled form for the unknown β_l^r and β_l^θ:

$$r^2 \frac{d^2\beta_l^r}{dr^2} + 2r\frac{d\beta_l^r}{dr} - \left[\gamma r^2 + l(l-1)\right]\beta_l^r = 0 \tag{10.0.2}$$

$$r^2 \frac{d^2\beta_l^\theta}{dr^2} + 2r\frac{d\beta_l^\theta}{dr} - \left[\gamma r^2 + (l+1)(l+2)\right]\beta_l^\theta = 0 \tag{10.0.3}$$

Now we discretize the derivatives appearing in the 10.0.2 by means of a

[1] see chapter 8 for the details of the analytic solution

centered finite difference scheme:

$$\frac{d\mathrm{F(x)}}{dx} = \frac{F_{i+1} - F_{i-1}}{2\Delta \mathrm{x}} + O(\Delta \mathrm{x}) \tag{10.0.4}$$

$$\frac{d^2\mathrm{F(x)}}{dx^2} = \frac{F_{i+1} - 2F_i + F_{i-1}}{\Delta \mathrm{x}^2} + O(\Delta \mathrm{x}^2) \tag{10.0.5}$$

After some algebraic calculation the 10.0.2 become:

$$\left(r^2 + 2r\Delta r\right)\beta^r_{l(i+1)} - \left(2r^2 + 2r\Delta r + l(l-1)\Delta r^2 - \gamma r^2 \Delta r^2\right)\beta^r_{l(i)} + r^2\beta^r_{l(i-1)} = 0 \tag{10.0.6}$$

$$\left(r^2 + 2r\Delta r\right)\beta^\theta_{l(i+1)} - \left(2r^2 + 2r\Delta r + (l+1)(l+2)\Delta r^2 - \gamma r^2 \Delta r^2\right)\beta^\theta_{l(i)} + r^2\beta^\theta_{l(i-1)} = 0 \tag{10.0.7}$$

with the boundary conditions:

$$\begin{aligned} \text{first set} \quad \rightarrow \quad & \beta^r_l(x_1) = \left(\frac{l}{2l+1}\right)^{1/2} \quad \beta^r_l(x_2) = 0 \\ & \beta^\theta_l(x_1) = \left(\frac{l+1}{2l+1}\right)^{1/2} \quad \beta^\theta_l(x_2) = 0 \end{aligned} \tag{10.0.8}$$

$$\begin{aligned} \text{second set} \quad \rightarrow \quad & \beta^r_l(x_1) = 0 \quad \beta^r_l(x_2) = \left(\frac{l}{2l+1}\right)^{1/2} \\ & \beta^\theta_l(x_1) = 0 \quad \beta^\theta_l(x_2) = \left(\frac{l+1}{2l+1}\right)^{1/2} \end{aligned} \tag{10.0.9}$$

The problem 10.0.2 is a linear system of equation for all the evenly spaced points x_i, with $i = 0, 1, \ldots, N$ along the radius. Denoting the coefficients of $\beta_{l(i+1)}$, $\beta_{l(i)}$ and $\beta_{l(i-1)}$ with the concise notation a, d and b respectively, and the right hand side of the equations with c:

$$b\beta_{l(i-1)} + d\beta_{l(i)} + a\beta_{l(i+1)} = c \tag{10.0.10}$$

the matrix form of the system is:

$$\begin{pmatrix} d_1 & a_1 & 0 & 0 & \ldots & 0 \\ b_2 & d_2 & a_2 & 0 & \ldots & 0 \\ 0 & b_3 & d_3 & a_3 & \ldots & 0 \\ 0 & 0 & b_4 & d_4 & & 0 \\ \vdots & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & b_N & d_N \end{pmatrix} \begin{pmatrix} \beta_{l(1)} \\ \beta_{l(2)} \\ \beta_{l(3)} \\ \beta_{l(4)} \\ \vdots \\ \beta_{l(N)} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ \vdots \\ c_5 \end{pmatrix} \tag{10.0.11}$$

The tridiagonal structure of this system suggests the use of the Thomas algorithm to solve the problem.

10.0.3 Spectral solution: Chebyshev expansion

To solve the metaharmonic problem 10.0.2 with a Chebyshev spectral method, the functions β_l^r and β_l^θ have to be expanded in terms of Chebyshev polynomials.

$$\beta_l^r(x) = \sum_{i=0}^{N} \beta_i^r T_i(x) \tag{10.0.12}$$

$$\beta_l^\theta(x) = \sum_{i=0}^{N} \beta_i^\theta T_i(x) \tag{10.0.13}$$

This expansion is allowed in the range $[-1;1]$, so we have firstly to introduce a coordinate transformation to resize the radial domain of the problem to such an interval:

$$x = x(r) = \frac{\ln\left(\frac{r^2}{r_1 r_2}\right)}{\ln\left(\frac{r_2}{r_1}\right)} \quad \rightarrow \quad r = r(x) = r_1 e^{(x+1)/\alpha} \text{ where } \alpha = \frac{2}{\ln\left(\frac{r_2}{r_1}\right)} \tag{10.0.14}$$

Besides mapping the domain $r_1 \leq r \leq r_2$ to Chebyshev one $-1 \leq x \leq 1$, this transformation make the coefficients of the differential operator in the 10.0.2 to become constant, in fact:

$$r\frac{d}{dr} \rightarrow x\frac{d}{dx} \quad r^2\frac{d^2}{dr^2} \rightarrow \alpha^2\frac{d^2}{dx^2} - \alpha\frac{d}{dx} \quad r^2 D_l \rightarrow \alpha^2\frac{d^2}{dx^2} - \alpha\frac{d}{dx} - l(l+1) \tag{10.0.15}$$

Substituting the 10.0.15 in the 10.0.2 we find:

$$\alpha^2\frac{d^2\beta_l^r}{dx^2} + \alpha\frac{d\beta_l^r}{dx} - \left[\gamma(r_1 e^{(x+1)/\alpha})^2 + l(l-1)\right]\beta_l^r = 0 \tag{10.0.16}$$

$$\alpha^2\frac{d^2\beta_l^r}{dx^2} + \alpha\frac{d\beta_l^r}{dx} - \left[\gamma(r_1 e^{(x+1)/\alpha})^2 + (l+1)(l+2)\right]\beta_l^\theta = 0 \tag{10.0.17}$$

Introducing the expansion 10.0.12 in the 10.0.16 we obtain:

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^r \frac{d^2 T_k(x)}{dx^2} + \alpha \sum_{k=0}^{N} \beta_{lk}^r \frac{dT_k(x)}{dx} \tag{10.0.18}$$

$$- \left[\gamma(r_1 e^{(x+1)/\alpha})^2 + l(l-1)\right] \sum_{k=0}^{N} \beta_{lk}^r T_k(x) = 0 \tag{10.0.19}$$

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^{\theta} \frac{d^2 T_k(x)}{dx^2} + \alpha \sum_{k=0}^{N} \beta_{lk}^{\theta} \frac{dT_k(x)}{dx} - \left[\gamma (r_1 e^{(x+1)/\alpha})^2 + (l+1)(l+2)\right] \sum_{k=0}^{N} \beta_{lk}^{\theta} T_k(x) = 0 \tag{10.0.20}$$

Finally we multiply the equations for the term $\frac{T_q(x)}{\sqrt{1-x^2}}$ and integrate between $[-1;1]$. In order to use the orthogonality properties of Chebyshev polynomials, first and second derivatives of these polynomials have to be written in terms of Chebyshev polynomials themselves, this was achieved by calculating the analytic form of the derivatives and introducing the recurrence relation:

$$2xT_n(x) = T_{n+1}(x) + T_{n-1}(x) \tag{10.0.21}$$

After some algebraic calculation the equations are reformulated such as:

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^{r} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} 2k \sum_{j=1}^{\frac{k-1}{2}} \sum_{i=1}^{j} 4jT_{2i-1}(x)\, dx & k \text{ odd} \\ 2k \sum_{j=1}^{\frac{k}{2}} (2j+1) \left[1 + 2\sum_{i=1}^{j} T_{2i}(x)\right] dx & k \text{ even} \end{cases}}_{II} + \tag{10.0.22}$$

$$+ \alpha \sum_{k=0}^{N} \beta_{lk}^{r} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} k \left[1 + 2 \sum_{j=1}^{\frac{k-1}{2}} T_{2j}(x)\right] dx & k \text{ odd} \\ 2k \sum_{j=1}^{\frac{k}{2}} T_{2j-1}(x) dx & k \text{ even} \end{cases}}_{I} +$$

$$- [l(l-1)] \sum_{k=0}^{N} \beta_{lk}^{r} \underbrace{\int_{-1}^{1} \frac{T_q(x) T_k(x)}{\sqrt{1-x^2}} dx}_{O1} - \gamma \sum_{k=0}^{N} \beta_{lk}^{r} \underbrace{\int_{-1}^{1} \frac{r^2(x) T_q(x) T_k(x)}{\sqrt{1-x^2}} dx}_{O2} = 0$$

