1. INTRODUCTION

"At the heart of Formula One lies a very simple engineering concept: To send a racing car over a pre-determined course in the shortest possible time. That's it, nothing more and nothing less. Ironically, this is never really achieved. Technology aside, it astonishes me how fallible a human being is, people make mistakes all the time and to me, that is how races are won most of the time. You win, if the deficiencies in your team are less than the shortcomings of your competitor." - Hans Fouché, 1995, the best engineer I know and a very good friend.

Hans goes on: "F1 is not about magic, it is about understanding the principles at play, harnessing them and using them to your advantage .... It's about the basics, if you don't understand them, you'll get nowhere. You go into a wind tunnel with a bright idea and come out with more questions than answers. There is so much that we don't understand yet."

It is against this background that a study into the principles of the aerodynamic forces at play while racing, was launched. It is hoped that by examining the flow beneath a Formula One car, a small contribution can be made to this very fascinating and exciting engineering industry.

Since the discovery of the beneficial effects of under-floor flow in Formula One by Colin Chapman[1] in 1975, race car designers have continuously sought to improve traction by optimising the flow under the floor. After the dangers of ground effects became apparent, the governing body laid down strict rules to limit traction. Despite these constraints and the ones laid down in the years to follow, designers still managed to increase grip and decrease lap times on a regular basis. The 1995 season saw the introduction of a set of rules which dramatically altered the shape of the car floor. A stepped-bottom floor appeared for the first time. As the majority of the floor now had to be at least 50 mm above the ground, it was felt that the experience and knowledge gained on smooth-bottomed floors might no longer be relevant. The interaction between the floor and the rest of the car therefore had to be re-established. Consequently, a study into the flow beneath a Formula One car was initiated. This investigation, which was conducted using experimental and computational techniques, had the following objectives:

- To verify and determine the centre-line pressure profile by using experimental techniques.
- To use experimental methods to determine whether the characteristic flow patterns beneath the floor were functions of the floor profile alone, or whether other components also played a major role.
- To reproduce the flow patterns numerically by using the isolated centre-line geometry of the floor.

The study was based on the Parmalat Forti Ford FG01-95 Formula One racing car. This car was first raced during the 1995 Formula One season.

The main aim of this document is therefore to present the findings of this research project as outlined above. A secondary and perhaps a more subtle goal can be found in the structure and contents of the literature study. After this chapter, a person with very little or no knowledge of aerodynamics should be in a position to understand fundamental automotive aerodynamics and most of the findings presented in later chapters.

The history of gasoline-fuelled motor racing and the major aerodynamic advances that have been made since the inception of the sport at the beginning of the 20th century, are briefly presented at the beginning of the literature study. From this it can be seen that the application of wings in Formula One more than thirty years ago and the discovery of ground effects approximately a decade later, have lead to the very distinctive aerodynamic shape of the modern open-wheel racing car. The evolution of the modern racing car is used as a basis for the introduction of general automotive aerodynamic terms and concepts. The more traditional but still highly reliable experimental methods are discussed in Chapter 2 before the use of computational methods are examined.

The discussion on experimental techniques is introduced by examining the basic outlay of the different types of wind tunnels found in the industry today. As wind tunnels are used to measure aerodynamic quantities, the definition and use of these terms are presented in conjunction with a discussion on aerodynamic measuring equipment. This section is concluded by looking at wind-tunnel testing techniques. Here, issues such as the effect of scale, tunnel blockage and model-mounting techniques are outlined.

The literature study is concluded with a more technical look at the Computational Fluid Dynamics techniques which are currently used to solve fluid-flow problems. After presenting the governing Navier-Stokes equations in their integral and differential forms, the principles of turbulent flows are outlined. This is followed by a more detailed discussion of the turbulent models which are currently in use. The next topic under discussion, grid generation, outlines the various techniques which are used to construct the domain over which the flow field is to be solved. This includes the general transformation from the irregularly-spaced physical domain to a uniformly-spaced computational domain. The remaining two sections of Chapter 2 are aimed at discussing the numerical-solution methods which are employed to solve the governing equations and the boundary conditions which are required to define the numerical problem.

The investigation of the flow beneath the Formula One car, as presented in Chapter 3, was conducted in two parts. Firstly, a 30% scale model of the Parmalat Forti Ford FG01-95 was manufactured and tested in the low-speed rolling-road wind tunnel at the Council for Scientific and Industrial Research, South Africa. It was demonstrated that the centre-line pressure profile was representative of this type of vehicle. A number of changes were then made to the model in order to determine which factors, external to the floor, played a fundamental part in producing the pressure distribution beneath the car. Measurements were in the form of surface-static pressures. It was found that although parameters such as ride-height and front-flap position affected the intensity of the pressure map, the fundamental shape remained unaltered. The wheels were found to be instrumental in inducing significant changes to the pressure distribution. Interestingly, the front or rear wheel sets were independently sufficient to establish these flow changes. Adding the other set only produced milder and more local changes. Although the net effect of the wheels on the floor were negative, areas were found where the wheels had a beneficial influence.

As stated previously, the computational part of the investigation was aimed at determining the success of using isolated centre-line floor cases to predict experimental results. The computational domain was constructed using an efficient algebraic grid generator. Two numerical codes were used for the simulations. The first was an in house version of Roe's Flux Splitting Difference algorithm[2] and the second was the commercial code STAR-CD of

Computational Dynamics[3], England. The boundary conditions were controlled using experimental data. It was found, that despite the experimentally-fixed boundaries, the absence of external components such as the rear-wing assembly made predictions difficult. Changes to the centre-line grid, did however, produce the correct trends. The use of Computational Fluid Dynamics in Reynolds-Number sensitivity calculations is also demonstrated.

To address the problem of the poor numerical correlation under the floor of the Forti, a second study was launched. This is the subject matter of Chapter 4. Here, a $1/12^{th}$ scale generic model of a BMW Touring Car was used to investigate the accuracy with which the centre-line profile of the complete vehicle could be simulated. The F1 model was not chosen for this investigation as the smaller-scale closed-wheel racing car offered a more efficient, more economical and less sensitive method of obtaining the results.

The experimental results required for numerical correlation were obtained by manufacturing a scale model of the BMW, instrumenting it with surface-static pressure tubes, and testing it in a wind tunnel. The low-speed scale-model wind tunnel at the University of Pretoria in South Africa was used for the tests. The tunnel was equipped with a stationary-road raised-platform test-section floor. The objectives of the experimental investigation were:

- To establish similarities between the BMW and the Forti
- To obtain a base-line set of results for use in the numerical investigation
- To demonstrate the effect of a rear-mounted aerodynamic device

After repeatability and Reynolds-Number independence are demonstrated, it is shown that the pressure distribution beneath the BMW floor exhibits the same characteristics as those found under the F1 floor. Continuity is therefore established. By super-imposing the test results of a rear-mounted spoiler over the base-line data of the BMW, the effect of this device becomes clearly visible. The related discussion outlines the underlying phenomena responsible for rear downforce on sedan-based racing cars.

For the numerical investigation as presented in section 4.2, the centre-line profile of the experimental model was duplicated using the structured pre-processor of STAR-CD. The numerical investigation was centred around the accuracy of simulating turbulence by using the three high Reynolds Number turbulence models as supplied by STAR-CD.

The results showed that although the inclusion of the complete centre-line profile in the simulations increased the accuracy of the solution to a degree, the correlation of the predicted pressure distribution under the floor was still poor. It was also found that the three turbulence models out-performed one another in different areas around the vehicle. The poor numerical performance under the floor, the uncertainties regarding the choice of turbulence model and supporting experimental evidence lead to the conclusion that the numerical simulation was not accurate enough. It was therefore suggested that a further study should be launched but this new investigation went beyond the scope of this script. This new study should be aimed at improving the accuracy of the computational simulations of the BMW. A proposed direction for the new study is outlined in section 4.2.3. In this section it is shown by additional simulations that an increase in grid density is likely to improve the solution. The effect of varying the turbulence parameters is also shown to be a plausible method of improving correlation. Finally, it is concluded that if the standard STAR-CD turbulence models are unsuitable for the flow around the BMW, other turbulence models, which are designed to capture turbulence under adverse pressure gradients, should be used.

In Chapter 5 a conclusion is presented where the methodology and findings of this study are summarised.

Chapter 6 contains a list of all the references that were consulted during the course of this study.

Two appendices are also included. Appendix A presents the algorithm for the implementation of Menter's Shear Stress Transport turbulence model[4] and in Appendix B, Roe's Flux Difference Splitting procedure[2] for the full Reynolds-Averaged Navier-Stokes equations is outlined.

2. LITERATURE STUDY

A brief historical outline of automobile racing and the evolution of race car aerodynamics serves as a general introduction to this chapter. This is followed by an experimental literature study covering topics such as, wind tunnels and wind-tunnel models, test techniques and related equipment. Finally, numerical aerodynamics or Computational Fluid Dynamics (CFD), is presented and includes grid generation methods, flow solvers, boundary conditions and turbulence models.

2.1 Race Car Aerodynamics

2.1.1 The History of Automobile Racing and Race Car Aerodynamics

Automobile racing, using gasoline-fuelled internal combustion engines began in 1894, as a reliability race from Paris to Rouen, France. The race covered a distance of 80 km with the winner maintaining an average speed of just 16.4 kph[5]. Motor racing and manufacturers grew rapidly and within three years racing cars were looking distinctly different from road cars. The first regulation, which was introduced in 1898 after contestant fatalities, simply divided the cars into lightweight and heavyweight classes.

The first annual trophy, the James Gordon Bennet Cup, was launched in 1901 and national automobile clubs could each enter three cars. This cup transformed racing from a gentleman's sport to the more modern philosophy of "fastest wins"[1]. After spectator fatalities in 1901 and 1903, the French government banned street racing, unless events were held in sparsely populated areas and on roads with barriers.

The French manufacturers boycotted the Bennet trophy in 1906, as they were not content with only three cars per country. This lead to the first Grand Prix to be held at Le Mans in the same year. Racing events were now contested between rival manufacturers and not between national teams. The French ban on street racing lead to the first purpose built track at Brooklands, England in 1907.

From 1911, after the recession and until the beginning of the First World War, motor racing in Europe underwent significant changes. The evolution of racing circuits from basic triangular or oval shapes, to more twisting and winding circuits, led to the changes in design philosophy. Racing cars were becoming less dependent on engine power and more dependent on manoeuvrability, better brakes and more flexible engines.

The first limitation on the body-work was introduced in 1912, when the maximum width of the cars was limited to 175 cm's. For the 1913 season, the first rule had surfaced which limited the use of aerodynamic devices. Streamlined tails behind the fuel tank were not permitted[1]. By 1914, racing cars were moulded into a form which remained almost unchanged for the next 40 years.

The rapid progress in engineering during the First World War, served as a basis for a period of technological advancement in European motor sport. These advances led, amongst other things, to the 1923 Fiat 805.405, which was the first Grand Prix car to be developed in a wind tunnel.

The American designers were less eager to adopt the new found knowledge and persisted in emphasising engine power as the key to success.

In 1947, *the Fédération Internationale de l'Automobile* (F.I.A.)[i] divided Grand Prix racing into two formulae: the F.I.1 and the F.I.2. Formula One and its subsidiaries were born. During the same year the National Association for Stock Car Auto Racing (NASCAR) was founded in the U.S.A. In 1950, a Formula One world championship for drivers was started, followed by a championship for constructors in 1955. Hot Rod and particularly Drag Racing, which originated in California in the 1930's, was recognised by the F.I.A. in 1965. Racing with midget cars, which lead to cart racing, began in the 1940's in the USA and expanded to several international events from the 1960's onwards.

Until the 1950's, engine, chassis and tyre technologies were still changing rapidly, but aerodynamic improvements were largely aimed at drag reductions. By 1961 the transition from front mounted, to rear-mounted engines in Formula One was complete. Cars were now becoming smaller, lighter and more streamlined. The Coopers (father and son)[1], started the move towards rear-mounted engines, but it was the Lotus designer Colin Chapman who capitalised on it. He was, in 1959, in a quest for the smallest frontal area, the first to change the driver's seat from an upright to a more reclined position.

The smaller, more streamlined rear-mounted engine cars were less readily accepted in the USA. It took the second place of the comparatively small Lotus of Jim Clark at the Indianapolis in 1963 and the win in 1965, where he increased the average lap speed by 10 mph[6], to convert the Americans.

One of the American pioneers in the rear-engine racing car design arena was Jim Hall, joint owner of Chaparral. He joined Lotus so that he could study Grand Prix racing at first hand. He returned to the USA in 1963 and started designing rear-engine cars for the Canadian-American (CanAm) racing series. His designs carried two notable aerodynamic innovations; the front end spoiler beneath the nose and the first successful inverted wing above the rear axle. The wing was first seen on the 1966 Chaparral 2E. This idea was not copied effectively by Grand Prix constructors until almost two years later. This was partially due to the introduction of a new racing formula in 1966. The new formula retained, amongst other rules, the 1961 rule which banned body-work around the wheels. This rule was becoming more significant, as the ever increasing tyre widths were affecting the aerodynamic performance of the cars.

From 1968 the use of wings in F1 had spread almost uncontrollably. The wings had variable angles of attack, numerous shapes and sizes and were attached to all areas of the body-work and suspension. The poor construction and attachment methods frequently caused the wings to break off. This lead to serious accidents and subsequently, in 1969, the F.I.A. laid down strict rules governing the use of wings. The front wings were limited to nasal fins, while the rear wings were confined to a small area fixed to the sprung mass of the car.

---

[i] The *Alliance Internationale des Automobiles Clubs Reconnus*, which was formed in 1904, was renamed *the Fédération Internationale de l'Automobile* in 1946.

