CHAPTER 2: Literature Study

2.1 Introduction

This section of the dissertation provides an overview of some of the major fields of interest either utilised or examined in this study. This chapter discusses some of the major publications made in the fields of interest, as well as contextualising them in terms of this study.

2.2 Sloshing

Sloshing is a phenomenon explained in the dictionary as "spill or splash copiously or clumsily, make a splashing sound". The event however has far reaching implications in the realms of engineering. Sloshing can be described as the motion of a fluid as it attempts to attain a state of equilibrium for the effective instantaneous acceleration (gravitational, translational, etc.) felt by the fluid. The momentum of the fluid and external loads on the fluid container will prevent this state of equilibrium. In turn, the motion of the fluid has a number of side effects. These include the acoustical affects due to the higher velocity motion and pressure fluctuations near the free surface, the impulsive loads the fluid may exert on the container or other structural bodies inside the container, and an effect on the dynamic stability of the container as a whole.

The phenomenon of sloshing has evoked the attention of scientific researchers since as early as the 1950s with the work of Graham and Rodriguez [1, 2] who studied the effects of fuel motion on airplane dynamics. In the 1960s the work was continued by the aerospace industry with work similar to that performed by Abramson of NASA [3]. Since then numerous other publications have been made, with wide-reaching applications [4-8, 14-18] (See section 2.4). The major fields of interest within the context of sloshing include fuel tanks of almost all vehicle types, liquid transport containers, seismic and wind induced oscillations of tall buildings, aerospace, maritime, liquefied natural gas (LNG) tanks, etc. More recently the source of many of these investigations is the study of numerical techniques for solving the physics

involved, in a hope that such techniques could assist in the design process of liquid containers.

Sloshing is often analysed in a simpler form where no overturn takes place, that is to say when the free surface stays intact. Assumptions like incompressibility, irrotational flow, inviscid, no ambient velocity, two-dimensional, and small amplitudes allow for a simplified analysis via linear wave theory [9]. The theory was first proposed by Lagrange (1776) and Airy (1845) and provides some insight into the behaviour of waves. The solution is then further subdivided into shallow and deep waves. Figure 2.1 below shows the general form of a wave, where C is the wave speed, h is the fluid level, L is the wave length, and η is the free-surface deviation from the mean water line.

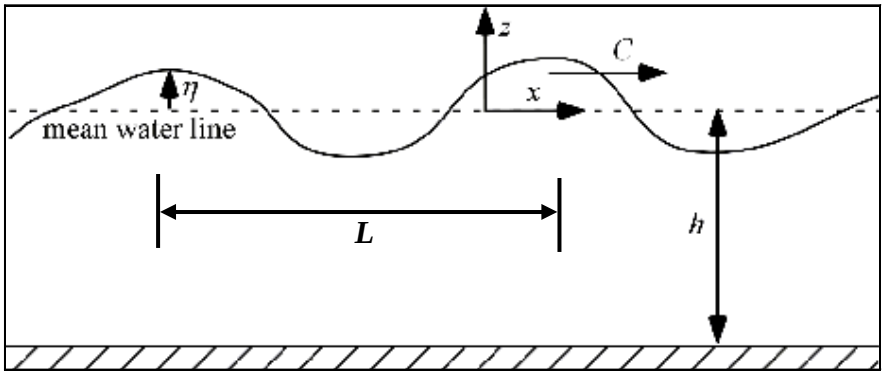

Figure 2.1: General wave form [9]

Mathematically, the difference between shallow and deep waves manifests as a wave-speed dependence on the wave number (k). For deep-wave theory, the wave speed is dependent on the wave number ($h/L > 0.5$) and the motion is dispersive. Conversely, when the wave speed displays independence from the wave number, usually for $h/L < 0.05$, we refer to this as shallow wave theory (non-dispersive). The wave number is simply the reciprocal of the wave length (L). Equation 2.1 below shows the actual relation for deep waves, while equation 2.2 shows that which applies to shallow waves for the wave speed.

$$C = \sqrt{\frac{g}{k}} \tag{2.1}$$

$$C = \sqrt{gh} \tag{2.2}$$

Visually one can see the difference between the behaviour of particles in "deep water" in Figure 2.2 below, a picture of particle trajectories in plane periodic water waves. The waves are travelling across the image with no reflection. The particle trajectories can be seen to be circular at the free surface and flattened toward the bottom. The particles midway between the free surface and the bottom are intuitively more elliptical in shape as the physical constraint of the bottom has more effect. Section 2.4 provides more detail on linear wave theory.

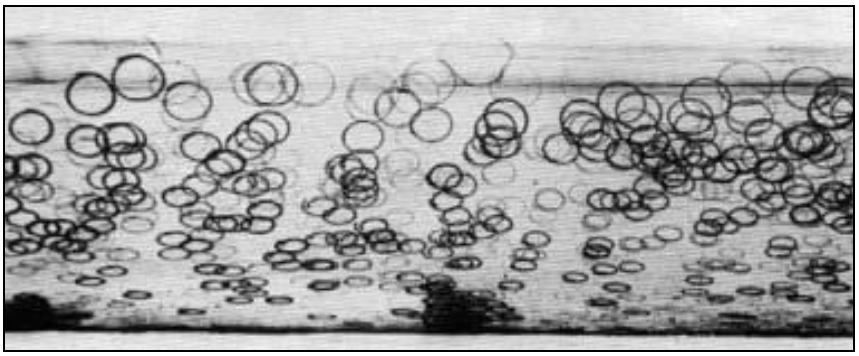

Figure 2.2: Particle Trajectories in plane periodic water waves [10].