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^{\theta} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} 2k \sum_{j=1}^{\frac{k-1}{2}} \sum_{i=1}^{j} 4jT_{2i-1}(x)\, dx & k\ odd \\ 2k \sum_{j=1}^{\frac{k}{2}} (2j+1) \left[1 + 2\sum_{i=1}^{j} T_{2i}(x)\right] dx & k\ even \end{cases}}_{II} +$$

$$+ \alpha \sum_{k=0}^{N} \beta_{lk}^{\theta} \underbrace{\int_{-1}^{1} \frac{T_q(x)}{1-x^2} \begin{cases} k \left[1 + 2 \sum_{j=1}^{\frac{k-1}{2}} T_{2j}(x)\right] dx & k\ odd \\ 2k \sum_{j=1}^{\frac{k}{2}} T_{2j-1}(x) dx & k\ even \end{cases}}_{I} + \tag{10.0.23}$$

$$- \left[(l+1)(l+2)\right] \sum_{k=0}^{N} \beta_{lk}^{\theta} \underbrace{\int_{-1}^{1} \frac{T_q(x)T_k(x)}{\sqrt{1-x^2}} dx}_{O1} - \gamma \sum_{k=0}^{N} \beta_{lk}^{\theta} \underbrace{\int_{-1}^{1} \frac{r^2(x)T_q(x)T_k(x)}{\sqrt{1-x^2}} dx}_{O2} = 0$$

where:

$$II = \begin{cases} k\ odd \begin{cases} q\ odd \rightarrow \pi k \delta_{(2i-1),q} \sum_{l=\frac{q+1}{2}}^{\frac{k-1}{2}} 4l \\ q\, even \rightarrow 0 \end{cases} \\ k\ even \begin{cases} q\ odd \rightarrow 0 \\ q\, even \rightarrow 2\pi k \delta_{2i,q} \sum_{l=\frac{q}{2}}^{\frac{k}{2}} (2l+1) \end{cases} \end{cases}$$

$$I = \begin{cases} k\ odd \begin{cases} q\ odd \rightarrow 0 \\ q\, even \rightarrow k\pi \delta_{2j,q} \end{cases} \\ k\ even \begin{cases} q\ odd \rightarrow k\pi \delta_{(2j-1),q} \\ q\, even \rightarrow 0 \end{cases} \end{cases} \tag{10.0.24}$$

$$O1 = \begin{cases} \frac{\pi}{2}\delta_{q,k} \\ \pi \delta_{q,0} \end{cases}$$

and the integral O2 can be evaluated by means of a Gauss-Chebyshev quadrature formula:

$$\int_{-1}^{1} \frac{r^2(x)T_q(x)T_k(x)}{\sqrt{1-x^2}} dx = \int_{-1}^{1} \frac{u(x)}{\sqrt{1-x^2}} dx = \frac{\pi}{N+1} \sum_{k=1}^{N+1} u(x_k) \tag{10.0.25}$$

and the points x_k are the Chebyshev nodes.

The 10.0.23 equations with the 10.0.24 define a system of equations with the variation of q between 0 and N, being $N + 1$ the number of Chebyshev nodes inside the range $[-1, 1]$. The matrix of coefficients of such a system is complete for the presence of the O2 terms, while other terms form a pattern in the upper triangular part of the matrix. The problem can be solved using a standard LU decomposition.

10.0.4 Spectral solution: Legendre expansion

To solve the metaharmonic problem 10.0.2 with a Legendre spectral method, the functions β_l^r and β_l^θ have to be expanded in terms of Legendre polynomials.

$$\beta_l^r(x) = \sum_{i=0}^{N} \beta_i^r P_i(x) \tag{10.0.26}$$

$$\beta_l^\theta(x) = \sum_{i=0}^{N} \beta_i^\theta P_i(x) \tag{10.0.27}$$

Similarly to the Chebyshev spectral solution, firstly we need to scale the domain $r_1 \leq r \leq r_2$ to the range $[-1; 1]$. This is achieved by the coordinate transformation 10.0.14.

Substituting the 10.0.15 in the 10.0.2 we find the equations 10.0.16 where the functions β_l^r and β_l^θ are expanded in terms of Legendre polynomials obtaining:

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^r \frac{d^2 P_k(x)}{dx^2} + \alpha \sum_{k=0}^{N} \beta_{lk}^r \frac{dP_k(x)}{dx} - \left[\gamma(r_1 e^{(x+1)/\alpha})^2 + l(l-1)\right] \sum_{k=0}^{N} \beta_{lk}^r P_k(x) = 0 \tag{10.0.28}$$

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^\theta \frac{d^2 P_k(x)}{dx^2} + \alpha \sum_{k=0}^{N} \beta_{lk}^\theta \frac{dP_k(x)}{dx} - \left[\gamma(r_1 e^{(x+1)/\alpha})^2 + (l+1)(l+2)\right] \sum_{k=0}^{N} \beta_{lk}^\theta P_k(x) = 0 \tag{10.0.29}$$

The expansion coefficients β_{lk}^r and β_{lk}^θ can be found by solving the system of equations obtained multiplying the equations 10.0.28 for the term $P_q(x)$ and integrating between $[-1; 1]$. This operation is necessary to take advantage of the orthogonality properties of Legendre polynomials, after expressing the first and second derivatives of Legendre polynomials in terms of Legendre polynomials themselves, by means of the recurrence relation:

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)P_n(x) \tag{10.0.30}$$

After some algebraic calculation the equations are reformulated such as:

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^r \underbrace{\int_{-1}^{1} P_q(x) \begin{cases} \sum_{j=1}^{\frac{k-1}{2}} (4j+1) \sum_{i=1}^{j} (4i-1) P_{2i-1}(x)\, dx & k\ odd \\ \sum_{j=1}^{\frac{k}{2}} (4j-1) \left[1 + \sum_{i=1}^{j-1} (4i+1) P_{2i}(x) \right] dx & k\ even \end{cases}}_{II} +$$

(10.0.31)

$$\alpha \sum_{k=0}^{N} \beta_{lk}^r \underbrace{\int_{-1}^{1} P_q(x) \begin{cases} 1 + \sum_{j=1}^{\frac{k-1}{2}} (4j+1) P_{2j}(x)\, dx & k\ odd \\ \sum_{j=1}^{\frac{k}{2}} (4j-1) P_{2j-1}(x) dx & k\ even \end{cases}}_{I} +$$

$$- [l(l-1)] \sum_{k=0}^{N} \beta_{lk}^r \underbrace{\int_{-1}^{1} P_q(x) P_k(x) dx}_{O1} - \gamma \sum_{k=0}^{N} \beta_{lk}^r \underbrace{\int_{-1}^{1} r^2(x) P_q(x) P_k(x) dx}_{O2} = 0$$

$$\alpha^2 \sum_{k=0}^{N} \beta_{lk}^\theta \underbrace{\int_{-1}^{1} P_q(x) \begin{cases} \sum_{j=1}^{\frac{k-1}{2}} (4j+1) \sum_{i=1}^{j} (4i-1) P_{2i-1}(x)\, dx & k\ odd \\ \sum_{j=1}^{\frac{k}{2}} (4j-1) \left[1 + \sum_{i=1}^{j-1} (4i+1) P_{2i}(x) \right] dx & k\ even \end{cases}}_{II} +$$

(10.0.32)

$$\alpha \sum_{k=0}^{N} \beta_{lk}^\theta \underbrace{\int_{-1}^{1} P_q(x) \begin{cases} 1 + \sum_{j=1}^{\frac{k-1}{2}} (4j+1) P_{2j}(x)\, dx & k\ odd \\ \sum_{j=1}^{\frac{k}{2}} (4j-1) P_{2j-1}(x) dx & k\ even \end{cases}}_{I} +$$

$$- [(l+1)(l+2)] \sum_{k=0}^{N} \beta_{lk}^\theta \underbrace{\int_{-1}^{1} P_q(x) P_k(x) dx}_{O1} - \gamma \sum_{k=0}^{N} \beta_{lk}^\theta \underbrace{\int_{-1}^{1} r^2(x) P_q(x) P_k(x) dx}_{O2} = 0$$

where:

$$
\begin{aligned}
II \;\; &= \begin{cases} k\ odd \begin{cases} q\ odd \;\rightarrow 2\delta_{(2i-1),q} \sum\limits_{l=\frac{q+1}{2}}^{\frac{k-1}{2}} (4l+1) \\ q\,even \;\rightarrow\; 0 \end{cases} \\ k\ even \begin{cases} q\ odd \rightarrow 0 \\ q\,even \;\rightarrow\; 2\delta_{2i,q} \sum\limits_{l=\frac{q}{2}}^{\frac{k}{2}} (4l-1) \end{cases} \end{cases} \\
I \;\; &= \begin{cases} k\ odd \begin{cases} q\ odd \rightarrow 0 \\ q\,even \;\rightarrow\; 2\delta_{2j,q} \end{cases} \\ k\ even \begin{cases} q\ odd \rightarrow 2\delta_{(2j-1),q} \\ q\,even \;\rightarrow\; 0 \end{cases} \end{cases} \\
O1 &= \tfrac{2}{(2q+1)}
\end{aligned}
\tag{10.0.33}
$$

and the integral O2 can be evaluated by means of a Gauss-Legendre quadrature formula:

$$
\int_{-1}^{1} r^2(x) P_q(x) P_k(x) dx = \int_{-1}^{1} u(x) dx = \sum_{k=1}^{N+1} w_k\, u(x_k) \tag{10.0.34}
$$

where w_k and x_k are the Legendre weights and nodes respectively.