The next aerodynamic revolution came to motor racing in the form of the 1978 Lotus, designed by Colin Chapman. He designed the under-tray of the car in the shape of two half venturis, one on either side of the cockpit. The under-tray of the car formed the upper surface of the venturi, while the lower side was made up by the road surface. The venturis, which generated large amounts of negative pressure, especially when sealed using moveable side skirts, increased cornering speeds tremendously. The dangers involved in the high speeds lead to the banning of side skirts in 1981. Subsequent rule changes over the years, have lead to significant reductions in the size of the venturis in Formula One. Where the rules permit, large under-body diffusers are still used in other forms of motor racing.

There have been many other aerodynamic innovations, but none have been so radical and so important as the use of wings and ground effects.

The governing bodies are continuously revising the aerodynamic regulations, in an attempt to make motor sport safer.

2.1.2 Aerodynamics of Modern Racing Cars

Figure 2-1 presents a Formula One racing car, indicating the most important aerodynamic devices found on a racing car.

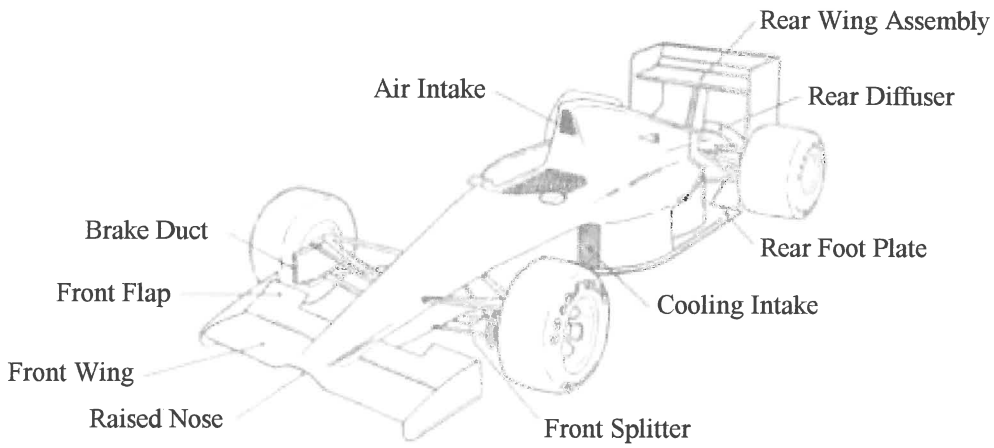

Figure 2-1 - Prominent Aerodynamic features of a 1992 Formula One Racing Car[7]

It can be seen from Figure 2-1, that race car aerodynamics can be divided into two classes. Internal aerodynamics, covering areas such as cooling ducts, heat exchangers, engine intakes and brake ducts and external areas such as wings, diffusers and body panels. To optimise these devices the aerodynamicist can use one of two techniques[8]:

- Experimental and
- Numerical or computational methods.

Traditionally, experimental methods were the only viable tool available to improve the aerodynamic performance of racing cars. With the advent of the high-speed digital computer, the aerodynamicist now has a relatively new tool on which he can rely. Potentially, CFD can be a very powerful and cost-effective tool to optimise the performance of cars. CFD has been used by some of the top Formula One teams in their development programmes[9,10]. Although the use of numerical flow solvers is becoming a more common practice, experimental aerodynamics is still regarded as the more reliable method for external flow problems.

2.2 Experimental Literature Study

Experimental aerodynamics can broadly be divided into two categories: road or track and wind-tunnel testing. Although both have advantages, wind tunnels are usually favoured when extensive development programmes are undertaken. This is largely due to the obtainable level of fluid control and the objective nature of tunnel tests. Furthermore, measuring equipment can be too bulky or too sensitive for track usage and can complicate experimental modifications. Track testing is normally used for verification, correlation and final "tuning", where driver preferences and track conditions become significant.

2.2.1 Wind Tunnels

Low-speed wind tunnels (LSWT) are commonly used for road and racing car development programmes. These tunnels are generally designed to operate at speeds below 300 mph (482 kph), where fluid compressibility is negligible[11]. This is the domain, in which most racing cars, except for certain types of drag racers and land-speed record vehicles, have to perform.

2.2.1.1 Low-Speed Wind tunnel Types

LSWT, where air is used as the test fluid, can be divided into two categories:

- Open-return tunnels and
- Closed-return tunnels

Open-return tunnels have their inlets and outlets venting to the atmosphere. These are usually cheaper to construct and are also preferable when engine development work or smoke generators are used frequently. They are more prone to noise pollution and testing can be dependent upon weather conditions if the tunnel is not enclosed in a larger building,.

In closed-return tunnels the air is circulated in a continuous loop. These are more expensive to build but are cheaper to operate. Although less dependent on weather fluctuations, they can be subject to rising air temperatures during prolonged tests. This problem can be overcome at additional costs, by installing air or heat exchangers.

The test sections of both types of tunnels are divided into either open-jet or closed-jet test sections. Open-jet test sections are usually enclosed in larger structures and do not, except for the road surface, have test section walls. Larger models can therefore be used, as the effect of tunnel blockage and pressure reflections are smaller. Although the stream-wise static pressure distribution is more constant in open-jet sections, they generally have to be shorter due to air dissipation to the surroundings. The reader is referred to section 2.2.3.2 for a definition of the terms used above.

2.2.1.2 Automotive Wind tunnel Testing

Unlike most studies in the aerospace industry where one medium, the test vehicle, is passing through another medium such as air, land vehicles are always in close proximity to a third medium. This third constraint, the road or running surface, also has to be duplicated or simulated in a wind tunnel. Several methods have been developed in an attempt to achieve this. As a first approximation, one of the tunnel walls could be used but an examination of the wall and road-boundary layers reveal significant differences. These are illustrated in Figure 2-2.

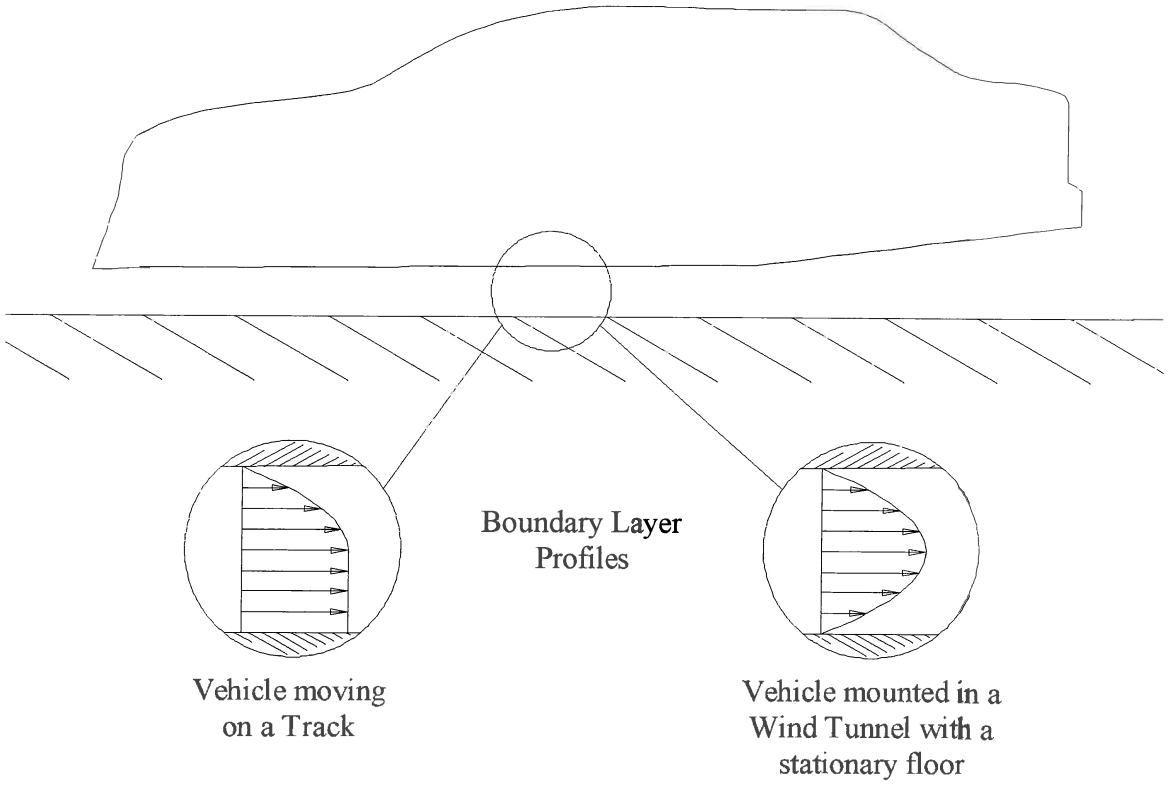

Figure 2-2: Boundary layer profiles of a vehicle moving on the road and mounted in a wind tunnel

In a real fluid, there is a viscous interaction between molecules of dissimilar velocities. This, combined with the condition of no-slip on the surface of an immersed solid[12], leads to a velocity gradient between the surface and the free stream fluid. This is known as the boundary layer and is discussed in more detail in section 2.2.3.1. On the track, under wind-still conditions, the road is stationary relative to the air. Here, boundary-layer development occurs only on the surface of the vehicle. In a wind tunnel however, where the vehicle is stationary relative to the fluid, there is boundary-layer growth on the vehicle surface, as well as on the tunnel floor.

An example of such a wall boundary layer was measured at the General Motors full-scale tunnel[8]. At the end of the test section, it was found to be approximately 130 mm thick at 95% of free stream velocity. Boundary-layer growth can become quite significant when developing racing cars, which often operate with ground clearances of the same order. Various methods aimed at addressing this problem have been developed. The most common solutions to this problem are shown in Figure 2-3.

The simplest and cheapest method of avoiding the boundary layer, is to mount the car on a raised platform situated in the free stream air as shown in Figure 2-3A. Although the tunnel boundary layer is avoided, the vehicle becomes subject to another one formed on the raised platform. This layer can be reduced by inclining the ground plane at a small negative angle[13]. Another method which can be employed, is to raise the model by an amount corresponding to the thickness of the layer. Care also has to be taken not to induce re-circulating flow at the end of the plate, due to an incorrect air split above and below the raised floor.

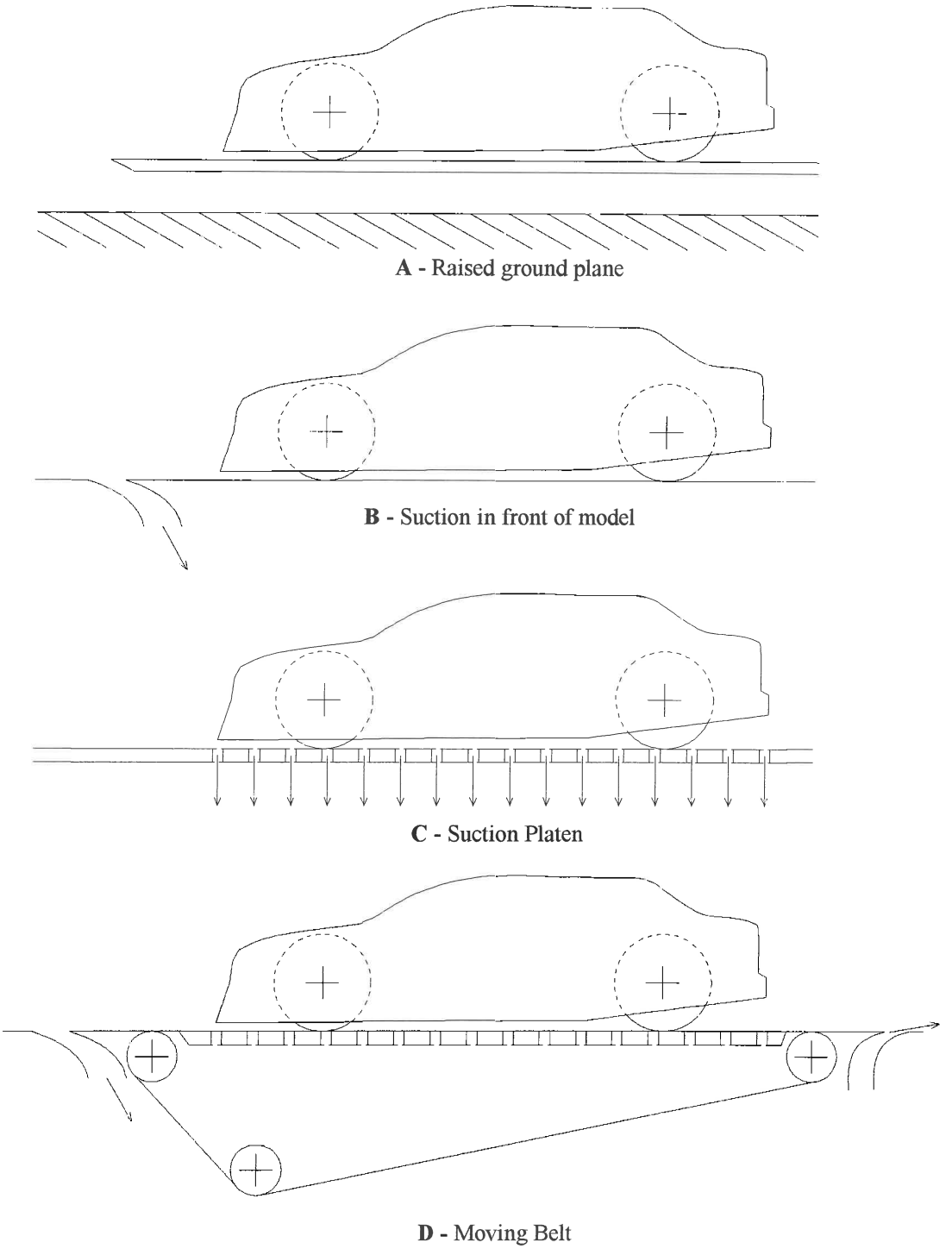

Figure 2-3 : Methods of simulating the moving ground

A slightly more complicated method is depicted in Figure 2-3 B. Here, a span-wise suction slot is placed in front of the model to remove the boundary layer. The amount of suction has to be controlled carefully. Too little will not remove the layer, while excess suction can cause flow reversal. Unfortunately, a new boundary layer will start forming immediately behind the slot, even if the wall layer is removed completely. To overcome this problem, a more

complicated, but more effective method is to introduce a suction platen under the entire test-section floor, as illustrated in Figure 2-3 C. Although this method has the potential to eliminate the boundary layer completely, discrepancies still remain over the amount and variation in length-wise suction which is required for different types of vehicles.