Reflected waves on the other hand behave somewhat differently. Figure 2.3 below shows the particle trajectories for a pure standing wave, 100% reflection. The trajectories now show the streamlines for pure standing waves. This is quite similar in behaviour to the oscillation of a liquid in a tank with no baffles.

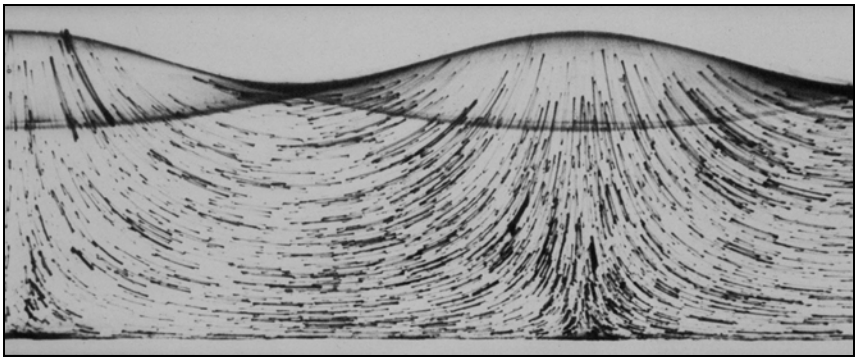

Figure 2.3: Particle trajectories in pure standing waves [10]

Often in cases of sloshing in containers the behaviour of the waves is somewhat non-linear. Figure 2.4 below (extracted from the experiments in chapter 3) shows a clear illustration of such an instance. Non-linear wave behaviour with turnover is a complex phenomenon and is not easily represented mathematically, but section 2.4 provides a further discussion of some methods used. This behaviour is typical of sloshing of fuel in fuel tanks, and other containers that may experience similar accelerations. The acoustical affects felt during sloshing are also associated with this level of fluid motion, something which is of interest to vehicle manufacturers due to the undesirability of driver-experienced noise levels. This forms part of the field of Noise, Vibration and Harshness (NVH), and an ongoing effort to make cars run quieter and more smoothly.

Figure 2.4: Non-linear sloshing of water in a rectangular tank

2.3 Fluid-Structure Interaction

Fluid-Structure Interaction (FSI) is a field which refers to the coupling of unsteady fluid flow and structural deformation. This is ideally a two-way coupling of pressure and deflection. A number of interesting applications have stimulated research in this field. Typical fields of interest would be airbag modelling, fuel tank sloshing, heart valve modelling, helicopter crash landings, etc. The important point, and the reason for the initiation of the FSI field of study, is that the fluid mechanics may affect and be affected by the structural mechanics, and vice versa. We might consider the case of an aeroplane wing. Once air starts to pass over a wing it typically produces some amount of lift, this will cause the wing to deflect and in turn alter the properties of the wing and the amount of lift it produces. The cycle may be continuous (e.g. flutter) or it may settle to some stable point of equilibrium. Other instances of FSI might be where the stress in a structural component may be induced by the fluid that surrounds it. Clearly in this case the coupling of the fluid's pressures and the motion of the structure is essential. Figure 2.5 below shows images of both the physical testing of airbags with crash test dummies as well as the simulations of the same event in a FSI-capable software package. One can appreciate that the cost effectiveness of being able to simulate the event with computer software would evoke much interest from vehicle manufacturers.

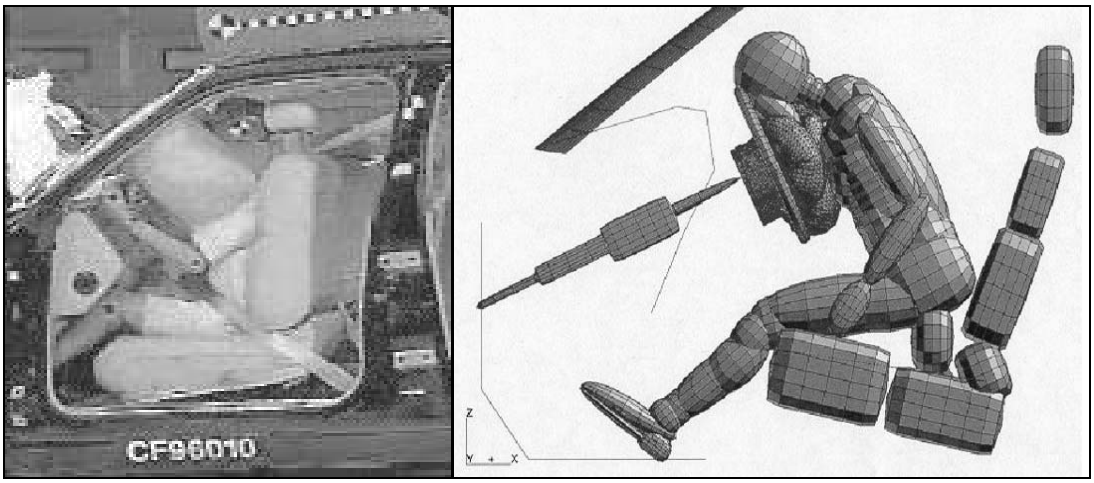

Figure 2.5: Airbag/Crash testing and simulation [11,12]

Section 2.4.4 gives a brief description of the mathematical methods involved in FSI and the suitability of certain methods.