The system of equations defined by 10.0.32 with the variation of q between 0 and N, being $N+1$ the number of Legendre nodes inside the range $[-1, 1]$ has a matrix of coefficients with the same pattern as Chebyshev spectral solution one. The problem can be solved using a standard LU decomposition.

10.0.5 Comparison between the different methods used to solve the metaharmonic problem

Figure 10.1 reports a comparison between the analytic solution of the metaharmonic problem in terms of Bessel function and the solution obtained by solving the system 10.0.11 with a finite difference technique varying the number of points (assumed evenly spaced) along the radius.

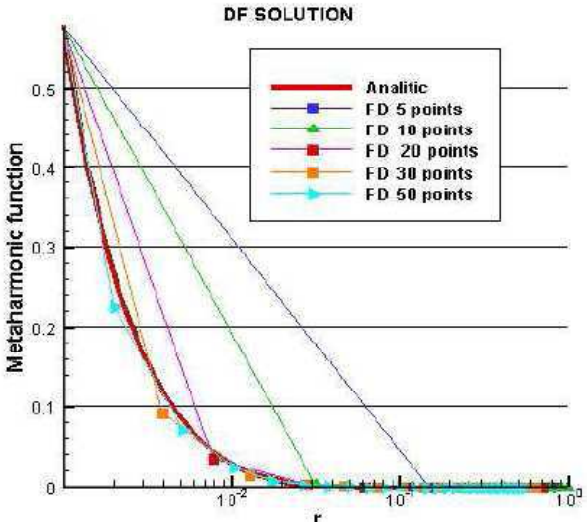

Figure 10.1: Comparison between the analytic solution for the radial component of the metaharmonic problem (red line) and the finite difference solution obtained using different number of points along the radius. Horizontal axis is plotted in log scale.

Horizontal axis has been plotted in log scale. Notice that the in the vicinity of the first boundary $r = r_1$ where the metaharmonic function assumes non zero values the accordance between the finite difference and the analytic solution is not very good when using few points, even if the FD solution is able to capture the general behaviour of the function. The accordance between the solution become satisfactory with 50 points.

Figures 10.2 and 10.3 report the comparison between the analytic solution and the spectral solution for the Chebyshev and the Legendre approximation respectively, obtained using a different number of nodes along the radius.

Remind that the number of nodes and the number of modes used to recon-

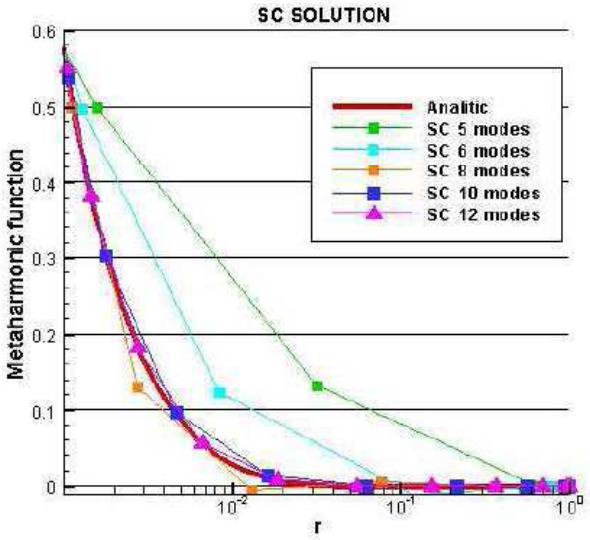

Figure 10.2: Comparison between the analytic solution for the radial component of the metaharmonic problem (red line) and the spectral Chebyshev solution obtained using 5 - 6 - 8 - 10 and 12 modes

struct the function are related together, in fact:

$$n_r = n_{modes} + 2 \tag{10.0.35}$$

The accordance between the approximating solution and the exact one is very good in both cases, even using few modes to reconstruct the solution, no relevant difference can be appreciated between the two methods. Notice, in fact, that 10 modes are sufficient to describe faithfully the exact solution. Let us indicate with f the analytic solution and with f_n the numerical solution calculated with different techniques. In order to verify the convergence between the numerical and the analytic solution the relative error $rerr$ defined according to:

$$rerr = \frac{\|f - f_n\|}{\|f\|} \tag{10.0.36}$$

is introduced.

$\|f\|$ denotes the L^2 norm of $f(x)$, which has been calculated in the three

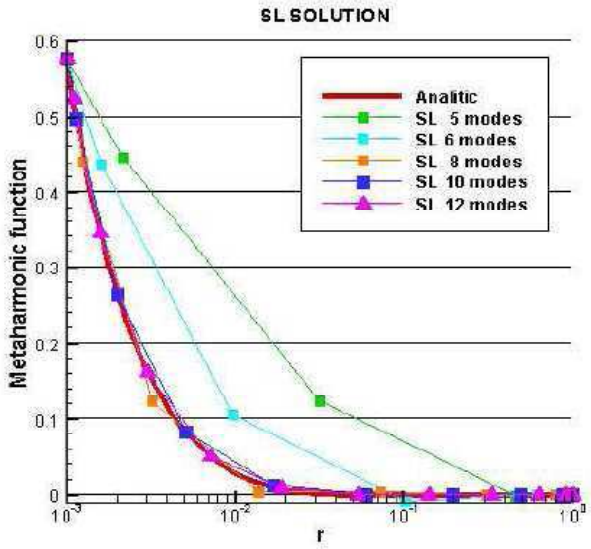

Figure 10.3: Comparison between the analytic solution for the radial component of the metaharmonic problem (red line) and the spectral Legendre solution obtained using 5 - 6 - 8 - 10 and 12 modes

examined case with the following expression:

$$\|f\|^2 = \frac{1}{N+1}\sum_{i=1}^{N+1}[f(x_i)]^2 \qquad Finite\ Difference \tag{10.0.37}$$

$$\|f\|^2 = \frac{\pi}{N+1}\sum_{i=1}^{N+1}[f(x_i)]^2 \qquad Spectral\,Chebyshev \tag{10.0.38}$$

$$\tag{10.0.39}$$

where, $x_i = cos[(i-1)\pi/N]$, $i = 1, 2, \ldots, N+1$ are the Chebyshev nodes.

$$\|f\|^2 = \sum_{i=1}^{N+1} w_i\,[f(x_i)]^2 \qquad Spectral\ Legendre \tag{10.0.40}$$

where, x_i and w_i are the Legendre nodes and weight respectively. The error committed approximating the exact solution with the numerical one depends on the number of the points along the radius (and also on their distribution, the best one is the Chebyshev nodes distribution) for finite difference technique and on the number of modes used to reconstruct the solution for spectral techniques. Figure 10.4 shows the relative error related to the number of nodes for

the three numerical methods used to solve the metaharmonic problem, referred to the reconstructed metaharmonic function. The errors trend confirms a low convergence of Finite Difference solution, while spectral methods are faster.

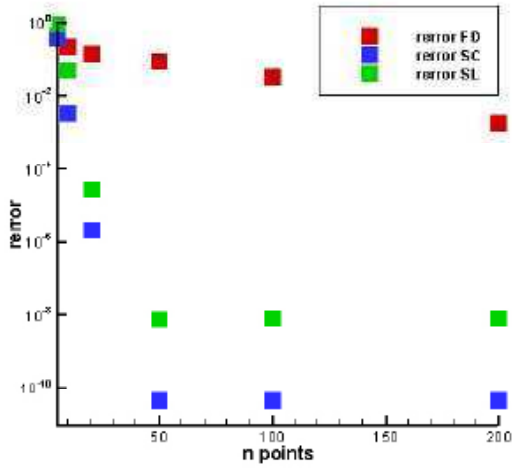

Figure 10.4: Relative error committed approximating the analytic solution with the numerical one related to the number of points along the radius. Finite Difference red dots, Spectral Chebyshev blue dots, Spectral Legendre green dots

Finally we report three tables containing a comparison of *rerr* on the single l mode of the metaharmonic function expansion, calculated for the three numerical solutions. Table 10.1 reports the numerical errors *rerr* for the Finite Difference solution calculated with an increasing number of points. The results confirm the slow convergence of the Finite Difference to the analytic solution.

Table 10.2 reports the numerical errors rerr for the Spectral Chebyshev solution calculated with an increasing number of points.

Table 10.3 reports the numerical errors rerr for the Spectral Legendre solution calculated with an increasing number of points.