The last illustration in Figure 2-3 is that of a rolling road system with suction slot. Despite the high initial expense, this system is used almost universally by all open-wheel racing teams. This is because the road and car floor interaction is critical to them. The boundary layer is removed by the front suction slot and then prevented from re-forming by synchronising the air and belt speeds.

Mercker and Knape[14] investigated the use of tangential blowing instead of a rolling road for passenger cars. Their results showed discrepancies when measuring the front wheel lift.

The moving ground plane or belt, also makes the simulation of the rotating wheels possible. This is essential to open-wheel racing cars, as the flow field in which the car is immersed, is strongly influenced by the large rotating wheels.

The moving ground plane, does however, introduce several complications. The overhead and rear mounted struts or stings which are required to secure the model (section 2.2.3.3), cause additional flow-field interference. The negative pressure under the racing car necessitates the use of a suction platen to prevent the belt from lifting. The friction induced by the belt running over the platen generates heat, and so platen cooling is required. A belt-steering mechanism is also needed to keep the belt in position at high speeds. The suction platen, belt steering or centring mechanism and related cooling problems, usually limit the operating speed of the tunnel to around 25 to 60 m/s. This can be well below the operating speed of the racing car.

The air which was removed from in front of the belt in Figure 2-3 D, is re-introduced immediately behind the belt. This is to delay the formation of the boundary layer behind the rolling road.

The methods presented in Figure 2-3 are all aimed at eliminating or correcting the tunnel-wall boundary-layer phenomena. The relatively large wake produced by bluff bodies such as motor cars, are not addressed in any of these methods. To compensate for this, Garry et al.[15] found that the rear end of a road vehicle on a raised platform, should be at least four times the square-root of its base area or $x \geq 4A^{1/2}$, from the end of the test section. He observed drag errors of 10 to 20% when this distance was less than $x = 2A^{1/2}$.

2.2.2 Wind-Tunnel Measurement Techniques

The effective development of a racing car depends on the ability of the aerodynamicist to measure and analyse experimental data with high degrees of accuracy and repeatability. The data types can for clarity, be divided into three categories:

- Body forces and coefficients
- Aerodynamic pressures
- Flow visualisation techniques

2.2.2.1 Body Forces and Coefficients

When a body such as a racing car moves through a fluid such as air, the fluid exerts a pressure and a shearing action on the surface of the vehicle. If integrated, a single resultant force can be calculated around the aerodynamic centre, or centre of pressure (COP) of the vehicle. This force can be divided into three independent components, known as the drag, the lift and the side force. If these forces are translated to another point such as the centre of gravity or centre of the wheelbase of the vehicle, three corresponding moments, the rolling, the pitching and the yawing moment appears. It should be noted that the centre of pressure of the side force, is generally not in the same longitudinal position as that of lift force[16]. The sign convention and aerodynamic moments and forces, as defined by the Society of Automobile Engineers (SAE)[17], is shown in Figure 2-4.

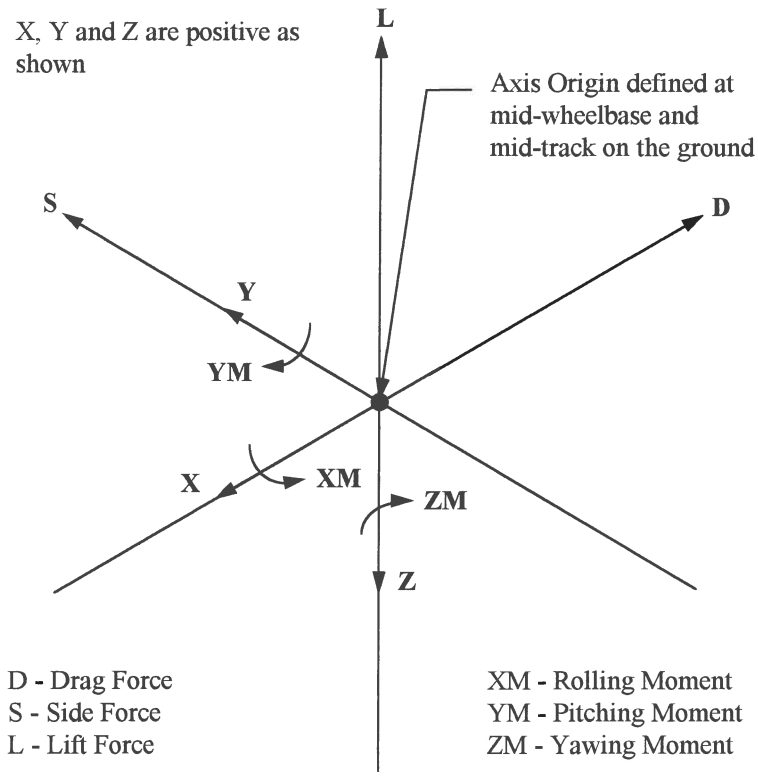

Figure 2-4 - SAE Aerodynamic Force and moment convention

The first of the forces to be discussed, is the drag force. This is a resistive force, which slows a body in motion and acts along the longitudinal axis of the vehicle. The drag force can be divided into two components[18], a pressure or form drag and a friction or surface drag. Pressure drag depends largely on the shape or form of the body and is caused by the pressures acting on the body. The surface drag arises due to shear forces acting on the body. The relative importance of the surface and pressure drag depends largely on the shape of the body. For very slim or streamlined shapes, the surface drag is the dominating factor. For blunt bodies, such as road vehicles and racing cars, the pressure drag is more important.

Lift is the force component acting either perpendicular to the direction of motion, or perpendicular to the free-stream velocity. The aerodynamic lift is primarily produced by the

pressure forces acting on the vehicle[19]. As most racing cars generate negative lift, the positive term "downforce", is almost universally used to describe negative lift.

The side force, is defined as a force acting perpendicular to both the lift and the drag.

The rolling moment acts around the longitudinal axis of the vehicle. This moment is created by asymmetrical lift and side forces between the starboard and port sides. The yawing moment is resolved along the vertical axis of the vehicle. It is the result of non-symmetrical drag and side force distributions along the longitudinal centre line. The pitching moment acts along the lateral axis and is produced by uneven lift and drag forces.

The body forces and moments are mainly measured using wind-tunnel balances[11]. Aerodynamic balances are very complicated, as the forces acting on the body, have to be de-coupled into independent components. Traditionally, this was achieved using intricate mechanically designed balances, but strain gauges have also been in use for a number of years. Balances can measure from one up to six components. Depending on their position in relation to the model, they are known as internal or external balances.

The forces and moments measured in a wind tunnel are seldom directly applicable to the track. This is because the wind-tunnel model scale and test speeds are often different to that experienced when racing. This dilemma is solved by non-dimensionalising the quantities and expressing them as aerodynamic coefficients. These non-dimensional aerodynamic coefficients are defined below.

The Force Coefficients:

$$C_i = \frac{F_i}{\frac{1}{2}\rho V_\infty^2 A} \text{ where } i = \begin{cases} D - \text{Drag} \\ Y - \text{Side} \\ L - \text{Lift} \end{cases}$$

Equation 2-1

The Moment Coefficients:

$$CM_i = \frac{M_i}{\frac{1}{2}\rho V_\infty^2 Al} \text{ where } i = \begin{cases} X - \text{Roll} \\ Y - \text{Pitch} \\ Z - \text{Yaw} \end{cases}$$

Equation 2-2

F_i and M_i are the measured forces and moments respectively, V_∞ is the free-stream velocity, A is the vehicle frontal area and ρ is the density of air. The wheelbase, or total vehicle length can be used as the additional dimension (l) which is required when calculating the moment coefficients.

After Reynolds-Number independence (section 2.2.3.1) is demonstrated prior to a test series, the coefficients can then be converted back to forces and moments, at track or any other operating conditions. This is achieved by applying the same relations as Equation 2-2, but by substituting the tunnel conditions for those experienced on the track. The coefficients can also be used directly, to compare the aerodynamic performance of different vehicles with the same

frontal area. Another application is the calculation of performance related figures. For example, the power requirements can be calculated at any speed, given the drag coefficient, or conversely, the vehicle speed can be determined at any given engine power. The tractive resistance equation[20] as defined in Equation 2-3, is used to derive these quantities.

$$P_w = \left[\mu_r mg\cos\theta + \tfrac{1}{2}\rho C_D A v^2 + mg\sin\theta + m\ddot{\theta}\right]v$$

Equation 2-3

or

$$\begin{pmatrix}\text{Power}\\ \text{Required}\end{pmatrix} = \begin{pmatrix}\text{Rolling}\\ \text{Resistance}\end{pmatrix} + \begin{pmatrix}\text{Drag}\\ \text{Resistance}\end{pmatrix} + \begin{pmatrix}\text{Climbing}\\ \text{Resistance}\end{pmatrix} + \begin{pmatrix}\text{Acceleration}\\ \text{Resistance}\end{pmatrix}$$

where

$$\mu_r = \mu_s + \mu_d v\ ;$$
$$P_w = P_e \eta_d$$

and

C_D = drag coefficient
P_e = engine power $[W]$
η_d = drive train efficiency
v = top speed $[m/s]$
μ_s = static rolling resistance
μ_d = dynamic rolling resistance
μ_r = effective rolling resistance
m = vehicle mass $[kg]$
ρ = density $[kg/m^3]$
A = vehicle frontal area $[m^2]$
$\ddot{\theta}$ = inertial rotation

During analysis, the aerodynamic coefficients are sometimes expressed as ratios in order to obtain a better understanding of the aerodynamic performance of various vehicle configurations. The Lift to Drag ratio (C_L/C_D) is frequently used to compare the efficiency with which different body components generate lift. When the changes in efficiency become very small, this ratio can be modified even further. The ratio of the change in lift to the change in drag ($\Delta C_L/\Delta C_D$) can then be used.

The dynamic pressure ($\frac{1}{2}\rho V^2$), as used to define the coefficients shown previously, can be obtained in a wind tunnel using a pitot-static tube. The pitot static consists of two concentric tubes. An inner tube, open at both ends and an outer tube, sealed at the leading edge barring a series of holes perpendicular to the longitudinal axis. The pitot-static tube is installed by pointing the tip into and parallel to the flow and connecting the other end to a manometer or pressure transducer. The inner tube measures the total air pressure, while the outer tube only measures the static pressure.

The measured pressures are related to one another through the dynamic pressure, as expressed in Bernoulli's equation[8]:

$$P_{tot} = P_{\infty} + \tfrac{1}{2}\rho V^2$$

Equation 2-4

or

Total Pressure = Static Pressure + Dynamic Pressure

The track performance or lap times of racing cars do not only depend on the size of the aerodynamic forces but are also strongly influenced by the position of the centre of pressure and the sensitivity of that point to changes in driving conditions. To determine this, the distribution of the forces has to be known. This is achieved by using a multi-component balance, which measures the aerodynamic moments around a known point, or by using two one-component force balances for each distribution direction. The force distribution can then be resolved using the position of the balance and the geometry of the vehicle.

2.2.2.2 Surface Pressures

The forces and coefficients discussed in the previous section, supply the aerodynamicist with information regarding the performance of the complete vehicle. While this information is adequate when looking at design targets, or when comparing different vehicles, very little insight is gained into the characteristics responsible for these coefficients. Understanding the flow field surrounding the car is often the key to a successful development programme.

The pressures acting on blunt bodies, such as racing cars, are as described earlier, predominantly responsible for the aerodynamic behaviour of the vehicle. Knowledge of these pressures, can therefore lead to a better understanding of the aerodynamic characteristics of the vehicle. One method of achieving this, is to measure the surface-static component of the total pressure. This pressure can be measured by applying the same principle as employed with the pitot-static tube. By measuring pressure perpendicular to the surface, the dynamic component can be excluded. The surface-static pressures can be measured by either placing small tubes or pressure taps through the skin of the body, or by using flush-mounted pressure discs. Care has to be taken to ensure that these devices do not produce local flow disturbances. As with the pitot-static tube, the tubes or discs are connected to a pressure transducer. After measurement, the pressures can then be reduced to non-dimensional pressure coefficients (C_P) as defined below[8].

$$C_P = \frac{P - P_{\infty}}{\frac{1}{2}\rho V_{\infty}^2}$$

Equation 2-5

where the reference static (P_{∞}) and dynamic pressures can be obtained using a pitot-static tube.