2.4 Mathematical modelling

Mathematical modelling in general is the representation of some physical event through mathematical equations. Due to the fact that most interesting physical events are very difficult to model exactly, i.e., that the equations provide an exact representation of the reality, many possible mathematical representations may exist for any given event. In general, this implies a trade-off between equation complexity and accuracy. However, innovations in a field where mathematical modelling is concerned may lead to simpler equations giving similar or even improved results. This trend is usually initiated when more complex governing equations are simplified for very specialised cases of the broader physical event type. The use of specialised turbulence models for external flow over a wing is one such example.

The motivation for the use of mathematical models stems from a need to understand trends without having to reconstruct the event. Furthermore, if we have reasonable confidence in the accuracy of the model, we may use it as a method with which to analyse the event itself. This may in turn form part of a design cycle and, more applicably, part of a mathematical optimisation cycle (Mathematical optimisation is further discussed in section 2.6). One must however always keep in mind that the success of these techniques is strongly dependent on the accuracy of the model.

The remainder of this section introduces some of the mathematical techniques that have traditionally been used in the modelling of sloshing events or events that are similar in physics to sloshing.

2.4.1 Linear Wave Theory

Linear wave theory, or Airy's wave theory, is one of the first types of mathematical modelling used to analyse wave motion [9]. Briefly introduced in section 2.2 it provides some insight into wave motion at a relatively simple level. The mathematical theory is based on the governing equation of continuity (equation 2.3) in two dimensions and potential flow assumptions. The velocity potential ϕ is defined as in equation 2.4 - 2.5, and the combined result gives the Laplacian form of the continuity equation (equation 2.6), for which well-established standard solutions exist. Again, all equations are defined within the coordinate system described in Figure 2.1 above.

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \tag{2.3}$$

$$u(x,z,t) = \frac{\partial \phi(x,z,t)}{\partial x} \tag{2.4}$$

$$w(x,z,t) = \frac{\partial \phi(x,z,t)}{\partial z} \tag{2.5}$$

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial z^2} = 0 \tag{2.6}$$

The simple case of linear pure progressive wave motion lends itself to a solution by separation of variables. Equation 2.7 below represents the assumed sinusoidal form of the free-surface position $\eta(x,t)$. After all boundary conditions are simplified and applied under the assumption that the wave amplitude ε is small, i.e., all $d\phi/dx$ product terms are neglected, we obtain a non-dimensional solution as in equation 2.8 below. In this equation C is the wave speed, ε is the amplitude, L is the wavelength and h is the mean free-surface level. Equation 2.7 is valid for $\varepsilon << h$ and $\varepsilon << L$.

$$\eta(x,t) = \varepsilon.\sin\frac{2\pi}{L}(x - Ct) \tag{2.7}$$

$$\frac{C^2}{gh} = \frac{L}{2\pi h}\tanh\frac{2\pi h}{L} \tag{2.8}$$

As discussed in section 2.2, two possible assumptions can be made at this point. The two possibilities are either $h<<L$ or only $L<<h$, which intuitively would describe a wave in shallow water and deep water respectively. If we consider the shallow water assumption, then $2\pi h/L$ in equation 2.8 will become small and the *tanh* term will approach the value of this ratio. We then obtain the relation in equation 2.9. The deep water assumption will conversely result in the *tanh* term approaching unity, which will result in the relation shown in equation 2.10.

$$\frac{C^2}{gh} = 1 \tag{2.9}$$

$$\frac{C^2}{gh} = \frac{L}{2\pi h} \tag{2.10}$$

Graphically, the difference can be seen in Figure 2.6 below, which shows the relationship between wave propagation speed and wave length in non-dimensional terms.

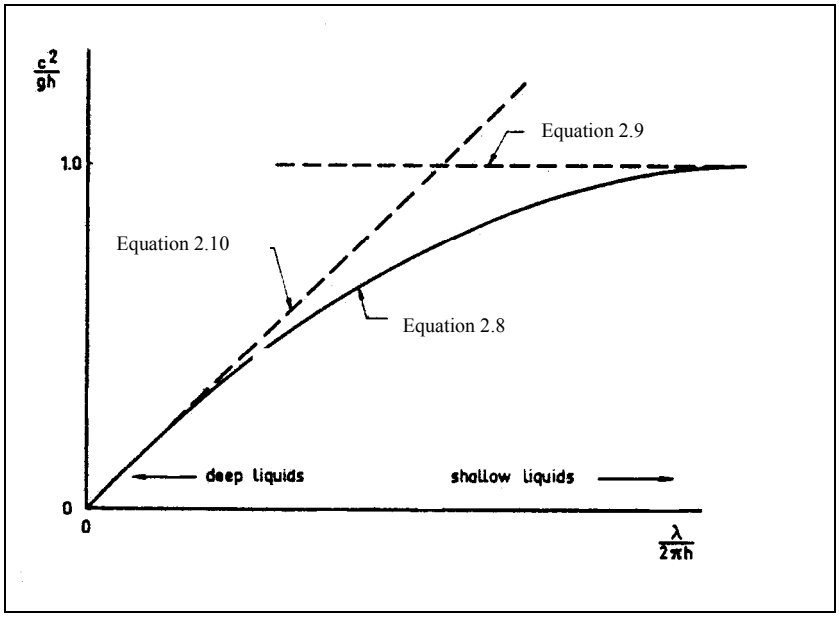

Figure 2.6: Relation between propagation speed and wavelength [13]