The results demonstrate that Spectral methods are able to gain an high precision using few (Chebyshev / Legendre) modes to reconstruct the single l modes. A slight worsening is noticed for l increasing maybe caused by the computer precision that conflicts with the solution trend. For great value of l, in fact, the analytic solution approximately assumes zero values on all the

N points	$rerr\ l = 1$	$rerr\ l = 3$	$rerr\ l = 10$
10	$7.4(10^{-2})$	$2.0(10^{-2})$	$4.6(10^{-2})$
12	$4.9(10^{-2})$	$3.3(10^{-2})$	$3.5(10^{-2})$
14	$2.8(10^{-2})$	$4.7(10^{-2})$	$4.4(10^{-2})$
16	$5.9(10^{-2})$	$6.0(10^{-2})$	$5.4(10^{-2})$
18	$1.0(10^{-2})$	$7.2(10^{-2})$	$6.7(10^{-2})$
20	$1.5(10^{-2})$	$8.2(10^{-2})$	$8.1(10^{-2})$
50	$7.3(10^{-2})$	$3.5(10^{-3})$	$6.5(10^{-2})$
100	$5.4(10^{-3})$	$8.6(10^{-3})$	$3.5(10^{-3})$
200	$2.9(10^{-3})$	$2.6(10^{-3})$	$8.8(10^{-3})$
500	$9.4(10^{-4})$	$7.1(10^{-4})$	$5.3(10^{-4})$

Table 10.1: Numerical error for Finite Difference solution for modes $l = 1$, $l = 3$ and $l = 10$

N points	$rerr\ l = 1$	$rerr\ l = 3$	$rerr\ l = 10$
8	$7.3(10^{-2})$	$4.4(10^{-1})$	$9.8(10^{-1})$
10	$5.1(10^{-3})$	$1.1(10^{-1})$	$8.1(10^{-1})$
12	$5.0(10^{-4})$	$2.1(10^{-2})$	$4.9(10^{-1})$
14	$1.5(10^{-4})$	$2.9(10^{-3})$	$2.4(10^{-1})$
16	$4.1(10^{-5})$	$3.2(10^{-4})$	$4.1(10^{-2})$
18	$1.1(10^{-5})$	$2.9(10^{-5})$	$1.5(10^{-2})$
20	$2.6(10^{-6})$	$2.1(10^{-6})$	$5.0(10^{-3})$
50	$4.1(10^{-11})$	$2.2(10^{-12})$	$5.4(10^{-13})$
100	$4.0(10^{-11})$	$2.0(10^{-12})$	$6.0(10^{-13})$

Table 10.2: Numerical error for Spectral Chebyshev solution for modes $l = 1$, $l = 3$ and $l = 10$

range.

N points	$rerr\ l = 1$	$rerr\ l = 3$	$rerr\ l = 10$
8	$9.1(10^{-2})$	$7.9(10^{-1})$	$7.1(10^{-1})$
10	$7.5(10^{-3})$	$2.2(10^{-1})$	$2.3(10^{-1})$
12	$9.4(10^{-4})$	$4.0(10^{-2})$	$1.2(10^{-1})$
14	$2.7(10^{-4})$	$5.7(10^{-3})$	$5.8(10^{-2})$
16	$7.4(10^{-5})$	$6.3(10^{-4})$	$2.5(10^{-2})$
18	$1.9(10^{-5})$	$5.6(10^{-5})$	$1.0(10^{-2})$
20	$4.3(10^{-6})$	$4.1(10^{-6})$	$3.8(10^{-3})$
50	$1.5(10^{-8})$	$1.6(10^{-8})$	$1.8(10^{-8})$
100	$1.0(10^{-9})$	$1.1(10^{-8})$	$1.2(10^{-8})$

Table 10.3: Numerical error for Spectral Legendre solution for modes $l = 1$, $l = 3$ and $l = 10$

Chapter 11

Conclusions and future development

The second part of this PhD Thesis is devoted to the formulation of a numerical model of the viscous flow inside an oscillating sphere. The purpose of the numerical is to describe quantitatively the three-dimensional behaviour of the flows, overcoming the limits of the physical model described in the first part of the present work.

The numerical model used a spectral technique to solve the Navier Stokes equations in primitive velocity-pressure form. The incompressibility constraint in the governing equation has been replaced by the Poisson equation for the pressure. This approach allows the derivation of an uncoupled formulation by joining the Poisson equation with an independent condition of integral character for the pressure. The latter requires the evaluation of a set of metaharmonic functions, which, in turn, are solutions of the metaharmonic (Helmoltz) problem.

The equations are discretized in time by means of a non-fractional-step scheme. Non-linear terms are evaluated explicitly while the linear terms are accounted implicitly. The pressure and velocity variables are expanded in series of spherical harmonics in the angular directions θ and φ. This operation allow to transform the original problem in a sequence of stationary problems for the coefficients of the expansion which depend only from the radial direction. The solution is calculated by means of a tau projection spectral method, after having expanded the coefficients in terms of Chebyshev polynomials.

The mathematical formulation of the numerical model required a special effort and allow us to built a numerical code, whose structure is described in chapter 9 which didn't provide any results because of the lack of time. The tests carried out on a simplified problem (the metaharmonic problem) demonstrate that spectral techniques are efficient and feasible.

In the near future the efforts made to formulate the complete problem will allow to obtain flow fields inside a rotating sphere and could be developed to extend the treatment to the deformed sphere.

Part III

Appendices

Appendix A

Differential operators in spherical coordinates

This appendix contains the expression of differential operators in spherical coordinates for scalar and vector functions. General definitions about basis, volume and area element in those coordinates are also included.

A.1 Definitions

The spherical basis consists of the vectors:

$$\begin{aligned} \mathbf{r} &= \sin\theta\cos\phi\mathbf{i} + \sin\theta\sin\phi\mathbf{j} + \cos\theta\mathbf{k} \\ \theta &= \cos\theta\cos\phi\mathbf{i} + \cos\theta\sin\phi\mathbf{j} - \sin\theta\mathbf{k} \\ \phi &= -\sin\phi\mathbf{i} + \cos\phi\mathbf{jk} \end{aligned} \tag{A.1.1}$$

The volume element in spherical coordinates is:

$$dV = r^2\sin\theta dr d\theta d\phi \tag{A.1.2}$$

The area element on a sphere with r constant is:

$$dS_r = r^2\sin\theta d\theta d\phi \tag{A.1.3}$$

The area element on a cone with θ constant is:

$$dS_\theta = r\sin\theta dr d\phi \tag{A.1.4}$$

The area element on the half plane with ϕ constant is:

$$dS_\phi = r d\theta dr \tag{A.1.5}$$

A.2 Differential operators

A.2.1 Gradient

The gradient of a scalar field $f(r, \theta, \phi)$ in spherical coordinates is given by:

$$\nabla f(r,\theta,\phi) = \frac{\partial f}{\partial r}\mathbf{r} + \frac{1}{r}\frac{\partial f}{\partial \theta}\theta + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\phi \tag{A.2.1}$$

A.2.2 Divergence

The divergence of the vector field $\mathbf{F} = F_r\mathbf{r} + F_\theta\theta + F_\phi\phi$ in spherical coordinates is:

$$\nabla \cdot F = \frac{1}{r^2\sin\theta}\frac{\partial}{\partial r}\left(\left(r^2\sin\theta F_r\right) + \frac{\partial}{\partial\theta}\left(r\sin\theta F_\theta\right) + \frac{\partial}{\partial\phi}\left(rF_\phi\right)\right) \tag{A.2.2}$$

A.2.3 Curl

The curl of the vector field above mentioned $\mathbf{F} = F_r\mathbf{r} + F_\theta\theta + F_\phi\phi$ in spherical coordinates is:

$$\begin{aligned}\nabla \times F = &\frac{1}{r\sin\theta}\left(\cos\theta F_\phi + \sin\theta\frac{\partial F_\phi}{\partial\theta} - \frac{\partial F_\theta}{\partial\phi}\right)r + \\ &+ \frac{1}{r\sin\theta}\left(\frac{\partial F_r}{\partial\phi} - \sin\theta F_\phi - r\sin\theta\frac{\partial F_\phi}{\partial r}\right)\theta + \\ &+ \frac{1}{r}F_\theta + r\frac{\partial F_\theta}{\partial r} - \frac{\partial F_r}{\partial\theta}\phi\end{aligned} \tag{A.2.3}$$

A.2.4 Laplace operator

The Laplace operator of the scalar field f in spherical coordinates is given by:

$$\Delta f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin\theta}\frac{\partial^2 f}{\partial\phi^2} \tag{A.2.4}$$

while the Laplace operator for the vector field $\mathbf{F}$ assumes the form:

$$\begin{aligned}\Delta F = \Delta F_r &- \left(\frac{2F_r}{r^2} - \frac{2}{r^2\sin\theta}\frac{\partial\left(\sin\theta F_\theta\right)}{\partial\theta} - \frac{2}{r^2\sin\theta}\frac{\partial F_\phi}{\partial\phi}\right)r + \\ &+ \left(\Delta F_\theta - \frac{F_\theta}{r^2\sin^2\theta} - \frac{2\cos\theta}{r^2\sin\theta}\frac{\partial F_\phi}{\partial\phi} + \frac{2}{r^2}\frac{\partial F_r}{\partial\theta}\right)\theta + \\ &+ \left(\Delta F_\phi - \frac{F_\phi}{r^2\sin^2\theta} - \frac{2\cos\theta}{r^2\sin\theta}\frac{\partial F_\theta}{\partial\phi} + \frac{2}{r^2\sin\theta}\frac{\partial F_r}{\partial\phi}\right)\phi\end{aligned} \tag{A.2.5}$$

Appendix B

Spherical Harmonics

Spherical harmonics (SH) are the angular portion of a set of solutions to Laplace's equation represented in a system of spherical coordinates. In this system they form an orthogonal basis. Spherical harmonics have many theoretical and practical applications, we can mention: the computation of atomic orbital electron configurations in quantum mechanics, the representation of gravitational and magnetic fields of planetary bodies and stars , and the characterization of the cosmic microwave background radiation in celestial mechanics. Here Spherical Harmonics are introduced from a theoretical perspective and then we show some attractive applications.