The aerodynamic forces can also be calculated from the surface pressures by surface integration, as defined by the following relationship [18]:

$$Dp = \oint p \sin\varphi dS$$

Equation 2-6

Although surface-static pressures can supply information regarding the flow on the surface of the vehicle, it is sometimes necessary to know what the conditions are at distances from the surface. The easiest method of achieving this, is to measure the total pressure using a pitot tube. Unfortunately, their construction and principle of operation limit their use. As with pitot-static tubes, pitot tubes also have to be aligned with the flow. Depending on their design, measurement errors can be incurred at misalignment angles of as little as 3 degrees[11]. Pre-existing knowledge of the flow is therefore required to ensure reliable results. A Kiel tube[11] is more suitable if prior knowledge of the flow field is unavailable, as flow angles in excess of 30° can still be measured with high degrees of accuracy.

The last of the pressure components, the dynamic pressure or flow velocity, can also be measured if required. One method of achieving this is to use a pitot-static tube. Unfortunately, despite the simplistic and robust construction of the tube, it is limited to use under free-stream conditions only. This is partly due to flow-angle sensitivity as described above, but more importantly, it is due to the sensitivity to static variations in the tunnel. The static fluctuations produced by the test object, can significantly influence the results if the tube is placed too close to the body under investigation.

To overcome the deficiencies of the pitot-static tube, multiple-component hot-wire probes can be used[18]. These probes work on the principle that the change in resistance due to cooling of electrically heated wires, can be related to the change in the air velocity responsible for the heat loss. Another method of determining the flow velocity is to use three, five or seven-hole probes. These probes can be described as an assembly of up to seven miniature pitot tubes, each tube measuring a component of total pressure. After measurement, these pressures are fed into a calibration matrix, where the magnitude of the velocity is determined in three dimensions. The precision and size of the calibration matrix is instrumental in determining the accuracy and maximum attainable flow angle.

2.2.2.3 Flow Visualisation

Flow visualisation techniques are used to observe different flow fields around a body[8,18]. They are used to gain qualitative rather than quantitative information and can be very cost effective.

One of the quickest, cleanest and most common methods employed are tufts. These are short, thin strips of string which when attached to the surface of the vehicle at one end, will align themselves with the airflow, thereby indicating local flow patterns. Other popular methods are lasers, smoke generators and oil-based fluorescent paints. The last two methods tend to soil the model and tunnel and are therefore usually only used at the end of a test series.

2.2.3 Wind-Tunnel Testing

One of the first decisions which has to be made when planning a wind-tunnel test programme, is the scale of the test model. When choosing the scale various factors have to be balanced. The operational and possible tunnel construction costs have to be considered. The amount of model detail and precision that is required, the ease and speed with which modifications can be made and the cost of manufacturing model components have to be taken into account. Finally, the desired accuracy of the results plays a major role in selecting the scale.

The first and often most obvious choice when considering the scale, would be to use full-scale models. Full-scale wind-tunnel testing is ideal, when the preservation of detail is important, as the actual vehicle under investigation or an identical model can be used. Full-scale testing does

however have several disadvantages. Vehicle transportation and tunnel installation, when using a model or purpose built racing car, can be expensive and cumbersome. Due to the size of the vehicle, component modifications can use up valuable tunnel time. Full-scale parts are also expensive to manufacture and can have long lead times. During the early phases of a project, sufficient detail is usually not available to warrant a full-scale vehicle. Furthermore, construction costs of the tunnel shell and power requirements tend to vary with the square of the tunnel width[11]. This means that a sedan-based racing car, which would require a tunnel with a cross-sectional area of more than 30 m^2, would need substantial initial investment and also a large operating budget. Currently, there are few tunnels in the world large enough for such tests and even fewer with the ability to simulate road conditions accurately. Full-scale tunnels are therefore usually only used for short test or verification programmes, while smaller-scale models are used when prolonged or continuous development is required.

When the scale of a car is changed it is important not only to reproduce the details of the vehicle, but also to duplicate the flow patterns and forces. The accurate scaling of the body and flow is achieved by a technique known as dimensional analysis.

2.2.3.1 Dimensional Analysis

Fluid flow problems can be characterised by fluid to fluid interactions and by the interaction of the fluid with its boundaries. In order to understand and compare similar fluid-flow phenomena and resulting forces originating under different physical conditions, it becomes necessary to remove the details of a particular flow problem, by non-dimensionalising the parameters defining the flow.[21] For most racing and passenger cars, where the flow is assumed to be incompressible, four non-dimensional relations are important. They are:

$$\begin{aligned}
\text{Finess Ratio} &= \frac{L}{D} = \frac{\text{Length}}{\text{Width}}; \\
\text{Aspect Ratio} &= \frac{S}{W} = \frac{\text{Span}}{\text{Width}}; \\
\text{Relative Roughness} &= \frac{\varepsilon}{D} = \frac{\text{Roughness Height}}{\text{Width}} \\
\text{Reynolds Number(Re)} &= \frac{\rho V L}{\mu} = \frac{\text{Inertial Forces}}{\text{Viscous Forces}}
\end{aligned}$$

Equation 2-7

where ρ is the density, μ is the viscosity, V is the characteristic velocity and L is the length of the vehicle. The first three ratios relate to the physical dimensions of the vehicle, while the fourth, the Reynolds Number, describes the interaction between the body and the fluid. If, when scaling, the full scale values of these ratios are applied to every component of the vehicle and flow, the case under investigation will be duplicated exactly, regardless of the scale employed. Unfortunately, this is rarely achieved. Fine detail and intricate body parts can disappear or not be viable to reproduce when reducing the scale. The surface roughness becomes in most cases impossible to scale as the full-scale surfaces are usually already at their practical minimum limits of smoothness. Although these problems diminish when the amount of detail on the full-scale vehicle is small, the Reynolds Number remains critically important.

External race car aerodynamics can as stated previously, be seen as fluid-flow problems where the flow is essentially incompressible. If the same working fluid is used for both the scale

model and the full-scale car, the Reynolds Number simply becomes proportional to the product of the size and speed of the vehicle. There are however, limits to where this simple product rule can be applied. When the Mach Number[11], defined as the ratio of the air speed to the speed of sound, rises above 0.3, the assumption of incompressibility becomes invalid and so additional corrections have to be made. Furthermore, physical limitations such as the attainable tunnel and rolling-road speeds also influence the maximum Reynolds Number. Scale-model tests are therefore usually performed at lowest possible Reynolds Numbers.

The Reynolds Number, not only defines, but also describes the type of interaction present at the solid-fluid boundary. As stated in section 2.2.1.2, a boundary layer is formed due to the viscous properties of the fluid. Initially, during the interaction, the fluid streamlines remain parallel and momentum exchange between layers is only on a molecular scale. This is known as laminar flow. Unless the free-stream fluid velocity or Reynolds Number is sufficiently low, this type of flow cannot be maintained and soon becomes turbulent. Here, considerable random motion and therefore, significant amounts of momentum exchange, exists between the different layers.

The transition from laminar to turbulent flow occurs at a finite distance from the leading edge of the surface. Therefore, as in the case of smaller scale models, if the length of the surface is reduced but the fluid properties are not adjusted, transition will still occur at the same absolute distance from the leading edge. This distance will however not be at the same relative position as on the full-scale case. This error is compounded even further if flow separation is present. The energetic nature of turbulent flow, will cause the point of separation to be postponed, resulting in a smaller wake than would be present in the laminar case. The body form therefore plays a major role in determining the sensitivity of the tests to the Reynolds Number.

Streamlined bodies such as tear-drop shapes are less sensitive to variations in the Reynolds Number, as the flow remains largely attached. Bluff bodies such as vertical plates are also insensitive as the separation point is fixed by the sharp edges. Unfortunately, cars which can be described as semi-streamlined or semi-bluff bodies, are very dependent on the Reynolds Number as there are few hard edges were the flow is forced to separate.

Tests performed by Barnard[16], showed that for a typical road car the critical Reynolds Number of $5x10^5$ corresponded to a $1/12^{th}$ scale model at 25 m/s. He also found that for a Group C racing car, little variation in the drag was found above a Reynolds Number of around $9.6x10^5$. He found a 25% scale model tested at a velocity of about 40 m/s to be a popular choice. Tests performed in the Swift tunnel in the U.S.A. on a 40% scale model of a open-wheel racing car, suggested that speeds in excess of 60 m/s were needed to obtain reliable downforce figures.[22]

For the purposes of external aerodynamics, most of the internal components such as the engine and gearbox can also be scaled using dimensional analysis. The cooling pack can however not be scaled directly using this approach. A cooling core, disregarding the thermodynamic effects, influences the flow in two major ways. The first is a pressure drop due to the obstruction caused by the core and the second is the flow straightening effect created by the core fins. The simulation of these phenomena, at a smaller scale, are usually achieved by constructing a wire mesh and honeycomb sandwich. The wire mesh is chosen to represent the pressure drop while the honeycomb is used to straighten the flow. The density of the mesh can be calculated using the approach outlined by Masaru[23]. He used the core pressure drop, the porosity coefficient and Reynolds Number to determine the mesh density.

In summary, dimensional analysis can be used to reproduce fluid-flow problems at any scale. The size of the vehicle can therefore be changed, within limits, to fit the requirements of the test facilities. For any given wind tunnel, restrictions are placed on the maximum model size. The reasons for, and effects of these limitations, are presented below.

2.2.3.2 Wind-Tunnel Blockage Effects

When a model is placed into a closed-jet test-section wind tunnel, the walls form an unnatural boundary around the body. This forces the fluid stream lines in the test section closer together, producing higher test-section speeds which results in an over-estimation of the aerodynamic forces. This is known as solid or tunnel blockage. The proximity of the walls can also cause pressure reflections which affect the results even further.

The amount of blockage (B) is defined as the ratio of the cross-sectional area of the model to that of the empty test section[11]. To minimise the blockage errors, the walls of the tunnel have to be placed as far as possible from the model. Construction and operational costs limit the practical extent to which this can be applied. Hucho[18] suggests a blockage ratio of not more than 5%. In addition, tunnel-blockage correction factors, which vary from direct data manipulation such as continuity corrections, to more complicated computations involving pressure measurements along the test-section walls, are also employed. One such correction, which is based on continuity, is that as proposed by Sykes[24]. He performed a series of tests relating to cars and found, that for tests with a blockage of up to 20% and a yaw angle of up to 20°, the correction factor f, to be

$$f = \frac{C_{f,c}}{C_f} \text{ where } f = 1 - 1.9B$$

Equation 2-8

$C_{f,c}$ is the blockage corrected coefficient, C_f is the measured coefficient and B is the blockage ratio. The pressure coefficients can be corrected in a similar way by using the following formula:

$$f = \frac{1 - C_{p,c}}{1 - C_p}$$

Equation 2-9

where $C_{p,c}$ is the corrected pressure coefficient and C_p is the measured pressure coefficient.

Furthermore, the boundary-layer growth along the test section walls also influences the results. The boundary layer effectively reduces the test-section area towards the outlet. This decreases the static pressure along the test section which tends to draw the model downstream, changing the apparent drag. This effect, which is known as longitudinal or horizontal buoyancy[11], can be corrected by having slightly expanding test sections, using slotted walls or open-jet test sections. Empirical corrections to the drag, based on the pressure gradient along the test section, can also be employed.

All the factors discussed so far have to be taken into consideration when determining the size of the model, but usually, it is the availability of wind tunnels and operational funds which are

the overriding factors. Formula One teams, which are at the pinnacle of motor sport, use wind tunnels which can accommodate models ranging from 30 to 50% of full scale.

2.2.3.3 Model Mounting

The methods used to secure models in wind tunnels depend on the tunnel, balance and model types, as well as on the model size. For stationary-road tunnels, the wheels can be fixed directly onto the balance platform itself. For full-scale cars, the wheels and suspension are simply locked, as the weight of the car is usually enough to prevent movement.

The methods employed when rolling-road tunnels are used are more complicated, as the floor cannot be used to fix the models. Struts or wires, which protrude from the model, then have to be used. This arrangement, if pitch automation is not required, usually consists of a single main strut. The strut, which either extends from the top or rear of the model, connects the model to the external balance, or the internal balance to the tunnel frame. It also provides a concealed path for the cabling of control, data acquisition and measurement equipment.

If pitch or ride-height adjustments are required during a test run, additional tail struts or wires are needed for roof-mounted systems. Similar to the main strut, these also extend from the top or rear of the vehicle. Additional fixtures are however not required when a tail-mounted main strut is used, as the boom can be designed to accommodate height adjustments. The mounting struts do unfortunately create additional disturbances to the airflow. These can be summarised as follows:

- The direct contribution of the strut to the drag; called tare. This only applies to external balances.
- The additional blockage caused by the strut.
- The disturbance caused by the strut on the flow field of the model.
- The disturbance caused by the model on the flow field of the strut.

The last two categories are usually grouped together and are called interference. All these effects have to be removed by calibration.

The tare contribution can be reduced by shielding the strut from the airflow, but it cannot be eliminated entirely. The remaining contribution is determined by measuring the drag under operating conditions, with only the strut in the tunnel. The value measured during the tare run, is then subtracted from the model-drag values.

The interference, which is more difficult to determine, can be obtained using an additional balance[11]. In this approach, the model, while left in position, is disconnected from the main system and secured to the second balance. Two sets of data are recorded. The first is taken with the primary strut in position and the second is made without it. The difference between the two readings is taken as the total system interference and includes the tare contribution.

When a rolling-road system is used, the running wheels produce an additional contribution to the tare. This is caused by the stiction in the bearings and the friction between the belt and the wheels. This is determined by running the belt without the wind.

2.2.3.4 Wheels

The aerodynamic performance of racing cars and in particular open-wheeled types, are strongly influenced by the flow patterns created by their rotating wheels. This extends to all regions of

the body-work and includes areas such as engine and brake cooling systems[25]. It is therefore very important to reproduce these effects accurately.