However, some assumptions made early in this formulation result in inaccuracies in these relations. For example, very closely packed waves (short wavelength) exhibit high propagation speeds and as a result are quickly dissipated. Figure 2.7 below shows a more realistic relation near this limit where the waves are termed 'capillary' waves.

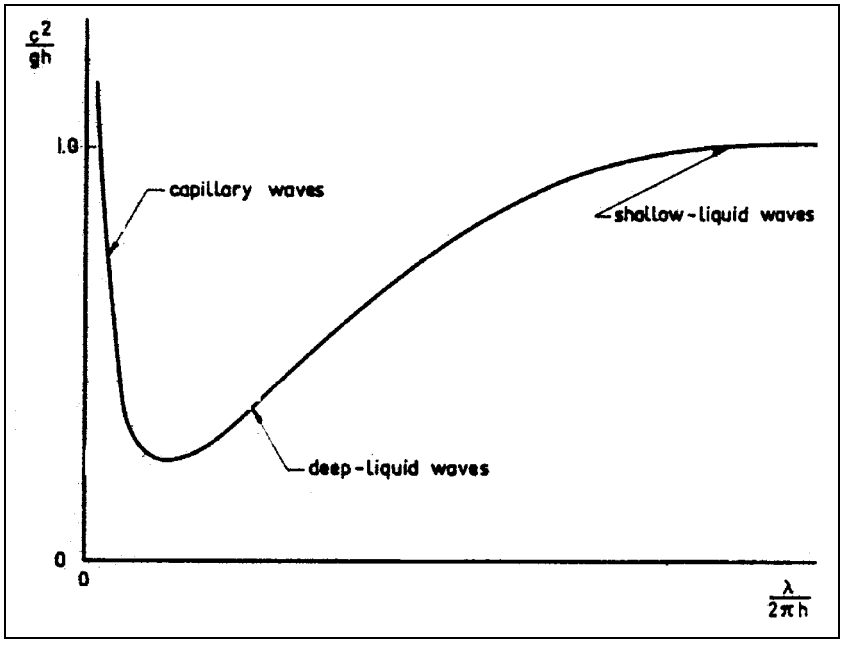

Figure 2.7: Dissipation effect of short wavelengths on wave propagation speed [13]

Linear wave theory for a 3-dimensional liquid container yields equation 2.11 below which represents the n^{th}-mode oscillation frequency 'ω_n' in a container of length 'a' and fluid height 'h' [13]. Note the independence on tank width at this stage.

$$\omega_n^2 = \frac{n\pi g}{a}\tanh\left(\frac{n\pi h}{a}\right) \tag{2.11}$$

The linear wave theory equations described above have been expanded in several studies, e.g., Warnitchai and Pinhaew [14] expanded the theory in an attempt to develop an equation for the 1st mode frequency in a container whilst considering flow damping devices or Tuned Liquid Dampers (TLD). The study yielded equation 2.12 below, which is an adaptation of equation 2.11 above for n=1 with specific consideration of the damping effect of a single rigid cylinder standing vertically in the centre of the tank. C is the coefficient of inertia of the cylinder, A is the wave amplitude, and b is the width of the 3-D tank.

$$\omega_1^2 = \frac{\pi g}{a}\tanh\left(\frac{\pi h}{a}\right)\left(1+\frac{A}{ab}C\left(1+\frac{2\pi\frac{h}{a}}{\sinh\left(2\pi\frac{h}{a}\right)}\right)\right)^{-1} \tag{2.12}$$

The effect is however small due to the relatively small magnitude of the A/ab term. Equation 2.11 is however still of interest when analysing sloshing in a tank without damping devices. Modi and Munshi [15] examined in detail the effect of liquid depth and modal frequency on their unique dampers, making the point that static tuned liquid dampers only operate optimally at a single depth and wave speed. Modal analysis received further attention by Schotte and Ohayon [16] with their demonstration of the apparent gravity approach for accelerating bodies.

Although linear solutions can work well for simple cases, they are limiting. Other solutions similar to the one described exist for several very simple flow problems, but simple cases like this do not represent the majority of engineering problems. In the 1970s, Faltinson [17] extended the field to approximate theoretical forms for inviscid sloshing in moving tanks. More recently, Frandsen [18] presented the combination of potential flow theory with physical-to-computational transformation techniques to simplify the tracking of the free surface. However, alternative methods of solving fluid sloshing need to be utilised in most cases that are not highly simplified. This necessity has led to the development of methods like Computational Fluid Dynamics (CFD), as discussed in section 2.4.3.