B.1 Definitions

B.1.1 Scalar spherical harmonics

Spherical harmonics satisfy the spherical harmonic differential equation given by the angular part of Laplace's equation written in spherical coordinates reported below. They are complex, continuous functions of θ and φ variables.

$$\Delta f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin\theta}\frac{\partial^2 f}{\partial\varphi^2} = 0 \qquad \text{(B.1.1)}$$

The combination of $\Phi(\varphi)$ and $\Theta(\theta)$ solutions obtained from the B.1.1 using separation of variables, scaled with a normalization factor defines the spherical harmonics:

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}\,P_l^{|m|}(\cos\theta)\,e^{im\varphi} \qquad \text{(B.1.2)}$$

where $P_l^{|m|}(\cos\theta)$ is an associated Legendre polynomial and the normalization is chosen such that:

$$\int_0^{2\pi}\int_0^{\pi} Y_l^m(\theta,\varphi)\bar{Y}_{l'}^{m'}(\theta,\varphi)\sin\theta d\theta d\varphi = \delta_{mm'}\delta_{ll'} \tag{B.1.3}$$

being $\bar{Y}_{l'}^{m'}(\theta,\varphi)$ the complex conjugative of $Y_{l'}^{m'}(\theta,\varphi)$.

SH have the following properties:

$$Y_l^0(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta) \tag{B.1.4}$$

$$Y_l^{-m}(\theta,\varphi) = (-1)^m\,\bar{Y}_l^{-m}(\theta,\varphi) \tag{B.1.5}$$

where $P_l(\cos\theta)$ is a Legendre polynomial.

B.1.2 Vector spherical harmonics

Vector spherical harmonics (VSH) are an extension of the scalar spherical harmonics for the use with vector fields. In scientific literature several conventions have been used to define the VSH. We follow that of Barrera et al.(1985)[27]. Being $Y_l^m(\theta,\varphi)$ a scalar spherical harmonic we define three VSH:

$$\mathbf{P}_{lm}(\theta,\varphi) = \hat{r}(\theta,\varphi)\,Y_{lm}(\theta,\varphi) \qquad l \geq 0 \tag{B.1.6}$$

$$\mathbf{B}_{lm}(\theta,\varphi) = \frac{1}{s_l} r\nabla Y_{lm}(\theta,\varphi), \qquad l > 0 \tag{B.1.7}$$

$$\mathbf{C}_{lm}(\theta,\varphi) = \frac{1}{s_l}\nabla\times[\vec{r}(\theta,\varphi)Y_{lm}(\theta,\varphi)], \qquad l > 0 \tag{B.1.8}$$

where $\hat{r}$ is the unitary vector along the radial direction, $\vec{r}$ represents the position vector of the point with spherical coordinate and $s_l = \sqrt{l(l+1)}$. The harmonics $\mathbf{B}_{00}$ and $\mathbf{C}_{00}$ are both equal to 0. The radial factors are included to guarantee that the dimensions of the VSH are the same as the ordinary spherical harmonics $Y_l^m(\theta,\varphi)$ and that the VSH are independent of the radial spherical coordinate. By the form of gradient and curl operator $\mathbf{B}_{lm}(\theta,\varphi)$ and $\mathbf{C}_{lm}(\theta,\varphi)$ spherical harmonics can be furthermore split in the the $\hat{\theta}$ and $\hat{\varphi}$ directions as shown below:

$$\mathbf{P}_{lm}(\theta,\varphi) = \hat{r}(\theta,\varphi)\,P_{lm}^{rr} \tag{B.1.9}$$

$$\mathbf{B}_{lm}(\theta,\varphi) = \hat{\theta}(r,\varphi)\,B_{lm}^{\theta\theta} + \hat{\varphi}(r,\theta)\,B_{lm}^{\theta\varphi} \tag{B.1.10}$$

$$\mathbf{C}_{lm}(\theta,\varphi) = \hat{\theta}(r,\varphi)\,C_{lm}^{\varphi\theta} + \hat{\varphi}(r,\theta)\,C_{lm}^{\varphi\varphi} \tag{B.1.11}$$

The vector spherical harmonics are orthonormal over the unit sphere:

$$\int_0^{2\pi}\int_0^{\pi} \mathbf{P}_{lm}(\theta,\varphi)\,\bar{P}_{l'm'}(\theta,\varphi)\sin\theta d\theta d\varphi = \delta_{ll'}\delta_{mm'} \tag{B.1.12}$$

$$\int_0^{2\pi}\int_0^{\pi} \mathbf{B}_{lm}(\theta,\varphi)\,\bar{B}_{l'm'}(\theta,\varphi)\sin\theta d\theta d\varphi = \delta_{ll'}\delta_{mm'} \tag{B.1.13}$$

$$\int_0^{2\pi}\int_0^{\pi} \mathbf{C}_{lm}(\theta,\varphi)\,\bar{C}_{l'm'}(\theta,\varphi)\sin\theta d\theta d\varphi = \delta_{ll'}\delta_{mm'} \tag{B.1.14}$$

$$\int_0^{2\pi}\int_0^{\pi} \mathbf{P}_{lm}(\theta,\varphi)\,\bar{B}_{l'm'}(\theta,\varphi)\sin\theta d\theta d\varphi = 0 \tag{B.1.15}$$

$$\int_0^{2\pi}\int_0^{\pi} \mathbf{B}_{lm}(\theta,\varphi)\,\bar{C}_{l'm'}(\theta,\varphi)\sin\theta d\theta d\varphi = 0 \tag{B.1.16}$$

$$\int_0^{2\pi}\int_0^{\pi} \mathbf{C}_{lm}(\theta,\varphi)\,\bar{P}_{l'm'}(\theta,\varphi)\sin\theta d\theta d\varphi = 0 \tag{B.1.17}$$

in which,as usual, being the overwritten bar denotes the complex conjugative functions. This vector basis is also orthogonal in the usual three-dimensional way:

$$\mathbf{P}_{lm}(\theta,\varphi)\cdot\mathbf{B}_{lm}(\theta,\varphi) = 0 \tag{B.1.18}$$
$$\mathbf{P}_{lm}(\theta,\varphi)\cdot\mathbf{C}_{lm}(\theta,\varphi) = 0 \tag{B.1.19}$$
$$\mathbf{B}_{lm}(\theta,\varphi)\cdot\mathbf{C}_{lm}(\theta,\varphi) = 0 \tag{B.1.20}$$

and this implies that:

$$\hat{r}(\theta,\varphi)\times\mathbf{C}_{lm}(\theta,\varphi) = \mathbf{B}_{lm}(\theta,\varphi) \tag{B.1.21}$$
$$\hat{r}(\theta,\varphi)\times\mathbf{B}_{lm}(\theta,\varphi) = -\mathbf{C}_{lm}(\theta,\varphi) \tag{B.1.22}$$

Like scalar spherical harmonics, VSH have the following symmetry property:

$$\mathbf{P}_{l-m}(\theta,\varphi) = (-1)^m\bar{P}_{lm}(\theta,\varphi) \tag{B.1.23}$$
$$\mathbf{B}_{l-m}(\theta,\varphi) = (-1)^m\bar{B}_{lm}(\theta,\varphi) \tag{B.1.24}$$
$$\mathbf{C}_{l-m}(\theta,\varphi) = (-1)^m\bar{C}_{lm}(\theta,\varphi) \tag{B.1.25}$$

B.2 Differential operators for functions expanded in SH

In differential problem characterized by a spherical geometry the expansion of the unknowns in scalar or vector spherical harmonics allows to transform the original partial differential problem in a set of ordinary differential equations for the coefficients of expansion, dependent only by the radial coordinate r. In order to obtain the representation of the problem in modal equations we need of explicit expression of the differential operator for functions expanded in spherical harmonics. Here the expression of differential operators for scalar and vector functions are presented:

B.2.1 Differential operators for scalar functions expanded in SH

Let $f(r,\theta,\varphi)$ be a scalar function expanded in scalar SH:

$$f(r,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} p_{lm}(r)\, Y_{lm}(\theta,\varphi) \tag{B.2.1}$$

Gradient

The gradient of the function $f(r,\theta,\phi)$ is given by:

$$\begin{aligned}\nabla f(r,\theta,\varphi) &= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\nabla f_{lm}(r)\, Y_{lm}(\theta,\varphi) + f_{lm}(r)\,\nabla Y_{lm}(\theta,\varphi)\right] \\ &= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\frac{df_{lm}(r)}{dr}\mathbf{P}_{lm}(\theta,\varphi) + f_{lm}(r)\,\frac{s_l}{r}\mathbf{B}_{lm}(\theta,\varphi)\right]\end{aligned} \tag{B.2.2}$$

which suggests the introduction of the vector differential operator:

$$\nabla_l = \begin{pmatrix} \frac{d}{dr} \\ \frac{s_l}{r} \\ 0 \end{pmatrix} \tag{B.2.3}$$

acting on the coefficient of B.2.1 expansion.