The scarceness of full-scale rolling-road wind tunnels necessitates the use of stationary floors when scale models are not used. Under these conditions, the wake produced by open wheels under dynamic conditions, can be simulated by forcing separation before the apex of the tyre[18]. This arrangement attempts to reproduce the wake of the rotating wheel. The success and accuracy of this method still remains unclear. Various other methods[18] have also been investigated to improve the accuracy of wheel simulation using stationary or rotating wheels. Cogotti[26] found it essential to ensure zero gap between the wheel and the road, as the gap produced a venturi effect which changed the sign of the lift coefficient. Mercker et al.[25] could not find a fixed relationship between stationary and rotating wheel tests. Quantitative as well as qualitative errors could therefore be introduced by incorrect simulation.

During scale-model test programmes, the rolling-road system is still preferred. Besides the problems discussed in section 2.2.1.2 relating to rolling roads, several additional problems arise when wheel rotation is considered. The correct surface roughness of the tyre and road is very difficult to reproduce or to scale. The tyre profile is not symmetrical under dynamic loading on the track. The contact patch area is flat, while at the apex of the tyre, camber is present. As the contact patch is usually considered critical for simulation, the tyre camber has to be omitted. A flat running surface is also required to reduce belt wear.

The weight of the model wheels are also important. Heavy wheels can cause excessive belt wear, while light wheels can induce bounce or vibration which can influence the results. These can be overcome by mounting the wheels off the model, using a side-mounted strut system for each wheel. Unfortunately, this system requires additional balances to measure the drag of each wheel. The side struts also introduce aerodynamic interference which has to be removed. Ride-height adjustments become more difficult, as additional equipment is required to move suspension and brake-duct components. Wheel-steering studies do however become easier when the wheels are mounted off the model.

On a rolling road, wheel lift becomes very difficult to measure due to the interaction with the running surface. Mercker[25] devised a method whereby the lift of full-scale production vehicles could be measured.

2.2.3.5 Full-Scale Correlation

Ultimately all wind-tunnel tests have to be related to the racing car on the track. In the top open-wheel racing formulae, this becomes crucial as the aerodynamic coefficients form an integral part of the set-up of the car. Accurate correlation between the track and the tunnel is therefore required.

A Formula One racing car operates with ride heights typically ranging between 5 and 50 mm. When this gap is scaled for smaller model tests, track correlation becomes essential, as the air gap does not scale linearly. Additional ride-height adjustments therefore have to be made. When these corrections are made accurately, the top speed of the racing car can be predicted to within 1 mph and the downforce to within 1 lb.[27]

The success of any wind-tunnel programme, using the test methods and procedures as described in this chapter, can still be undermined by repeatability problems. Conclusions can

never be drawn from data, which have discrepancies smaller than the repeatability of the experiment. It is therefore essential, to construct precision models and test them in quality wind tunnels with highly repeatable systems.

2.3 Computational Literature Study

Computational Fluid Dynamics is the study of aerodynamic and fluid-mechanic problems using numerical solutions[28].

The governing fluid-flow equations and underlying assumptions are solved by discretising the equations and then implementing a numerical scheme over a predetermined solution domain. The governing equations are therefore discussed first. This is followed by the construction of the solution domain and finally, two different numerical-solution schemes are presented.

2.3.1 The Governing Equations

Fluids[ii] consist of molecules which are continuously in a state of random motion[29]. The large scale motion of fluids add a uniform velocity vector to the motion of each molecule. If the sample of molecules under investigation is large enough, the individual molecular motions are not detectable and so only the large scale motion needs to be investigated.

The governing equations of fluid dynamics are based on three universal laws of conservation[30]:

- Conservation of Mass (Continuity equation)
- Conservation of Momentum (Newton's Second Law)
- Conservation of Energy (First Law of Thermodynamics)

Although Navier and Stokes were only responsible for the equations relating to the conservation of momentum, all the governing equations are currently, collectively known as the Navier-Stokes equations. These equations can be expressed either in integral or differential form.

In integral form the conservation of mass over an arbitrary fixed control volume Ω can be expressed as:

$$\frac{\partial}{\partial t}\iiint_{\Omega}\rho d\Omega + \iint_{S} n \bullet \rho V dS = 0$$

Equation 2-10

or

$$\begin{pmatrix}\text{Net outflow of mass leaving the control}\\ \text{volume}\end{pmatrix} = \begin{pmatrix}\text{Time rate of mass change in the control}\\ \text{volume}\end{pmatrix}$$

where t is the time co-ordinate, ρ is the fluid density, V is the velocity vector, S is the boundary surface surrounding Ω and n is the unit normal vector to S. This equation is valid for any continuous perfect or real fluid.

[ii] A fluid can be defined as any substance which cannot sustain a shear force when at rest.

The conservation of momentum in integral form can be expressed over the fixed control volume Ω as:

$$\frac{\partial}{\partial t}\int\int\int_{\Omega}\rho V d\Omega + \int\int_{S}\rho V\,(n \bullet V)dS = \int\int\int_{\Omega}\rho f d\Omega + \int\int_{S} n \bullet T_{ij} dS$$

Equation 2-11

or

$$\begin{pmatrix}\text{Time rate of change in}\\ \text{linear momentum within}\\ \text{the volume}\end{pmatrix} + \begin{pmatrix}\text{The net flow of}\\ \text{momentum across}\\ \text{the volumes surface}\end{pmatrix} = \begin{pmatrix}\text{Body forces}\\ \text{e.g. gravity}\end{pmatrix} + (\text{Surface forces})$$

where f is the body force acting on the fluid per unit mass and T_{ij} is the total or Newtonian flow stress tensor. The Newtonian stress tensor is given by

$$T_{ij} = -\delta_{ij} p + \sigma_{ij}$$

where δ_{ij} is the Kronecker Delta defined as

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Equation 2-12

and p is the fluid pressure characterised as a normal stress and σ_{ij} is the viscous stress tensor.

The momentum equation is valid for steady and unsteady flows of continuous liquids or gases which obey the assumptions of Stokesian and Newtonian fluids. Stokesian fluids obey four conditions:

- The fluid must behave as a continuum in which the stress tensor is only a function of the rate of the strain tensor and of the thermodynamic state of the fluid.
- The fluid must be homogeneous.
- The fluid must be isentropic.
- When the fluid is at rest the viscous stress tensor and the rate of the strain tensor must both vanish.

A Newtonian fluid is a fluid where the stress tensor is linearly proportional to the rate of the strain tensor.

Using the Stokesian-Newtonian fluid assumption as well as Stokes' hypothesis[31], as defined in Equation 2-14, the viscous stress tensor can be written as

$$\sigma_{ij} = \mu\left[\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \frac{2}{3}\delta_{ij}\frac{\partial u_k}{\partial x_k}\right]$$

Equation 2-13

or

$$\begin{pmatrix}\text{Viscous stress}\\ \text{tensor}\end{pmatrix} = \begin{pmatrix}\text{Normal and shear stress}\\ \text{due to fluid friction}\end{pmatrix} - \begin{pmatrix}\text{Normal stress contribution}\\ \text{to the fluids volume change}\end{pmatrix}$$

where μ is the first coefficient of viscosity, u and x are the Cartesian velocity and position co-ordinates respectively. i, j and k are in standard tensor notation, each cycled over 1, 2 and 3.

Stokes' hypothesis which states that the bulk viscosity which must vanish, is defined as

$$\mu' = \lambda + \frac{2}{3}\mu$$

Equation 2-14

The conservation of energy can be expressed in integral form as:

$$\frac{\partial}{\partial t}\int\int\int_{\Omega} E d\Omega + \int\int_{S} n \bullet EV dS = \frac{\partial}{\partial t}\int\int\int_{\Omega} q_{ext} d\Omega - \int\int_{S} n \bullet q dS + \int\int\int_{\Omega} \rho f \bullet V d\Omega + \int\int_{S} n \bullet \left(T_{ij} \bullet V\right) dS$$

Equation 2-15

or

$$\begin{pmatrix}\text{Time rate of increase in}\\ \text{total energy per unit}\\ \text{volume in}\end{pmatrix} + \begin{pmatrix}\text{Net increase in total}\\ \text{energy per unit volume}\\ \text{by convection across S}\end{pmatrix} = \begin{pmatrix}\text{Time rate of increase in heat}\\ \text{added to } \Omega \text{ from sources}\\ \text{e.g. radiaton}\end{pmatrix}$$

$$- \begin{pmatrix}\text{Heat loss due to}\\ \text{conduction across S}\end{pmatrix} + \begin{pmatrix}\text{increase in energy due to the}\\ \text{work done on } \Omega \text{ by the body}\\ \text{forces}\end{pmatrix} + \begin{pmatrix}\text{Work done by the}\\ \text{stress tensor on S}\end{pmatrix}$$

The total Energy per unit volume (E) is defined as:

$$E = \rho\left(e + \frac{u^2 + v^2 + w^2}{2}\right)$$

Equation 2-16

e is the internal energy per unit mass, q_{ext} is the heat added from external sources and q is the heat flux due to conduction.

The heat conduction in a fluid is usually assumed to be governed by Fourier's heat conduction law which is defined as:

$$q = -k\nabla T$$

Equation 2-17

where k is the coefficient of thermal conductivity and T is the fluid temperature in an absolute temperature scale such as Kelvin.

If the body-force terms and the external-heat source terms are omitted, the conservation laws described above can be expressed in a special vector format:

$$\frac{\partial}{\partial t}\iiint_{\Omega} U d\Omega + \iint_{S} (F + G + H) \bullet n dS = \iint_{S} (F_v + G_v + H_v) \bullet n dS$$

where

$$U = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}; \quad F = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ (E+p)u \end{bmatrix} i; \quad G = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vw \\ (E+p)v \end{bmatrix} j; \quad H = \begin{bmatrix} \rho w \\ \rho uw \\ \rho vw \\ \rho w^2 + p \\ (E+p)w \end{bmatrix} k;$$

$$F_v = \begin{bmatrix} 0 \\ \sigma_{xx} \\ \sigma_{yx} \\ \sigma_{zx} \\ \beta_x \end{bmatrix} i; \quad G_v = \begin{bmatrix} 0 \\ \sigma_{xy} \\ \sigma_{yy} \\ \sigma_{zy} \\ \beta_y \end{bmatrix} j; \quad H_v = \begin{bmatrix} 0 \\ \sigma_{xz} \\ \sigma_{yz} \\ \sigma_{zz} \\ \beta_z \end{bmatrix} k$$

and

$$\beta_x = u\sigma_{xx} + v\sigma_{xy} + w\sigma_{xz} + k\frac{\partial T}{\partial x}; \quad \beta_y = u\sigma_{yx} + v\sigma_{yy} + w\sigma_{yz} + k\frac{\partial T}{\partial y};$$

$$\beta_x = u\sigma_{zx} + v\sigma_{zy} + w\sigma_{zz} + k\frac{\partial T}{\partial z}$$

Equation 2-18

Q is the dependent or conserved variable vector. F, G and H are the inviscid flux vectors and F_v , G_v and H_v are the viscous flux vectors.

The first row of the vectors corresponds to the continuity equation. The second, third and fourth rows describe the momentum equations, while the last row represents the energy equation.

By transforming all the surface integrals to volume integrals using Gauss' Divergence theorem[31], the integral equations which were defined above can be transformed to a set of differential equations. When this is done the following flux-vector form of the differential equation results:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} = \frac{\partial F_v}{\partial x} + \frac{\partial G_v}{\partial y} + \frac{\partial H_v}{\partial z}$$

Equation 2-19

By inspection, it can be seen that the five equations presented in the integral or differential vector format above, contain seven unknowns (p, ρ, T, e, u, v, w). Two additional relations are therefore required to close the system of equations. This is achieved from the equation of state, by establishing a relationship between the thermodynamic variables (p, ρ, T, e), as well as to relate the transport properties (μ and k) to the thermodynamic properties. According to the state principle of thermodynamics, the local thermodynamic state is fixed by any two

independent thermodynamic variables, provided that the chemical composition is not changed due to diffusion or to finite rate reactions[30]. For example, if e and ρ are chosen as the two independent variables, equations of state of the following forms are required:

$$p = p(e,\rho); \qquad T = T(e,\rho)$$

Equation 2-20

2.3.2 Reynolds Equations for Turbulent Flows

The unsteady Navier-Stokes equations can be regarded as the governing equations for turbulent flow, but due to the space scales involved, computers are not yet capable of solving these equations directly for practical problems. A common estimate is that 10^5 grid points are required to resolve 1 cm^2 of a typical turbulent flow field[29].

An alternative approach to solving these equations directly is to break the dependent variables in the conserved equations, into time-mean and fluctuating components and then to take time averages of the governing equations. Currently the majority of turbulent research is directed towards the solution of these time-averaged or Reynolds-Averaged equations.

There are two time-averaging procedures currently in use. They are the standard or conventional and the mass or Favré averaging approaches. The mass-averaging approach is often preferred for compressible and external flows due to its compact form. For flows where the density fluctuations can be neglected, the two approaches become identical.

In the conventional averaging approach, the dependent variables are split into mean and fluctuating parts:

$$f = \overline{f} + f'$$

Equation 2-21

where

$$\overline{f} = \frac{1}{\Delta t}\int_{t_o}^{t_o+\Delta t} f dt$$

f is any dependent variable such a u or v. $\overline{f}$ represents the time-averaged or mean part and f' is the fluctuating component. Δt should be large compared to the period of random fluctuations in the turbulent flow-field, but small compared to the slower flow-field variations of ordinary unsteady flows.

In the mass-weighted averaging approach only the velocity and thermal variables are defined as:

$$\tilde{f} = \frac{\overline{\rho f}}{\overline{\rho}}$$

Equation 2-22

The remaining variables are averaged as in the standard approach.