2.4.2 Equivalent Mechanical Systems

Equivalent Mechanical Systems (EMS) are mechanical systems that respond in the same way as the original/actual systems would. Sometimes referred to as lumped parameter models (LPM), these systems would typically consist of a series of masses, springs, and dampers that when excited will exhibit similar feedback forces. These models are a legacy of a time when computational power was restricted to a pen and paper, and are often restricted by small angle assumptions. Models of this nature do however still exist [1,2,4] and may still give some level of insight into a problem. It must also be said that even today, computational power can still become a factor when considering the cost of solving complex flow or structural problems.

A very simple mechanical model was introduced in 1951 by Graham [2] who represented the fuel in the tank as an equivalent pendulum. This analogy has more recently been investigated by Fernando [4] where some experimental work was performed to determine the spectrum of validity. Figure 2.8 shows a schematic representation of the pendulum analogy.

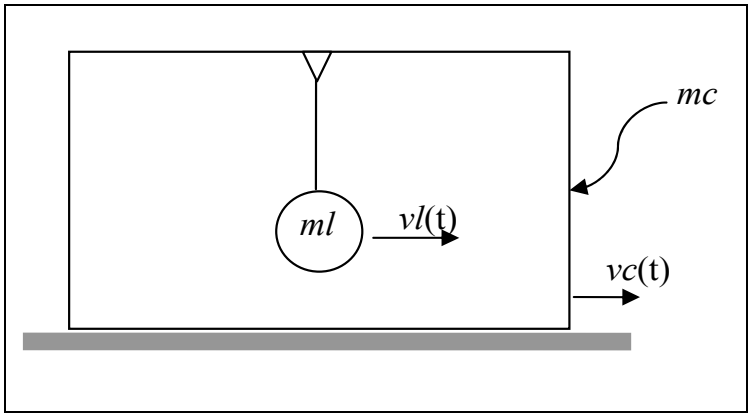

Figure 2.8: Schematic of pendulum analogy

The mass of the liquid in the container is represented by the mass of the pendulum (*ml*), while the pendulum's velocity (v*l*) represents the velocity of the liquids centre of mass. The container retains its mass (*mc*) and velocity (v*c*). Fernando's study [4] centres on the motion of the container relative to the surface it is resting on. The motion is influenced by the behaviour of the fluid in the container. By comparing experimental data with the analytical model, he shows the regimes within which the model exhibits acceptable performance. As one would expect, the model's spectrum of validity is limited to smaller excitation forces, since non-linear free-surface behaviour cannot be accounted for by a pendulum. The model does however provide some interesting insight into level of sloshing as a function of the fluid level. Continuing with the analogy as introduced above, Fernando considers the frictional force (F_f), and the force on the container due to the liquid sloshing (F_c). Figure 2.9 below illustrates the variation in forces F_f and F_c with respect to the mass of fluid in the container m_l. F_{f1} and F_{f2} represent two frictional force curves corresponding to the two container masses (mc_1 and mc_2). For the case represented by the F_{f1} curve, the tank will oscillate over the interval $m_{l1} < m_l < m_{l2}$, i.e., when $F_c > F_f$. When $F_c < F_f$, as at points A and D, there will be no relative motion between the container and the surface. The shaded area indicates where motion will occur for all container masses greater than mc_1. Some of the major observations that can be made include the non-linear behaviour of F_c. This can be explained by the fact that not all the fluid is involved in the oscillation. For low fill levels, fluid mass below point B, the force F_c is minimal due to the lack of fluid in the tank, i.e., inertial effects are low. For high fill levels, fluid masses above point C, insufficient room (gas) is in fact available in the

tank for complete oscillation, and when the container is full (m_{l2}) no oscillation takes place at all. A further observation is that the point of zero gradient of $Fc(ml)$ point E, or maximum Fc is in fact not at the 50% fill point. The data in fact suggests a value between 65 to 70%. This value will prove important at a later point in this study.

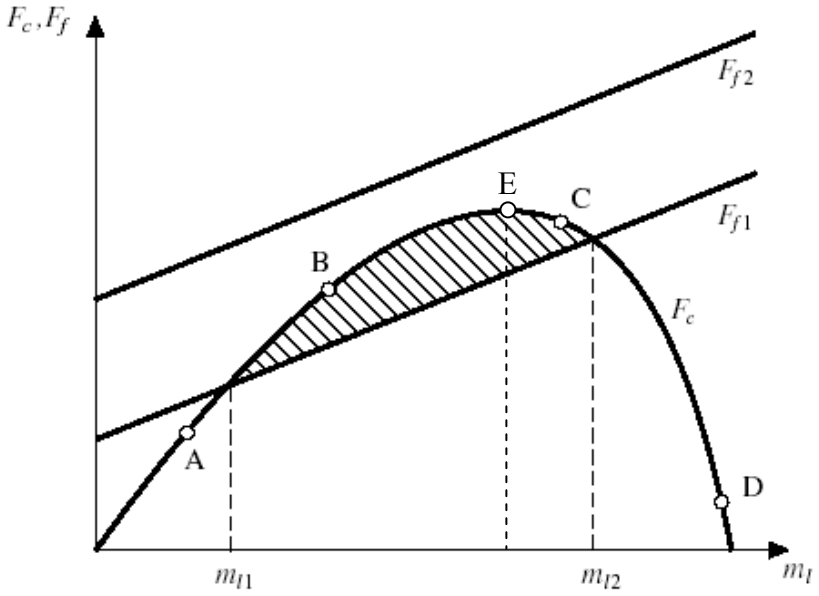

Figure 2.9: The relation between Fc/Ff and ml [4]

A more complex LPM was introduced by Graham and Rodriguez [1] in 1952. The model was developed in an attempt to analyse the forces induced on an aeroplane by the motion of the fuel in the fuel tanks during flight. Figure 2.10 below shows a schematic of this LPM. The various masses are used to model the various wave harmonics.