Laplace operator

The Laplace operator of the scalar field f is such as:

$$\nabla^2 f(r,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}-\frac{s_l^2}{r^2}\right) f_{lm}(r)\, Y_{lm}(\theta,\varphi)\right]$$
$$= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\left(D_l f_{lm}(r)\right) Y_{lm}(\theta,\varphi)\right] \tag{B.2.4}$$

having assumed:

$$D_l = \left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}-\frac{s_l^2}{r^2}\right) \tag{B.2.5}$$

B.2.2 Differential operators for vector functions expanded in VSH

Let $\mathbf{F}(r,\theta,\varphi)$ be a vector field expanded in scalar SH:

$$\mathbf{F}(r,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[F_{lm}^{r}(r)\,\mathbf{P}_{lm}(\theta,\varphi) + F_{lm}^{\theta}(r)\,\mathbf{B}_{lm}(\theta,\varphi) + F_{lm}^{\varphi}(r)\,\mathbf{C}_{lm}(\theta,\varphi)\right] \tag{B.2.6}$$

Divergence

The divergence of the vector field $\mathbf{F}$ is:

$$\nabla\cdot\mathbf{F}(r,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\left(\frac{d}{dr}+\frac{2}{r}\right)F_{lm}^{r}(r)+\frac{s_l}{r}F_{lm}^{\theta}(r)\right]Y_{lm}(\theta,\varphi)$$
$$= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\nabla_l\cdot\mathbf{F}_{lm}(r)\right]Y_{lm}(\theta,\varphi) \tag{B.2.7}$$

where:

$$\nabla_l = \begin{pmatrix} \left(\frac{d}{dr}+\frac{2}{r}\right) \\ -\frac{S_l}{r} \\ 0 \end{pmatrix} \tag{B.2.8}$$

Curl

The curl of the vector field above mentioned vector field is:

$$\begin{aligned}\nabla \times \mathbf{F}(r,\theta,\varphi) = &\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{s_l}{r}F_{lm}^{\varphi}(r)\,\mathbf{P}_{lm}(\theta,\varphi) + \\ &+\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(\frac{d}{dr}+\frac{1}{r}\right)F_{lm}^{\varphi}(r)\,\mathbf{B}_{lm}(\theta,\varphi) + \\ &+\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\frac{s_l}{r}F_{lm}^{r}(r)-\left(\frac{d}{dr}+\frac{1}{r}\right)F_{lm}^{\theta}(r)\right]\mathbf{C}_{lm}(\theta,\varphi) = \\ &=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}(R_l\mathbf{F}(r,\theta,\varphi))\mathbf{VSH}_{lm}(\theta,\varphi) \qquad \text{(B.2.9)}\end{aligned}$$

where:

$$R_l = \begin{pmatrix} 0 & 0 & \frac{s_l}{r^2} \\ 0 & 0 & \left(\frac{d}{dr}+\frac{1}{r}\right) \\ \frac{s_l}{r^2} & -\left(\frac{d}{dr}+\frac{1}{r}\right) & 0 \end{pmatrix} \qquad \text{(B.2.10)}$$

and the term $\mathbf{VSH}_{lm}$ indicates the vector spherical harmonics $(\mathbf{P}_{lm}, \mathbf{B}_{lm}, \mathbf{C}_{lm})$.

Laplace operator

Finally the Laplace operator for the vector field $\mathbf{F}$ assumes the form:

$$\begin{aligned}\nabla^2\mathbf{F}(r,\theta,\varphi) = &\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[\left(D_l-\frac{2}{r^2}\right)F_{lm}^{r}(r)+\frac{2s_l}{r^2}F_{lm}^{\theta}(r)\right]\mathbf{P}_{lm}(\theta,\varphi) \\ &+\left[\frac{2s_l}{r^2}F_{lm}^{r}(r)+D_lF_{lm}^{\theta}(r)\right]\mathbf{B}_{lm}(\theta,\varphi)+\left[D_lF_{lm}^{\varphi}(r)\right]\mathbf{C}_{lm}(\theta,\varphi) = \\ &=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}(\mathbf{D}_lF(r,\theta,\varphi))\mathbf{VSH}(\theta,\varphi) \qquad \text{(B.2.11)}\end{aligned}$$

in which:

$$\mathbf{D}_l = \begin{pmatrix} D_l-\frac{2}{r^2} & \frac{2s_l}{r^2} & 0 \\ \frac{2s_l}{r^2} & D_l & 0 \\ 0 & 0 & D_l \end{pmatrix} \qquad \text{(B.2.12)}$$

B.3 Applications of SH

Spherical Harmonics arise from a generalization of Fourier Series in one dimension higher. Consequently, many beautiful properties of Fourier Series are valid also for Spherical Harmonics. Like Fourier Series are a basis for the circle, Spherical Harmonics are a basis for the sphere. Both these functions have the property that eigenspaces are rotation-invariant, this results is a consequence of the Laplace operator (whose solutions generate these functions) invariance with the coordinates. The rotation invariance implies that if we know the expansion of a function, we can compute the expansion of any of its rotations by applying a sparse matrix. For spherical harmonics expansion the rotation matrix (called Wigner rotation matrices) consists of blocks along the diagonal, with a $(2n+1) \times (2n+1)$ block for each n. If we have a band width limitation $n < N$ then the matrix has to be applied $\sum (2n+1)^2 = \left(\frac{4N^3-N}{3}\right)$ times instead of $\left[\sum (2n+1)\right]^2 = N^4$ times. Moreover, if we work with operators with are linear and rotation-invariant Fourier Series and Spherical Harmonics diagonalize such operators making them trivial to apply. The wide use of these function in a large class of differential problems, involving round or spherical geometry, is justified by the fact that many operators corresponding to physical processes are coordinate invariant.

Appendix C

Chebyshev polynomials

This appendix resumes Chebyshev polynomials definitions, properties and applications, for futher informations the reader is referred to the Special Function section of Abramovitz and Stegun [25].

C.1 Definitions

Chebyshev polynomials are a set of orthogonal polynomials solution of the Chebyshev differential equations:

$$\left(1-x^2\right)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + n^2 y = 0 \tag{C.1.1}$$

in which n is a non-negative integer. Chebyshev Polynomials are usually denoted by $T_n(x)$.

They can be defines also by the trigonometric identity:

$$T_n(x) = \cos(n \arccos x) = \cosh(n \arccos hx) \tag{C.1.2}$$

whence:

$$T_n(\cos\theta) = \cos(n\theta) \tag{C.1.3}$$

having assumed $x = \cos\theta$.

The polynomials are either even or odd functions of x for even or odd orders n, that is to say:

$$T_{2n}(x) = T_{2n}(-x) \tag{C.1.4}$$
$$T_{2n+1}(x) = -T_{2n+1}(-x) \tag{C.1.5}$$

according to the invariance of the C.1.1 respect to the transformation which exchange x with $-x$. Chebyshev Polynomials are defined, continuous and limited in the range $[-1, 1]$:

$$\max_{x\in[-1,1]} |T_n(x)| = 1 \tag{C.1.6}$$

The first few polynomials are shown below:

$$\begin{aligned}
T_0(x) &= 1 & T_4(x) &= \left(8x^4 - 8x^2 + 1\right)\\
T_1(x) &= x & T_5(x) &= \left(16x^5 - 20x^3 + 5x\right)\\
T_2(x) &= \left(2x^2 - 1\right) & T_6(x) &= \left(32x^6 - 48x^4 + 18x^2 - 1\right)\\
T_3(x) &= \left(4x^3 - 3x\right) & T_7(x) &= \left(64x^7 - 112x^5 + 56x^3 - 7x\right)
\end{aligned}$$

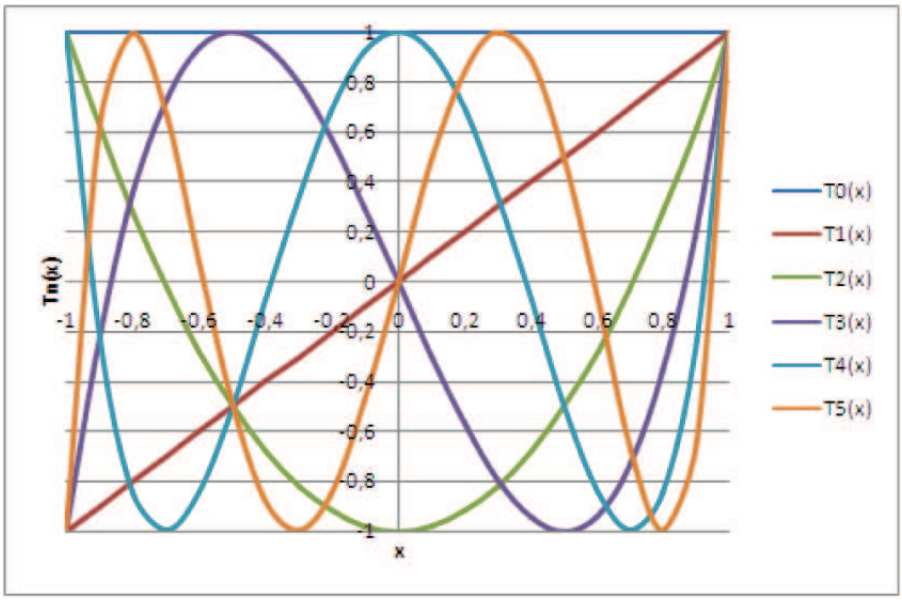

Figure C.1: Graphic of the first Chebyshev Polynomials in the range [-1,1]

Notice that on the interval $[-1, 1]$ all of the extrema have values that are either -1 or 1.