To simplify the Reynolds-Averaged Navier-Stokes Equations, it is assumed that triple correlations are smaller than double correlations and are therefore neglected. The fluctuating

components of viscosity, thermal conductivity and specific heat are usually small and are also neglected. The Favré-averaged Navier-Stokes equations can then be written as:

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}+\frac{\partial H}{\partial z}=\frac{\partial F_v}{\partial x}+\frac{\partial G_v}{\partial y}+\frac{\partial H_v}{\partial z}+\frac{\partial F'}{\partial x}+\frac{\partial G'}{\partial y}+\frac{\partial H'}{\partial z}$$

Equation 2-23

where

$$U=\begin{bmatrix}\bar{\rho}\\ \overline{\rho\tilde{u}}\\ \overline{\rho\tilde{v}}\\ \overline{\rho\tilde{w}}\\ \bar{E}\end{bmatrix};\ F=\begin{bmatrix}\overline{\rho u}\\ \overline{\rho\tilde{u}^2}+\bar{p}\\ \overline{\rho\tilde{u}\tilde{v}}\\ \overline{\rho\tilde{u}\tilde{w}}\\ \overline{(\bar{E}+\bar{p})\tilde{u}}\end{bmatrix};\ F_v=\begin{bmatrix}0\\ \tilde{\sigma}_{xx}\\ \tilde{\sigma}_{yx}\\ \tilde{\sigma}_{zx}\\ \tilde{\beta}_x\end{bmatrix};\ F'=\begin{bmatrix}0\\ -\overline{\rho u''u''}\\ -\overline{\rho u''v''}\\ -\overline{\rho u''w''}\\ -\overline{(E+p)u''}+\overline{\sigma_{xx}u''}+\overline{\sigma_{xy}v''}+\overline{\sigma_{xz}w''}\end{bmatrix}$$

The other fluxes are expanded in a similar way. The viscous stress tensor in is given by

$$\tilde{\sigma}_{ij}=\mu\left[\frac{\partial\tilde{u}_i}{\partial x_j}+\frac{\partial\tilde{u}_j}{\partial x_i}-\frac{2}{3}\delta_{ij}\left(\frac{\partial\tilde{u}_k}{\partial x_k}\right)\right]$$

Equation 2-24

The $-\overline{\rho u_i'' u_j''}$ and $-\overline{(E+p)v''}$ terms are new apparent turbulent stresses and heat-flux quantities. These are collectively known as Reynolds stresses. To close this system of equations, either additional equations involving the new unknowns have to be found, or assumptions have to be made which relate the new quantities to the time-mean flow variables. This is most often done with the use of turbulence models.

2.3.3 Turbulence Models

The turbulent-boundary layer is often sub-divided according to the dominance of the shearing action that is present[33]. In one definition, the laminar sub-layer is classified as the very small area next to the surface where viscous shear is important. The outer region, known as the fully turbulent zone, is defined as the layer where turbulent shearing is the overriding consideration. The region connecting the two zones is called the buffer zone. In another definition, the boundary layer is divided into two regions. Here, the inner region consists of the laminar sub-layer, the buffer zone and part of the fully turbulent zone. The remaining part of the turbulent boundary layer is then known as the outer region. In order to identify the upper limit of the inner zone a non-dimensional space co-ordinate, y^+ is used. This term is defined as:

$$y^+=y\frac{u^*}{v}$$

Equation 2-25

where u^* is the friction velocity.

u^* is defined as:

$$u^* = \left(\frac{\tau_\omega}{\rho}\right)^{\frac{1}{2}}$$

Equation 2-26

and τ_ω is the wall-shear stress

A popular estimate for the start of the outer region is a y^+ value of about 400[32].

Generally, turbulence models relate the fluctuating correlations to the mean-flow properties by empirical constants. As implied by the name, these schemes do not solve but only model the turbulent flow. Turbulence models can broadly be classified into two categories:

- Eddy-viscosity models
- Reynolds-stress models

Eddy-viscosity models use the Boussinesq approximation, while the Reynolds-stress models do not assume that the turbulent shearing stress is proportional to the rate of mean strain. Reynolds-stress models are more expensive computationally, as they solve a partial differential equation (PDE) for each of the turbulence terms and are therefore, currently not yet viable as engineering tools. As a result, eddy-viscosity models are the most widely used turbulence models for practical flows today.

The turbulent shear-stress is defined as the stress resulting from momentum exchange in the fluid. It is assumed that the fluid travels a finite distance before exchange takes place. This distance is named after its founder and is known as the Prandtl mixing length[32]. The Prandtl's mixing length concept is defined as[28]:

$$\mu_{turb} = \rho \ell^2 \left|\frac{\partial u}{\partial y}\right|$$

Equation 2-27

where ℓ, the mixing length is the distance away from the wall where particles maintain their original momentum and μ_{turb} is turbulent viscosity.

2.3.3.1 Eddy-Viscosity Models

The Boussinesq assumption relates the Reynolds stresses[28] to the mean flow quantities by

$$-\overline{\rho u_i'' u''}_j = \mu_T \left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i} - \frac{2}{3}\delta_{ij}\frac{\partial \tilde{u}_k}{\partial x_k}\right) - \frac{2}{3}\delta_{ij}\rho k_T$$

Equation 2-28

and

$$-c_p \overline{\rho T'' u_i''} = k_{Tc}\frac{\partial \tilde{T}}{\partial x_i}$$

Equation 2-29

k_{Tc} is the turbulent conductivity which is related to the eddy viscosity by $k_{Tc} = c_p\mu_T/Pr_T$, where Pr_T is the turbulent Prandtl number, μ_T is the turbulent viscosity and k_T is the turbulent kinetic energy.

Substituting the Eddy-viscosity model equations into the Reynolds-averaged equations

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}+\frac{\partial H}{\partial z}=\frac{\partial F_v}{\partial x}+\frac{\partial G_v}{\partial y}+\frac{\partial H_v}{\partial z}$$

yields the following:

$$U=\begin{bmatrix}\bar{\rho}\\ \overline{\rho\tilde{u}}\\ \overline{\rho\tilde{v}}\\ \overline{\rho\tilde{w}}\\ \bar{E}\end{bmatrix};\ F=\begin{bmatrix}\overline{\rho u}\\ \overline{\rho\tilde{u}^2}+\bar{p}\\ \overline{\rho\tilde{u}\tilde{v}}\\ \overline{\rho\tilde{u}\tilde{w}}\\ \left(\bar{E}+\bar{p}\right)\tilde{u}\end{bmatrix};\ F_v=\begin{bmatrix}0\\ \tilde{\tau}_{xx}\\ \tilde{\tau}_{yx}\\ \tilde{\tau}_{zx}\\ \tilde{\beta}_x\end{bmatrix}\quad \text{etc.}$$

Equation 2-30

and

$$\tilde{\tau}_{ij}=\left(\mu+\mu_T\right)\left[\frac{\partial\tilde{u}_i}{\partial x_j}+\frac{\partial\tilde{u}_j}{\partial x_i}-\frac{2}{3}\delta_{ij}\left(\frac{\partial\tilde{u}_k}{\partial x_k}\right)\right]-\frac{2}{3}\delta_{ij}\overline{\rho k}_T$$

$$\tilde{\beta}_x=\tilde{u}\tilde{\tau}_{xx}+\tilde{v}\tilde{\tau}_{xy}+\tilde{w}\tilde{\tau}_{xz}+\left(k+k_T\right)\frac{\partial\tilde{T}}{\partial x};$$

The same form as that of the laminar Navier-Stokes equations as defined in Equation 2-19 can therefore be retained.

It has been shown experimentally that for attached flows up to Mach numbers of five[32], the turbulent structure remains unchanged and is similar to that of incompressible flows. Under these conditions, density fluctuations can be neglected.

Eddy-viscosity models are generally classified according to the number of supplementary PDE's utilised. The most common models are zero and two-equation models.

In addition to the underlying Boussinesq assumption, zero-equation or algebraic models further assume that the local rate of turbulence production is equal to the rate of turbulence dissipation. These models also omit the convection of turbulence. A study performed by Marvin[33] suggests that algebraic models in their standard forms fail to capture the flow adequately when subjected to moderate-to-adverse pressure gradients. In certain cases, modifications to the standard model using a relaxation equation improved the situation somewhat, but the results were dependent on the choice of the relaxation length. The accuracy of solutions generated by algebraic models are strongly dependent on the type of flow field, the type of model and subsequent model modifications.

Unlike zero-equation models, two-equation models include the past history of the flow. This is achieved by deriving transport equations, which are based on the Navier-Stokes equations.

Unfortunately, the accuracy of these models are dependent on the assumptions made when evaluating the terms in the transport equations. This is a problem which is shared by all higher-order closure schemes. Research has therefore, been directed at improving the accuracy of the underlying assumptions and constants. These models are the most widely used eddy-viscosity models today.

The most popular two-equation model in current use, is the k-ε turbulence model or derivatives of it. These models solve differential transport equations for the turbulence energy (k) and dissipation rate (ε). The differences between most of these models lie in the form of the equations or the treatment of the near-wall region. In the standard version of the STAR-CD[3] model, high Reynolds-Number forms of the k and ε equations are used in conjunction with an algebraic "law of the wall" function.

Two variations of the standard k-ε model which are supplied by STAR-CD, are the Renormalisation Group (RNG) model and Chen's k-ε model. The RNG model is claimed[3] to be more accurate and general due to the fundamental nature of the approach. This model includes an additional term in the dissipation equation. Chen's variation of the k-ε model includes the production time scale and the dissipation time scale in closing the dissipation equation. This allows the energy transfer mechanism of turbulence to respond to the mean strain rate more effectively.

The STAR-CD version of the standard k-ε model is suitable for fully turbulent, incompressible or compressible flows. Buoyancy is also taken into account to a certain degree.

Wall functions are required as the k-ε models are not valid near the wall where molecular and turbulence effects are of the same order of magnitude. The following assumptions are made when wall functions are employed:

- One dimensional behaviour of the flow. Variations in velocity and other flow parameters are assumed to be normal to the wall.
- Uniform shear stress in the boundary layer. The effects of pressure gradient and body forces are assumed to be small.
- Shear stress and velocity vectors are unidirectional.
- The rate of the production of turbulence energy equals the rate of dissipation.
- Linear variation of the turbulence length scale.

The accuracy of the wall functions depends on the degree to which the assumptions and approximations as presented above, correlate to the physical problem.

Wall functions effectively provide turbulent-boundary conditions at the outer edge of the layer for the turbulent-transport equations. These are based on known distribution of velocity, temperature and turbulence parameters in a one dimensional contexts.

The alternative approach proposed by STAR-CD to using wall functions is to use low Reynolds-Number mixing-length models. The models calculate the turbulent viscosity by extending Prandtl's mixing-length hypothesis to three dimensions. Transport equations are therefore not required. A switch is made from the mixing-length model to the high Reynolds-Number turbulence model at a user defined distance from the wall. This distance should be in the range of $y^+ \approx 30$ to 150 to ensure that the normal k-ε transport equations

apply[3]. The switching criteria is based on the equality, within a specified tolerance, of the turbulent viscosity as calculated by the near wall and far-wall models.

Despite the extensive usage of the various k-ε models, Cambers and Wilcox[34] demonstrated two flaws for wall-bounded flows. They found the model to be ill-behaved when approaching a solid wall and inaccurate for flows with adverse pressure gradients. Attempts to correct this problem by introducing damping functions, have resulted in other deficiencies as stated by Yang and Shih[35]. They attempted to correct the problem of near-wall singularities, by developing a near-wall k-ε model. This model failed to capture the adverse pressure gradients properly.

To overcome the deficiencies of the various k-ε models, a significant number of alternative models have been developed. One of the more successful models to be developed was the k-ω model of Wilcox[36]. However, Menter[37] found this model to be strongly dependent on the value of the free-stream dissipation rate (ω_f), when applied to free-shear layers. He changed the eddy viscosity by as much as 100% by simply reducing ω_f. Despite this deficiency, the k-ω model is superior to the various k-ε models in other areas, as:

- No damping functions are required.
- Adverse pressure gradients are captured with higher accuracy.
- Simple Diriclet boundary conditions can be imposed.
- It is numerically more efficient and more stable. Wilcox[38] claims that the k-ω model is about 30% more efficient.
- Surface roughness and mass injection can be employed without changing the governing equations.

To eliminate the free-stream deficiency of the k-ω model, Menter[4] devised a model which he called the Baseline or BSL model. This model, although based on the k-ω, is a combination of the standard k-ω and k-ε models. The BSL model however, still suffered the same deficiency as all eddy-viscosity models, as the transport of the principle turbulent shear stress $-\rho\overline{u_i'' u_j''}$ was not accounted for. To include this within the framework of the eddy-viscosity models, Menter modified his BSL model to include the transport of turbulent shear stress. He called it the "Shear-Stress Transport" or SST model. Menter showed that the SST model correlated better to experimental data than the k-ε, k-ω and BSL models for a variety of flow situations. This method is discussed in more detail in Appendix A.

2.3.4 Grid Generation

The governing equations, whether in differential or integral form, are solved by discretising the equations over the flow field of interest. To achieve this, the computational domain has to be divided up into discrete points. A grid is therefore required.

Numerical algorithms and discretising procedures are best suited to be solved over uniformly-spaced orthogonal grids. These are difficult to achieve, as most physical flow problems have irregular boundaries. Interpolating procedures are therefore required at the perimeter. This generally means a loss in grid precision, but more importantly, a loss in computational accuracy becomes unavoidable. This problem can be overcome by transforming the irregular non-orthogonal physical domain to an equally-spaced orthogonal computational domain[32]. This also facilitates grid clustering in regions were high flow gradients are expected.