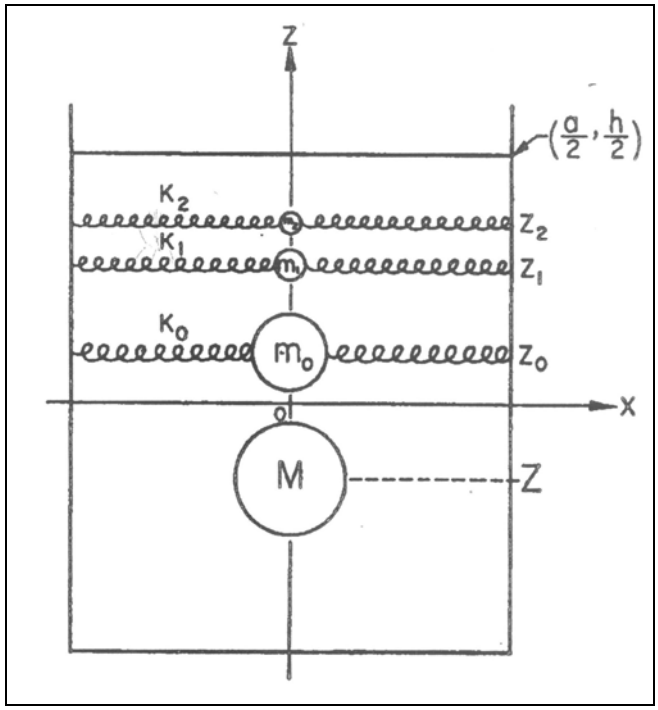

Figure 2.10: Graham and Rodriguez's LPM [1]

Graham and Rodriguez's model is restricted to small accelerations compared to gravity and only very small angular displacements. In fact, due to the application, the model is restricted to three basic tank motions; x-axis translation, pitching in the x-z plain, and yaw about the z-axis (figure 2.10). However, even when considering significant simplifications, the equations that represent the model are very involved and the author suggests that a model that would respond to all tank motions "cannot" be constructed. One major drawback of the system is that no damping exists, however the location of natural frequencies can be extracted. Figure 2.11 [1] below shows a force ratio (force produced by the mechanical system due to a horizontal oscillation divided by the force that would be produced if the fuel were a solid) to dimensionless frequency for a 0.25 aspect ratio tank. This is the typical format of the data one would be able to extract from the model and clearly shows natural frequencies. This can ultimately be used in the design of the aeroplane's fuel tank to ensure that certain natural frequencies will not adversely interfere with the behaviour of the plane.

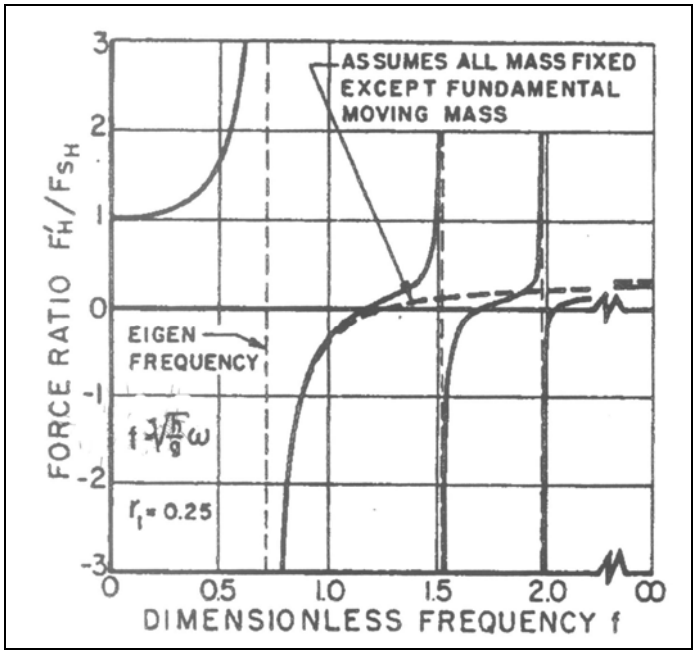

Figure 2.11: Force ratio vs. dimensionless frequency for Graham and Rodriguez's LPM [1]

2.4.3 Navier-Stokes Methods

Computational Fluid Dynamics (CFD) is the numerical solution of the discritised form of the partial differential equations that govern fluid flow (Navier-Stokes equations). The three principle governing equations or major conservation laws are momentum, mass, and energy conservation. CFD is now a widely used tool for solving and visualising fluid flow processes in the engineering world. A number of CFD codes are commercially available and for the purposes of the study, due to availability, Fluent v6.x, as distributed by Fluent Incorporated [19], is utilised. A brief explanation of the equations solved by CFD codes will follow.