C.1.1 Recurrence formula

A convenient tool to define Chebyshev polynomials is the recurrence relation:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \tag{C.1.7}$$

C.1.2 Properties

Chebyshev Polynomials are orthogonal functions in $[-1, 1]$ with respect to the weight $\frac{1}{\sqrt{1-x^2}}$:

$$\int_{-1}^{1} T_m(x)T_n(x)\frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0 & n \neq m \\ \pi & n = m = 0 \\ \pi/2 & n = m \neq 0 \end{cases} \tag{C.1.8}$$

A Chebyshev polynomial with degree n has n different simple roots, also called Chebyshev nodes, in the interval $[-1, 1]$. The roots of $T_n(x)$ are:

$$x_i = \cos\left(\frac{\pi}{2}\,\frac{2i-1}{n}\right), \quad i = 1, \ldots, n. \tag{C.1.9}$$

C.1.3 Applications

Chebyshev polynomials are important in approximation theory because their roots, also called Chebyshev nodes, are the best nodes in polynomial interpolation (see Chapter). In fact, the resulting interpolation polynomial minimizes the problem of Runge's phenomenon and provides an approximation that is close to the polynomial of best approximation to a continuous function under the maximum norm. Chebyshev polynomials form a complete orthogonal basis, so that a function defined on the interval $[-1, 1]$ can be expanded in Chebyshev series:

$$f(x) = \sum_{i=0}^{\infty} f_i T_i(x) \tag{C.1.10}$$

whose partial sum:

$$f(x) = \sum_{i=0}^{N} f_i T_i(x) \tag{C.1.11}$$

is very useful in the approximation of functions and in solution of differential equations with spectral methods. The coefficient f_i of D.1.13 expansion can be determined through the application of an inner product thank to the Chebyshev polynomials orthogonality. Since a Chebyshev series is related to a Fourier cosine series through a change of variables, all of the theorems, identities, etc that apply to Fourier series have a Chebyshev counterpart.

Appendix D

Legendre functions

This appendix is dedicated to the Legendre functions presentation. Definitions, properties and applications are briefly discussed, for a wide treatment the reader is referred to the Special Function section of Abramovitz and Stegun [25].

D.1 Legendre Polynomials

D.1.1 Definitions

The Legendre polynomials, sometimes called Legendre functions of the first kind, Legendre coefficients, or zonal harmonics, are solutions to the Legendre differential equation:

$$\left(1-x^2\right)\frac{d^2f}{dx^2} - 2x\frac{df}{dx} + n\left(n+1\right)f = 0 \tag{D.1.1}$$

When n is a real number the solutions of D.1.1 are called Legendre Functions of degree n while if n is a non-negative integer, the Legendre Functions are often referred to as Legendre Polynomials. In this last case, the polynomials may be denoted by $P_n(x)$.

Legendre polynomials are also generated by the Rodriguez formula:

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left(x^2-1\right)^n \tag{D.1.2}$$

or by the contour integral:

$$P_n(x) = \frac{1}{2\pi i}\oint \frac{1}{\sqrt{1-2tx+t^2}}t^{-n-1}dt \tag{D.1.3}$$

where the contour encloses the origin and is traversed in a counterclockwise direction. The polynomials are either even or odd functions of x for even or odd orders n, that is to say:

$$P_{2n}(x) = P_{2n}(-x) \tag{D.1.4}$$
$$P_{2n+1}(x) = -P_{2n+1}(-x) \tag{D.1.5}$$

according to the invariance of the D.1.1 respect to the transformation which exchange x with $-x$. The first few polynomials are shown below:

$$P_0(x) = 1 \qquad P_4(x) = \frac{1}{8}\left(35x^4 - 30x^2 + 3\right)$$
$$P_1(x) = x \qquad P_5(x) = \frac{1}{8}\left(63x^5 - 70x^3 + 15x\right)$$
$$P_2(x) = \frac{1}{2}\left(3x^2 - 1\right) \qquad P_6(x) = \frac{1}{16}\left(231x^6 - 315x^4 + 105x^2 - 5\right)$$
$$P_3(x) = \frac{1}{2}\left(5x^3 - 3x\right) \qquad P_7(x) = \frac{1}{16}\left(429x^7 - 693x^5 + 315x^3 - 35x\right)$$

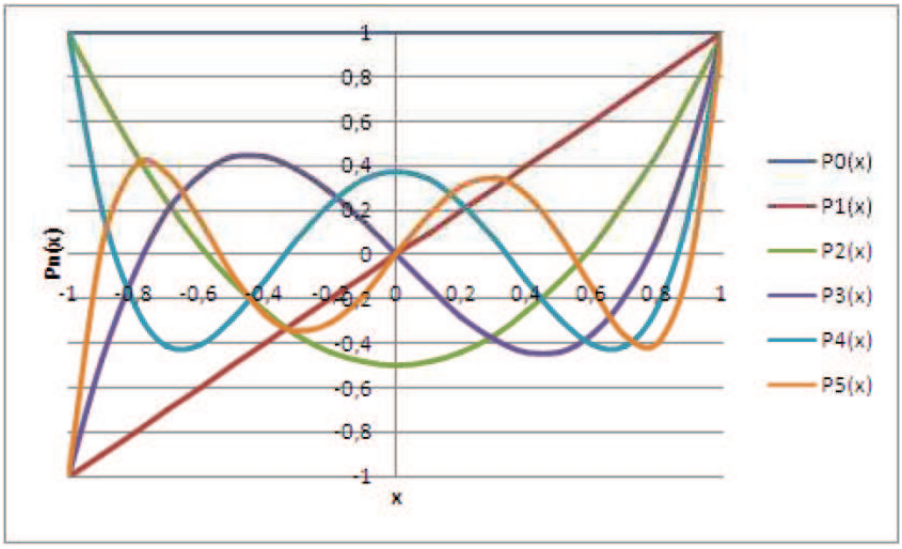

Figure D.1: Graphic of the first Legendre Polynomials in the range [-1,1]

D.1.2 Recurrence formulas

Legendre polynomials and their derivatives agree with the following recursive formulas.

$$(n+1)\,P_{n+1}(x) = (2n+1)\,xP_n(x) - nP_{n-1}(x) \tag{D.1.6}$$

$$P'_{n+1}(x) - xP'_n(x) = (n+1)\,P_n(x) \tag{D.1.7}$$

$$xP'_n(x) - P'_{n-1}(x) = nP_n(x) \tag{D.1.8}$$

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)\,P_n(x) \tag{D.1.9}$$

$$\left(x^2-1\right)P'_n(x) = nxP_n(x) - nP_{n-1}(x) \tag{D.1.10}$$

The first one is an easy tool, useful in numerical computation, to obtain Legendre polynomials starting from $P_0(x)$ and $P_1(x)$. The others allow to compute Legendre polynomials derivatives by means of known derivatives or polynomials.

D.1.3 Properties

Legendre Polynomials are defined and continuous over the whole real axis, but restricting our attention on the range $[-1,1]$ Legendre Polynomials are orthogonal functions.

$$\int_{-1}^{1} P_l(x)P_n(x)dx = 0 \qquad m \neq n \tag{D.1.11}$$

$$\int_{-1}^{1} \left(P_n(x)\right)^2 dx = \frac{2}{2n+1} \tag{D.1.12}$$

Moreover, in the range $[-1,1]$ Legendre Polynomials $P_n(x)$ have n distinct roots, all internal to the interval. In this context the variable x is often replaced by $cos\theta$ which automatically scales the domain to this interval.

D.1.4 Applications

Legendre polynomials form a complete orthogonal basis, so that a function defined on the interval $[-1,1]$ can be expanded in Legendre series:

$$f(x) = \sum_{i=0}^{\infty} f_i P_i(x) \tag{D.1.13}$$

whose coefficient f_i can be determined through the application of an inner product thank to the Legendre polynomials orthogonality.

D.2 Associated Legendre functions

From the Legendre polynomials can be generated another important class of functions for physical problems, the associated Legendre functions. These functions are of great importance in quantum physics because they appear in the solutions of the Schrodinger equation in spherical polar coordinates. Moreover, the associated Legendre functions can be used to construct another important set of functions, the spherical harmonics, which are used in many physical problem involving spherical symmmetry[1].

D.2.1 Definitions

The associated Legendre functions are solutions of the associated Legendre differential equation:

$$\left(1-x^2\right)\frac{d^2f}{dx^2}-2x\frac{df}{dx}+\left(n\left(n+1\right)-\frac{m^2}{1-x^2}\right)f=0 \tag{D.2.1}$$

Here we restrict our attention to the cases in which n and m are non negative integers. The associated Legendre functions may be denoted by $P_n^m(x)$.