The general transformation that is used is an independent variable transformation from the physical domain (x, y, z), to the computational domain (ξ, η, ζ) involving non-orthogonal co-ordinates[39]. This transformation can be written as:

$$\begin{aligned}\tau &= t\\ \xi &= \xi(t,x,y,z)\\ \eta &= \eta(t,x,y,z)\\ \zeta &= \zeta(t,x,y,z)\end{aligned}$$

Equation 2-31

A unique single-valued relationship between the generalised and physical co-ordinates is assumed.

Using the chain rule of partial differentiation, the metrics of the transformation (ξ_x, η_x, ζ_x, ...) can be expressed, after mathematical manipulation, in terms of primitive metrics (x_ξ, x_η, x_ζ, ...) as follows[iii]:

$$\begin{aligned}\xi_x &= J(y_\eta z_\zeta - y_\zeta z_\eta)\\ \xi_y &= J(x_\zeta z_\eta - x_\eta z_\zeta)\\ \xi_z &= J(x_\eta y_\zeta - x_\zeta y_\eta)\\ \eta_x &= J(y_\zeta z_\xi - y_\xi z_\zeta)\\ \eta_y &= J(x_\xi z_\zeta - x_\zeta z_\xi)\\ \eta_z &= J(x_\zeta y_\xi - x_\xi y_\zeta)\\ \zeta_x &= J(y_\xi z_\eta - y_\eta z_\xi)\\ \zeta_y &= J(x_\eta z_\xi - x_\xi z_\eta)\\ \zeta_z &= J(x_\xi y_\eta - x_\eta y_\xi)\\ \xi_t &= -(x_\tau \xi_x + y_\tau \xi_y + z_\tau \xi_z)\\ \eta_t &= -(x_\tau \eta_x + y_\tau \eta_y + z_\tau \eta_z)\\ \zeta_t &= -(x_\tau \zeta_x + y_\tau \zeta_y + z_\tau \zeta_z)\end{aligned}$$

Equation 2-32

The last three equations can be rewritten using the first nine metric expressions

$$\begin{aligned}\xi_t &= -J(x_\tau(y_\eta z_\zeta - y_\zeta z_\eta) + y_\tau(x_\zeta z_\eta - x_\eta z_\zeta) + z_\tau(x_\eta y_\zeta - x_\zeta y_\eta))\\ \eta_t &= -J(x_\tau(y_\zeta z_\xi - y_\xi z_\zeta) + y_\tau(x_\xi z_\zeta - x_\zeta z_\xi) + z_\tau(x_\zeta y_\xi - x_\xi y_\zeta))\\ \zeta_t &= -J(x_\tau(y_\xi z_\eta - y_\eta z_\xi) + y_\tau(x_\eta z_\xi - x_\xi z_\eta) + z_\tau(x_\xi y_\eta - x_\eta y_\xi))\end{aligned}$$

Equation 2-33

[iii] Duplicated from: Craig, K.J., "Numerical Thermo Fluids - MSM780", Grid Generation Lecture Notes, University of Pretoria, 1994. Courtesy of Professor K.J. Craig

The Jacobian of the transformation, J, is defined as:

$$J = \frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)} = \begin{vmatrix} \xi_x & \xi_y & \xi_z \\ \eta_x & \eta_y & \eta_z \\ \zeta_x & \zeta_y & \zeta_z \end{vmatrix} = \frac{1}{J^{-1}} = 1 \Big/ \frac{\partial(x,y,z)}{\partial(\xi,\eta,\zeta)} = 1 \Big/ \begin{vmatrix} x_\xi & x_\eta & x_\zeta \\ y_\xi & y_\eta & y_\zeta \\ z_\xi & z_\eta & z_\zeta \end{vmatrix}$$
$$= 1/\left[x_\xi\left(y_\eta z_\zeta - y_\zeta z_\eta\right) - x_\eta\left(y_\xi z_\zeta - y_\zeta z_\xi\right) - x_\zeta\left(y_\xi z_\eta - y_\eta z_\xi\right)\right]$$

Equation 2-34

When these mapping procedures are employed the physical domain can be of an irregular nature without affecting the computational accuracy.

The success of the transformations do however place a number of constraints on the grid in the physical domain. These are summarised below:

- The mapping must be one to one. This will ensure, that grid lines of the same family do not cross.
- The grid lines should be smooth. This will ensure the continuous transformation of derivatives.
- The grid should have clustered points in the physical domain where large flow gradients are expected.
- Excessive grid skewness should be avoided.

These requirements have lead to the evolution of a number of grid topologies. In two dimensions these are known as the "C"-, "H"- and "O" grids. A number of techniques have been developed to construct these grids. The nature of the generation methods, make it possible to place them in different groups. These are:

- Algebraic methods
- Partial Differential methods
- Unstructured methods
- Adaptive methods

2.3.4.1 Algebraic Grid Generators

Algebraic grid generation methods are the least complicated techniques that can be employed. Although computationally fast, they are difficult to implement for complicated geometry. These meshes cannot always control interior skewness or smoothness. Discontinuities at the boundaries can also propagate into the interior, leading to errors caused by sudden changes in the metrics.

Algebraic methods employ interpolating schemes in order to obtain the boundary and interior cell points. If stretching algorithms are used for construction, grid clustering can easily be achieved. Stretching procedures can also be applied to ensure grid smoothness, when the grid is initially generated in multiple blocks. Multiple-block construction techniques become necessary when a combination of grid topologies are utilised, or when multiple obstacles are present in the domain.

Stretching functions range from simple geometric methods, where there is only control over the spacing of the first cell, to more sophisticated algorithms such as that of Vinokur[39]. He

employed hyperbolic relations to develop a two-sided procedure, giving control over the first and last spacing, the number of grid points as well as the total line distance.

To improve the orthogonal nature of algebraic grids close to solid boundaries, blending functions[40] can be used. Blending can be achieved in two steps. Firstly, by projecting an existing arc length onto a surface normal, and then by blending the existing grid and the surface normals. This procedure is outlined below.

$$x(\ell) = x_1(\ell) f_1 + x_2(\ell) f_2$$

Equation 2-35

where

$$f_1 = \frac{2}{\pi} \tan^{-1}\left[\left(\frac{\ell}{L_t}\right)^p\right]$$

$$f_2 = \frac{2}{\pi} \tan^{-1}\left[\left(\frac{L_t}{\ell}\right)^p\right]$$

Equation 2-36

L_t is the transition index, p is the power of the blending function, x_1 and x_2 are the generic co-ordinates of the original and projected points respectively and ℓ is the index. The number of intervals to blend, the transition index and the power of the blending function has to be specified.

2.3.4.2 Partial Differential Grid Generators

Partial Differential Equation grid generators, are the most commonly used among grid-generation methods because of their control capabilities. In these methods, the grid points in the physical space are solved based on a uniformly spaced grid in the computational domain. These methods[39] are categorised as elliptic, parabolic and hyperbolic systems of PDE's[iv].

Eliptic grid generators are currently widely used. These are suited to cases where all the boundaries of the domain are known. The boundaries can either be specified as grid-line slopes or as discrete grid points. The Thompson, Thames and Mastin (TTM)[41] method is one of the more popular schemes.

[iv] A second order PDE can be expressed in the following general format:

$$A\phi_{xx} + B\phi_{xy} + C\phi_{xy} + D\phi_x + E\phi_y + F\phi + G = 0$$

This general equation can be classified according to the to the sign of the expression $B^2 - 4AC$:

Elliptic if	$B^2 - 4AC < 0$
Parabolic if	$B^2 - 4AC = 0$
Hyperbolic if	$B^2 - 4AC > 0$

For the TTM method, the Poisson equations are used to transform the mesh from the physical to the computational domain. In 3D the equations can be written as:

$$\xi_{xx} + \xi_{yy} + \xi_{zz} = P(\xi,\eta,\zeta)$$
$$\eta_{xx} + \eta_{yy} + \eta_{zz} = Q(\xi,\eta,\zeta)$$
$$\zeta_{xx} + \zeta_{yy} + \zeta_{zz} = R(\xi,\eta,\zeta)$$

Equation 2-37

where P,Q and R are control functions

2.3.4.3 Unstructured Grid Generators

Algebraic and PDE grid generators create meshes where the cells in the physical domain remain adjacent to the same cells after transformation. This structured nature can make it difficult to generate very complex grids, or to create adaptive meshes. To overcome these deficiencies, a variety of techniques which are collectively known as unstructured grid-generation methods, have been developed. When these are employed, the grid points can be placed anywhere in the domain. The vertices are then connected to one another by a process known as tessilisation, resulting in tetrahedral-shaped cells when working in three dimensions.

2.3.4.4 Adaptive Grid Generators

The accuracy of any computational solution can be undermined by the ability of the grid to capture the flow field sufficiently. When the engineer generates a grid, he/she has to anticipate regions of high flow gradient and cluster the grid cells accordingly. As prior knowledge of the flow field is not always known, the concentration of cells might not be in the correct position to capture adverse gradients. Adaptive grids have the ability to ensure this by relocating the cells during the solution process. These mesh generators look at intermediate results and place cells where gradients are present. The grid does not stabilise until the steady state solution has been reached.

2.3.5 Numerical Solution Methods

The numerical solution of the governing fluid-flow equations can be divided into either space-marching, or time-marching schemes. Space-marching methods are designed to solve the governing equations where the flow is inherently stable. Time-marching methods on the other hand, can be used to obtain solutions for both steady state and time-dependent problems. Time-marching routines approach their steady-state solutions asymptotically with increasing time.

Marching schemes are solved by using either explicit or implicit algorithms. Explicit procedures are easy to implement and require the minimum number of operations per cycle. Although the implicit approach is more complicated, it is inherently more stable, meaning that larger time steps can be taken. These methods are therefore computationally more efficient.

Numerical methods which discretesize the differential form of the Navier-Stokes equations, are generally called finite-difference (FD) methods. Finite-volume (FV) methods are based on the integral form of these equations. The major differences between these two approaches are listed in Table 2-1.

Finite Difference	Finite Volume
$\frac{df(x)}{dx} \approx \frac{f(x+\Delta x)-f(x)}{\Delta x}$	$\iint_S f dS \approx \sum_s f \Delta S$
• Taylor series expansion is used. • Δx is finite leading to a truncation error. • The power of Δx determines the order and therefore the accuracy of the method. • Variables are evaluated at the cell corners. • Special one-sided boundary treatment is required, as a row of grid points lie on the boundary. Once convergence is obtained the solution is known at the surface. • Free stream capture only through difference approximations.	• S is the surface surrounding the volume V. • For non-uniform, non-orthogonal meshes care has to be taken when calculating cell face areas and cell volumes. • Variables are evaluated at the cell faces. • Slave cells are needed at all boundaries. These facilitate tangency and no slip implementation, but the solution at the surface has to be obtained through averaging. • Automatic free-stream capture capability.

Table 2-1 Differences between Finite-Difference and Finite-Volume Methods

From Table 2-1 it can be seen that finite differences can be grouped according to the order of the method that is used. Examples are shown in Figure 2-5 by graphically depicting a backward and a forward first-order case as well as a second order central-difference method.

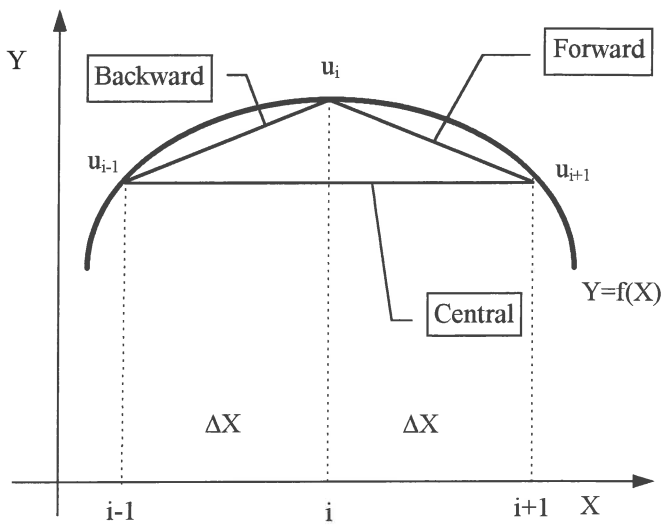

Figure 2-5 Difference schemes[28]

The numerical implementation of PDE's result in two classes of round-off errors. The first is a property of the numerical processor, while the second is caused by the discretising process. As the second inaccuracy is a function of the numerical scheme, the algorithm should be designed to ensure bounded and consistent errors. This means that the discrepancy should not only stay constant during the solution process, but also as the time and spatial increments tend towards zero, the difference equations should approach the differential equations. If these conditions are not met, numerical instability can result.

A few of the most likely causes of instability are listed below:

- Inconsistent spatial or temporal discretising schemes.
- Physically ill-posed boundary conditions.
- Inconsistent implementation of boundary conditions.
- Time or space-step size violations.
- Mismatch between the numerical and physical domains of dependence.
- The attempt to capture strong gradients with too little dissipation.

Numerical schemes have therefore not only been developed to increase numerical efficiency, but also to limit potential instabilities.

Two numerical schemes are used in this study. The first is Roe's Flux difference splitting method for the Full Reynolds Averaged Navier-Stokes Equations (RANS) and the second is the Semi-Implicit Method for Pressure Linked Equations (SIMPLE) by Patankar and Spalding[30].

2.3.5.1 Roe's Flux difference vector splitting for the Euler Equations

The 2D Euler equations, which are an inviscid case of the RANS equations, will be used to explain the concept of Roe's flux-vector splitting method as the physical interpretation of his method is not entirely clear[2]. These equations are used as they are less complicated but still representative of the full 3D RANS equations. The discussion of the Euler equations is conceptual in nature and the reader is referred to Appendix B for a detailed presentation of Roe's algorithm for the full RANS.