2.4.3.1 Conservation of Mass

The mass conservation or continuity equation solved is as follows:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho \vec{V}) = S_m \tag{2.13}$$

where ρ is the density, t is time, and $\vec{V}$ the velocity vector. S_m is a mass source term, e.g., the mass added by one phase from another phase, e.g., during the vaporisation of liquid droplets. Equation 2.13 is a general form of the continuity equation and is valid for both compressible and incompressible flow.

2.4.3.2 Conservation of Momentum

The conservation of momentum in the i-th direction, assuming a non-accelerating reference frame, is given by:

$$\frac{\partial}{\partial t}(\rho V_i) + \frac{\partial}{\partial x_i}(\rho V_i V_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + \rho g_i + F_i \tag{2.14}$$

where τ_{ij} is the stress tensor given by equation 2.15, g is the component of gravitational acceleration, p is the static pressure, and F_i is an external body force and behaves much like the gravitational body force, but allows for user-defined source terms, e.g., momentum source. Momentum source has the units of kg/m²s² and is the multiple of the density of a specific mesh cell and the instantaneous acceleration. This term will prove very useful in the implementation of load curves in this study. The stress tensor is as follows:

$$\tau_{ij} = [\mu(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i})] - \frac{2}{3}\mu\frac{\partial V_l}{\partial x_l}\delta_{ij} \quad (2.15)$$

where μ is the molecular viscosity and the term after the negative sign is an effect due to volume dilation (size change). The Kronecker delta function δ_{ij} represents a value of zero if $i \neq j$ or a value of one if $i=j$. The momentum conservation equations are also known as the Navier-Stokes equations.

To take into account the effect of turbulence fluctuations, it is customary to perform a time averaging of equation 2.14, resulting in the Reynolds-Averaged Navier-Stokes equations. The equations contain turbulent stress terms that are modelled (See section 2.4.3.4 below).

2.4.3.3 Conservation of Energy Equation

The energy conservation is solved in terms of conservation of enthalpy, h, given by:

$$h = \sum_i m_i h_i \quad (2.16)$$

where

$$h_i = \int_{Tref}^{T} c_{p,i}\, dT \quad (2.17)$$

where m is mass, c_p is specific heat, and T_{ref} is a reference temperature. The energy equations can be written in terms of h as in equation 2.18 below.

$$\frac{\partial}{\partial t}(\rho h)+\frac{\partial}{\partial x_i}(\rho u_i h)=\frac{\partial}{\partial x_i}(k\frac{\partial T}{\partial x_i})-\frac{\partial}{\partial x_i}\sum_{j1}h_{j1}J_{j1}+\frac{\partial p}{\partial t}+u_i\frac{\partial p}{\partial x_i}+\tau_{ij}\frac{\partial u_i}{\partial x_j}+S_h \quad (2.18)$$

where T is the temperature, τ_{ij} is the viscous stress tensor, J_{j1} is the flux of species $j1$, and k is the mixture thermal conductivity. S_h is a source term that includes sources like energy due to chemical reaction, radiation, and heat exchange. The second last term is a viscous heating term that is an optional term that should be activated when assuming compressible flow.

2.4.3.4 Turbulence Modelling

This study utilises Eddy Viscosity Models (EVM) to close the Reynolds-Averaged Navier- Stokes equations. An example of this is the k-ε turbulence model that assumes proportionality between Reynolds stresses in the fluid and mean velocity gradients. Although the form of the momentum equations remain the same, the viscosity term becomes an effective viscosity μ_{eff}, and is determined by the sum of the molecular viscosity μ and a turbulent viscosity μ_t. The turbulent viscosity can be determined from equation 2.19 below.

$$\mu_t = \rho C_\mu \frac{k^2}{\varepsilon} \quad (2.19)$$

where k, is the turbulent kinetic energy, ε the turbulent dissipation rate and C_μ is an empirically-derived constant of proportionality (default value equals 0.09 in Fluent).

A further example of an EVM is Wilcox's k-ω turbulence model [20]. This is a relatively new model, more capable of handling higher curvature flows than the standard k-ε model.

In the implementation of the *k*-ε turbulence model and other EVMs, the 'near-wall functions' model near-wall turbulence. This method prescribes a certain shape or function that describes the near-wall velocity profiles but satisfies no-slip conditions. This model proves very effective provided a certain y^+ range is adhered to. y^+ is a parameter that describes the size of a grid volume relative to the assumed velocity profile, in non-dimensional terms and can be compared to a wall Reynolds number. It is generally used to establish the correctness of near-wall grid density according to the requirements stipulated for wall functions.