Associated Legendre functions can be derived by the Rodriguez formula:

$$P_n^m(x)=\frac{\left(1-x^2\right)^{m/2}}{2^n n!}\frac{d^{m+n}}{dx^{m+n}}\left(x^2-1\right)^n=\left(1-x^2\right)^{m/2}\frac{d^m}{dx^m}P_n(x) \tag{D.2.2}$$

Notice that:

$$P_n^0(x)=P_n(x) \tag{D.2.3}$$

$$P_n^m(x)=0 \quad if\, m>n \tag{D.2.4}$$

The first few polynomials are shown below:

$$\begin{aligned} P_1^1(x)&=(1-x^2)^{1/2} & P_3^1(x)&=\frac{3}{2}(5x^2-1)(1-x^2)^{1/2}\\ P_2^1(x)&=3x(1-x^2)^{1/2} & P_3^2(x)&=15x(1-x^2)\\ P_2^2(x)&=3(1-x^2) & P_3^3(x)&=15(1-x^2)^{3/2} \end{aligned} \tag{D.2.5}$$

D.2.2 Recursive formulas

Associated Legendre functions respect the following recursive formulas.

$$\left(n+1-m\right)P_{n+1}^m(x)=\left(2n+1\right)xP_n^m(x)-(n+m)P_{n-1}^m(x) \tag{D.2.6}$$

$$P_n^{m+2}(x)=\frac{2(m+1)x}{(1-x^2)^{1/2}}P_n^{m+1}(x)-(n-m)(n+m+1)P_n^m(x) \tag{D.2.7}$$

[1]Spherical Harmonics theory, properties and applications are given in appendixB

D.2.3 Properties

Associated Legendre functions are orthogonal in the range $[-1, 1]$:

$$\int_{-1}^{1} P_l^m(x) P_n^m(x) dx = 0 \qquad m \neq n \tag{D.2.8}$$

$$\int_{-1}^{1} \left(P_n^m(x)\right)^2 dx = \frac{2}{2n+1} \frac{(n+m)!}{(n-m)!} \tag{D.2.9}$$

Bibliography

[1] R.J. Adrian, K.T.Cristensen, Z.C. Liu *Analysis and interpretation of instantaneous turbolent velocity fields*. Experiments in fluids, **29** (2000), 275-290.

[2] J.C.Agui, J.Jimenez *On the performance of particle tracking*. J. of Fluid Mech. **185** (1987), 447-468

[3] W.Becker, *Metrics, The neurobiology of saccadic eye movements*. Eds. Wurtz e Goldberg. Elsevier Science Publisher BV (Biomedical Division), 1989.

[4] C.R.Canning, J.N.Dewynne, A.D.Fitt and M.J.Greaney, *Fluid ow in the anterior chamber of a human eye*, J. Math. Appl. Med. Bio. **2** no 5 (1990), 31-60.

[5] M.S.Chong and A.E.Perry, *A general classification of three dimensional flow fields*. Phys. in fluids, **12** (2001), 765-777.

[6] U.Cicekli, *Computational model for heat transfer in the human eye using the finite element method*. PhD. Thesys, Luisiana State University.

[7] T.David, S.Smye e T.James, *A model for the fluid motion of vitreous humour of the human eye during saccadic movements*. Phys. Med. Biol., **43** (1998),1385-1399.

[8] V.M.M.Flyckt, B.W.Raaymakers and J.J.W.Lagendijk, *Modelling the impact of blood ow on the temperature distribution in the human eye and in the orbit: fixed heat transfer coefficients versus the Pennes bioheat model versus discrete blood vessels*, Phys. Med. Biol. **51** (2006), 5007-5021

[9] B.L.Hua, P.Klein, *An exact criterion for the stirring properties of nearly two-dimensional turbolence*. Physica, **113** (1998), 98-110.

[10] J.Jeong, F.Houssain, *On the identification of a vortex*. Journal of Fluid Mechanics, **285** (1995), 69-94.

[11] L.Graftieaux, M.Michard and N. Grosjan, *Combining PIV, POD and vortex identification algorithms for the study of unsteady turbulent swirling flows*. Meas. Sci. Technol., **12** (2001), 1422-1429.

[12] B.Lee, M.Litt e G.Buchsbaum, *Rheology of the vitreous body. Part I: Viscoelasticity of human vitreous*. Biorheology **29** (1992),521-533.

[13] K.Lindner, *Über die Herstellung von Modellen zu Modellversuchen der Netzhautabhebung*. Klin. Mbl. Augenheilk, **90** (1933),289-300.

[14] H.Lund-Andersen, *Adler's physiology of the eye*. Tenth edition Mosby, (2003).

[15] RG. Miller, *Biometrika***61** (2006),1-17. (2003).

[16] J.M.Ottino, *The kinematics of mixing: stretching, chaos, and transport* (Cambridge Texts in Applied Mathematics) Cambridge University Press (1989).

[17] M.Raffel, C.Willert, J. Kompenhans *Particle Image Velocimetry: a practical guide*. Eds. Springer.

[18] R.Repetto, *An Analytical Model of the Dynamics of the Liquefied Vitreous Induced by Saccadic Eye Movements* . Meccanica **41** (2006),101-117.

[19] R.Repetto, I.Ghigo, G.Seminara e C.Ciurlo, *A simple hydro-elastic model of the dynamics of a vitreous membrane*. J.Fluid Mech. **503** (2004),1-14.

[20] R.Repetto, A.Stocchino and C.Cafferata, *Experimental investigation of vitreous humour motion within a human eye model* . Physics in Medicine and Biology **50** (2005),4729-4743.

[21] B.Rosengren e S.Östrelin, *Hydrodynamic events in vitreous space accompanying eye movements*. Ophtalmologica **173** (1976),513-524.

[22] J.A. Scott, *A finite element model of heat transport in the human eye*. Physics in Medicine and Biology **33**,n°2 (1988),227-242.

[23] A.Stocchino, R.Repetto and C.Cafferata, *Eye rotation induced dynamics of a Newtonian fluid within the vitreous cavity: the effect of the chamber shape* . Physics in Medicine and Biology **52** (2007),2021-2034.

[24] H.Stuer, S.Blaser *Interpolation of scattered 3D PTV data, to a regular grid*. Flow Turbolence Combut.**64** (2000), 215-232

[25] M. Abramowitz, I. Stegun, *Handbook of mathematical functions*. Dover Pubblications, INC,New York.

[26] K. Atkinson, *Numerical integration on the sphere*. Journal of Australian Mathematical Society **23**, Series B (1982), 332-347.

[27] R.G. Barrera, G.A. Estévez and J. Giraldo, *Vector spherical harmonics and their application to magnetostatic* Eur. J. Phys. **6**(1985), 287-294

[28] J.P. Boyd, *Chebyshev and Fourier Spectral Methods* DOVER Publications, Inc., 2nd edition (2000), p 595

[29] C. Canuto, M.Y. Hussaini, A. Quarteroni and T.A. Zang, *Spectral methods in fluid Dynamics* Springer-Verlag (1988), 339-367.

[30] L. Fox, I.B. Parker, *Chebyshev Polynomials in numerical Analysis*. Oxford University Press.

[31] L. Gatteschi, *Funzioni speciali*. UTET.

[32] L. Gatteschi, *Lezioni di Analisi numerica*. Libreria Editrice Universitaria Levrotto e Bella, Torino.

[33] M. Graf, S. Kunis, D. Potts, *On the computation of nonnegative quadrature weights on the sphere*. Appliede and Computational Harmonic Analysis **27** (2009), 124-132.

[34] M.Y. Hussaini, T.A. Zang, *Spectral methods in fluid Dynamics* Annual Review of Fluid Mechanics **19** (1987), 2nd printing pp.567

[35] J. Keiner, D. Potts, *Fast evaluation of quadrature formulae on the sphere*. Mathematics of computation **77**, no 261, (2008), 397–419.

[36] S. Kunis, D. Potts, *Fast spherical Fourier alghorithms*. Journal of Computational and Applied Mathematics **161** (2003), 75–98.

[37] M. J. Mohlenkamp, *A fast Transform for Spherical Harmonics*. The Journal of Fourier Analysis and Applications **5** 2/3 (1999), 159-184.

[38] M. J. Mohlenkamp, *A User's Guide to Spherical Harmonics*. `http://www.math.ohiou.edu/~mjm/research`

[39] AA.VV, *NAG Fortran Library Manual*. http://www.nag.co.uk/numeric/F1/manual/xhtml.

[40] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes*. Cambridge University Press.

[41] L.Quartapelle, M. Verri, *On spectral solution of the three-dimensional Navier-Stokes equations in spherical and cylindrical regions*. Computer physics **90** (1995), no1, 1-43 (33 ref.).Elsevier Science, Amsterdam, PAYS-BAS (1969) (Revue)

[42] L.Quartapelle and M.Napolitano, *Integral conditions for the pressure in the computation of incompressible viscous flow*. J. of Computational Physics **62** (1986), 340-348.

[43] L.Quartapelle, *Numerical solution of the incompressible Navier-Stokes equations*. Birkhäuser, Basel, (1993).

[44] A. Quarteroni, *Modellistica numerica per problemi differenziali*. 3rd edition, (2008) Springer.

[45] A. Quarteroni, R. Sacco, F. Saleri, *Matematica numerica*. 2nd edition, Springer.

[46] M.R. Spiegel, *Manuale di Matematica*. Collana Schaum, ETAS libri.

[47] J.C. Tannehill, D.A. Anderson, R.H. Pletcher, *Computational Fluid Mechanics and Heat Tranfer*. 2nd edition,Taylor & Francis.

[48] Mark Wieczorek, Archive of Fortran 95 based software. `http://www.jpgp.fr/~wieczor/SHTOOLS/SHTOOLS.html`