The 2D Euler equations, after transformation to the rectangular computational domain, can be written as[v]

$$\tilde{U}_t + \tilde{F}_\xi + \tilde{G}_\eta = 0$$
$$\tilde{U} = UV = UJ^{-1}$$

Equation 2-38

where V is the volume of the finite volume element and corresponds to the inverse of the Jacobian of the transformation. This system of equations possesses the following properties[32]:

- The flux vectors are homogeneous
- The eigenvalues are real and can be positive and negative
- The sign of the eigenvlaues indicates the direction of data propagation
- The flux Jacobian matrices can be diagonalised

The algorithm used to explain Roe's procedure is given by

$$U_{i,j}^{n+1} = U_{i,j}^{n} + \Delta U_{i,j}^{n}$$

where

$$\Delta U_{i,j}^{n} = -\frac{\Delta t}{V_{i,j}}\left(\frac{D_+}{\Delta\xi}\tilde{F}_- + \frac{D_-}{\Delta\xi}\tilde{F}_+ + \frac{D_+}{\Delta\eta}\tilde{G}_- + \frac{D_-}{\Delta\eta}\tilde{G}_+\right)_{i,j}^{n}$$

Equation 2-39

[v] Duplicated from: Craig, K., "Computational Fluid Dynamics - MBV780", Lecture Notes, University of Pretoria, 1994. Courtesy of Professor K.J. Craig.

The plus and minus fluxes in the ξ and η directions are determined using a suitable flux-splitting method. The two methods that are presented below are explained with the aid of Figure 2-6.

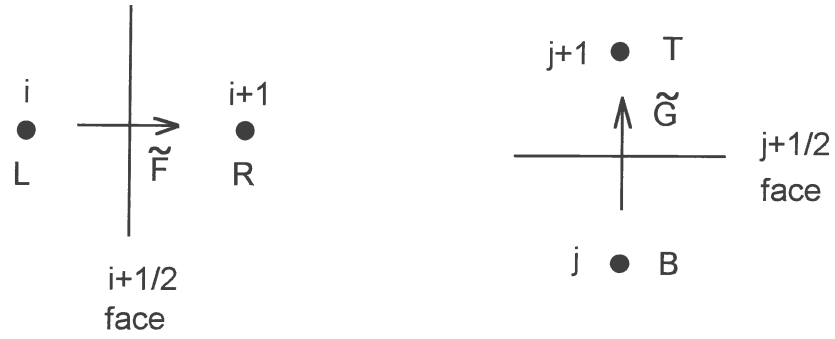

Figure 2-6 Fluxes through the cell face in two directions

Steger and Warming[28] formulated the characteristics theory[vi] into conservation law form by splitting the flux vector according to the sign of the characteristic speeds of the flow. By considering the schematic illustration in Figure 2-6 their methods of flux splitting in two directions can be written as[42]:

$$\tilde{F}_{i+\frac{1}{2}} = \frac{\tilde{F}_R + \tilde{F}_L}{2} - \tfrac{1}{2}\left[\left(A_+ - A_-\right)U_R - \left(A_+ - A_-\right)U_L\right]$$
$$\tilde{G}_{i+\frac{1}{2}} = \frac{\tilde{G}_T + \tilde{G}_B}{2} - \tfrac{1}{2}\left[\left(B_+ - B_-\right)U_T - \left(B_+ - B_-\right)U_B\right]$$

Equation 2-40

where A and B are the Jacobians of the transformation.

Roe[2] found that by using the geometric averages to split the fluxes, the characteristics could be captured in the right direction. Referring back to Figure 2-6, Roe's approach can be written as[42]

$$\tilde{F}_{i+\frac{1}{2}} = \frac{\tilde{F}_R + \tilde{F}_L}{2} - \tfrac{1}{2}\left(\hat{A}_+ - \hat{A}_-\right)\left(U_R - U_L\right)$$
$$\tilde{G}_{j+\frac{1}{2}} = \frac{\tilde{G}_T + \tilde{G}_B}{2} - \tfrac{1}{2}\left(\hat{B}_+ - \hat{B}_-\right)\left(U_T - U_B\right)$$

Equation 2-41

where $\hat{A}$ and $\hat{B}$ are the *geometric-averaged* Jacobians. Sudden changes, such as shock waves could be apprehended with higher accuracy by using Roe's method.

Using Roe's method the algorithm of Equation 2-39 can now be re-written as

$$\Delta U_{i,j}^n = -\frac{\Delta t}{V_{i,j}}\left(\frac{\tilde{F}_{i+\frac{1}{2}} - \tilde{F}_{i-\frac{1}{2}}}{\Delta\xi} + \frac{\tilde{G}_{j+\frac{1}{2}} - \tilde{G}_{j-\frac{1}{2}}}{\Delta\eta}\right)_{i,j}^n$$

Equation 2-42

[vi] The precise treatment of the flow physics in the classic method of characteristics was lost during the early development of finite difference procedures.

Splitting of the Jacobians $\hat{A} = \hat{A}_+ + \hat{A}_-$ and $\hat{B} = \hat{B}_+ + \hat{B}_-$ of Equation 2-39 is through polarity of the rotated eigenvalues. As the eigenvalues and associated eigenvectors are real, the Jacobian can de diagonalised. A similarity transformation therefore exists such that[32]

$$\tilde{A}_\pm = X^{-1}\Lambda_{A_\pm}X$$
$$\tilde{B}_\pm = X^{-1}\Lambda_{B_\pm}X$$

where Λ is a diagonal matrix of eigenvalues of $\tilde{A}$ and X is the eigenvector matrix. Λ_+ and Λ_- are then determined by

$$\Lambda_+ = \frac{\Lambda + |\Lambda|}{2}$$
$$\Lambda_- = \frac{\Lambda - |\Lambda|}{2}$$

Equation 2-43

As stated previously, a thorough presentation of Roe's algorithm for the full 3D RANS equations is given in Appendix B. In the appendix it is seen that the algorithm is expanded from a thin-layer procedure. This was necessary as the underlying assumptions of the thin-layer approach does not necessarily hold for bluff bodies such as racing cars where large areas of separation is expected. The thin-layer equations are designed for flow cases where the normal-stress gradients are much larger than the streamwise and circumferential gradients[32].

2.3.5.2 The SIMPLE Algorithm of Patankar and Spalding

The incompressible SIMPLE algorithm for the Navier-Stokes equations is used as implemented by STAR-CD[3]. This procedure is based on a cyclic series of guess-and-correct operations. The velocity components are first calculated from the momentum equations using a guessed pressure field. The pressures and velocities are then corrected so as to satisfy continuity. This process continues until the solution converges. The outline of the SIMPLE algorithm is presented below:

- Guess the pressure at each grid point.
- Solve the momentum equations to find the intermediate velocity components.
- Solve the pressure-correction equation at each grid point to find the pressure corrections.
- Correct the pressures and velocities.
- Replace the intermediate values of pressure and velocity with the new corrected values.
- Repeat the process until the solution converges.

2.3.6 Boundary Conditions

In order to obtain a unique solution for the PDE's, two supplementary sets of conditions are required. These are initial and boundary conditions.

Initial conditions are required as the dependent variables have to be specified at some initial state. Fortunately, the dependency on these vanish as the solution evolves through time.

The relations which the dependent variables or their derivatives have to satisfy on the edge of the solution domain, are known as boundary conditions. As the outcome of the solution depends on these, it is vital to specify them correctly.

Boundary conditions can be specified in two basic ways:

- Dirichlet boundary conditions, where the dependent variables are specified along the boundary.
- Neumann boundary conditions, where the normal gradients of the dependent variables are specified.

Irrespective of the way in which boundary conditions are specified, they still have to relate to the physical-flow problem. They can therefore, depending on the prevailing flow conditions, be classed as follows:

- Impermeable boundaries
- Axis of symmetry or reflection boundaries
- Entrance and Exit boundaries

2.3.6.1 Impermeable Boundaries

Impermeable boundaries are constraints such as solid walls, where the flow has to be parallel to the surface. For viscous flow fields, velocity, pressure and wall temperature relations can be as specified in Table 2-2.

Independent Variable	Boundary Condition	Boundary Type	Description
Pressure	$\frac{\partial p}{\partial n} = 0$	Neumann	One-dimensional boundary layer equation
Velocity	$u_n = 0$ $u = 0$	Dirichlet	Tangential flow No-slip condition
Temperature	$k\frac{\partial T}{\partial n} = 0$ $T = T_{wall}$	Dirichlet Neumann	Adiabatic wall Isothermal wall

Table 2-2 Impermeable Boundary Conditions

where n is the unit normal vector.

A study by Mazeheri and Roe[43] suggested that the assumptions of one-dimensional analysis and waves travelling normal to the boundaries, were not always valid. They also showed that the path to convergence followed a simple, well-defined repetitive pattern, characterised by the passage of acoustic waves travelling between the inner and outer boundaries. Based on this, they proposed a new "soft wall" boundary condition. In the steady limit this coincided with the conventional rigid boundary. They found that by replacing

$$u_n = 0$$

with

$$\frac{\partial u_n}{\partial t} + \frac{u_n}{\tau} = \frac{\mu}{\rho a}\frac{\partial p}{\partial t}$$

Equation 2-44

the computational time could be reduced by more than 40% for two-dimensional lifting flows. μ and τ are parameters chosen to minimise reflected disturbances. They found that μ could roughly be set equal to 0.5 as the precise value was not very critical. The parameter τ is loosely bounded by two considerations. The value should not be much less than the time step and secondly, it should not be greater than the time for a round trip of a wave crossing the domain.

2.3.6.2 Reflection Boundaries

Reflection boundaries can be employed when the flow is symmetrical around a line or plane. By definition this means that the flow cannot cross, or have any gradient normal to the boundary. These conditions are expressed in mathematical form in Table 2-3.

Independent Variable	Boundary Condition	Boundary Type
Pressure	$\frac{\partial p}{\partial n} = 0$	Neumann
Velocity	$\frac{\partial u}{\partial n} = 0;$ $u_n = 0$	Neumann Dirichlet
Temperature	$\frac{\partial T}{\partial n} = 0$	Neumann

Table 2-3 - Reflection Boundary Conditions

2.3.6.3 Entrance and Exit Boundaries

To implement the inflow and outflow boundary conditions for the RANS equations, it is easiest to look at the hyperbolic Euler equations. For this type of PDE, there has to be as many boundary conditions as there are characteristics. The two-dimensional characteristics for the Euler equations are plotted Figure 2-7[v] for a subsonic flow field.

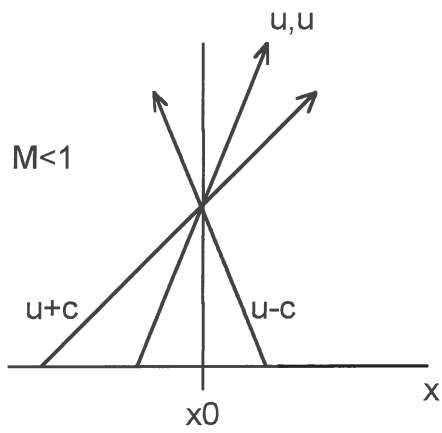

Figure 2-7 - Characteristics for the subsonic Euler equations

Characteristics originating from inside the domain do not have to be specified, as they can be resolved internally. From Figure 2-7, it can be seen that three external boundary conditions have to be specified at the inlet, while only one is required at the outlet. These values are normally determined through measurement. For external-flow problems, the computational

domain can be constructed so that the inlet and outlet boundaries are far from the body under investigation. Free-stream conditions can then be specified.

Wei Shyy[44] developed an alternative approach to specifying the unknown outlet condition. He calculated the remaining characteristic using a line-iterative method based on the internal flow. He proposed four different methods:

Explicit step space extrapolation - $\Phi_{NI,j}^{n+1} = \Phi_{NI-1,j}^{n}$ **Equation 2-45**

Implicit step space extrapolation - $\Phi_{NI,j}^{n+1} = \Phi_{NI-1,j}^{n+1}$ **Equation 2-46**

Explicit linear space extrapolation - $\Phi_{NI,j}^{n+1} = 2\Phi_{NI-1,j}^{n} - \Phi_{NI-2,j}^{n}$ **Equation 2-47**

Implicit linear space extrapolation - $\Phi_{NI,j}^{n+1} = 2\Phi_{NI-1,j}^{n+1} - \Phi_{NI-2,j}^{n+1}$ **Equation 2-48**

where NI refers to the right-boundary index, n is the iteration number, and Φ is any of the flow variables.

He found the implicit linear space extrapolation method to be computationally expensive, and that a unique solution was not always guaranteed. The explicit zero-order space extrapolation method appeared to be the most robust.

2.4 Conclusion

The aim of this chapter was to introduce the reader to experimental and numerical road-car aerodynamics and to present him/her with a summary of the existing techniques and methods, as found in the literature.

As an introduction, the history of racing cars, with the accent on aerodynamics, was briefly presented. This was followed by a more detailed discussion on the aerodynamics of modern racing cars. General Aerodynamic relationships and experimental techniques followed. Finally, the computational techniques were outlined.

Some of the techniques and methods included in this chapter were presented in "lay-man" terms. Certain elements had no direct bearing on the case studies presented in later chapters. The author felt that at the time of his introduction to race-car aerodynamics a few years ago, there was a large gap between "beginner" literature and highly technical reports. The author also found the information segregated and very few books, such as the one by Katz[8], were dedicated exclusively to race-car aerodynamics. It was therefore the intention to present a more complete picture in moderately technical terms, than was required for this study.