2.4.3.5 Volume of Fluid Method

A further model that needs to be discussed, due to its relevance to this study, is the multiphase volume of fluids method (VOF). The model was originally developed by Hirt, et al. [21]. This model is used for the tracking of the interface between phases. The tracking is done via the solution of a continuity equation as shown in equation 2.20 below.

$$\frac{\partial \alpha_q}{\partial t} + u_i \frac{\partial \alpha}{\partial x_i} = S_{\alpha q} \tag{2.20}$$

where α is the volume fraction of the q-th phase, and the source term on the right-hand side allows for the use of cavitation models. Intuitively, the following constraint will apply for n phases:

$$\sum_{q=1}^{n} \alpha_q = 1 \tag{2.21}$$

When determining the properties in each control volume, an equation of the form shown in 2.22 below is solved, where A is a typical property like density or viscosity.

$$A = \sum \alpha_q A_q \tag{2.22}$$

The VOF model is a simple and efficient model and works well for the tracking of large free surfaces, as opposed to mixing and dispersing flow situations [22]. Some current restrictions of this model include not being compatible with Large Eddy Simulations (LES), or species mixing and reacting flows. Although other methods exist for the tracking of free surfaces it has been shown through studies like those performed by Cariou and Casella [6] that the VOF model is both the most efficient and most commonly used method.

Many years of development and verification have led to the point that modern day CFD codes are considered to be reliable engineering design tools, used even for standardisation and pre-release testing in the aerodynamics industry. However this does not mean that experimental verification is not necessary. The nature of the technique allows for a lot of adjustments in solver setup, a factor which can influence both the speed and accuracy of the solution.

2.4.4 Structural Finite Element Methods

Finite Element Methods (FEM) has become a powerful and widely used method for the numerical analysis of engineering problems. Although the method is often confined to the analysis of structures, more complex problems like fluid flow or magnetic fluxes can also be tackled with this method. The basic approach involves the discretisation of a potentially complex region that defines the continuum, into simple geometric shapes called finite elements. The properties of the appropriate material are considered in conjunction with the governing equations and expressed in terms of unknown values at the element corners or nodes. These unknowns are in turn determined by the loading and constraints that apply to that portion of the continuum. Solution of the resulting equations will give us an approximation to the behaviour of the continuum, with results manifesting in the form of stresses and strains [23].

Of more direct interest in this study are the models that are used when analysing Fluid-Structure Interaction (FSI) problems. Typical instances of computational FSI are in aero-elastics, where flow over an elastic aerofoil or oscillating cylindrical objects are solved, or in the bio-mechanics field where elastic behaviour of micro-pumps or artificial valve membranes may be of interest. FSI in general is the coupling of fluids and structures and can be done in either a partitioned solution or in a single code. There are advantages and disadvantages in both techniques. The most significant point in the development of these methods is one of accuracy versus performance. This is in part due to the variation in formulation of the fluids (Eulerian) and structures (Lagragian) equations, and thus the treatment of a moving boundary in a fluid domain. It is accepted that partitioned solutions exhibit the most accurate results, since separate yet well-established solution techniques can be used for each sub-problem. It does however mean that each sub-problem is solved separately and coupling data are exchanged at the end of each iteration or time step. This in turn translates to quite a computationally expensive technique, and from a practical perspective requires the use of two computational solution codes, that can become financially expensive. A good example of the use of partitioned techniques can be seen in Matties and Steindorf [24]. An alternate option is one where both continua are solved simultaneously and linked within a single solution code. An example of this approach is the Multi-Material Arbitrary Lagragian Eulerian (ALE [25]) formulation that can be seen in a commercial capacity in LS-DYNA [26]. An example of the use of this technique can be seen in Le Sourne, et al. [27]. Even when solving both problems in one code, a coupling formulation must still exist. The two coupling techniques that are in use in LS-DYNA are the Constraint-based formulation and the Penalty-based formulation [28]. Although these are not the only coupling formulations that exist (other methods include Nitsche's method [29]) they are well established and accepted within their suggested applications. The constraint-based formulation is an algorithm that alters the velocities of the nodes of the solid and shell elements implicitly and forces them to follow each other. In this study, the solid elements are used to model the fluid, while the shell elements are used to model the container structures. The method attempts to conserve momentum but not energy, and is regarded as quite stable. The Penalty-based formulation on the other hand applies nodal forces explicitly by tracking the relative motion of a given point. The method conserves energy but is not as stable as the constraint-based formulation.

The ALE method allows for the use of both Eulerian meshes and Lagrangian structures within the same model simultaneously. This is particularly useful in this study where a simple crash analysis of a fuel tank requires Lagragian shell elements for the baffles and a rigid body for the tank that will "contain" the multi-material Eulerian mesh for modelling the partially-filled container. More detail on the modelling procedure used will be given in section 5.2.

2.5 Experimental Methods

In this section, two fields of experimental methods are discussed. The first group covers those methods that are used to verify the integrity of products according to certain industrial safety standards. Another group of experimental methods is intended to assist in the verification of the validity of numerical models. Both groups are very important and unlikely to disappear in the near future, but high costs are often associated with these techniques.

Vehicle safety is one of the foremost issues in motor vehicle manufacture, primarily because people demand safer vehicles but also because certain industrial safety standards must be met before a vehicle may be sold. The National Highway Traffic Safety Administration (NHTSA) has been issuing Federal Motor Vehicle Safety Standards (FMVSS) since 1967 to which motor vehicle manufacturers in the United States must conform and certify compliance. NHTSA crashes vehicles from many manufacturers every year in an attempt to improve the safety standards themselves. Car manufacturers in turn crash up to 100 of the same vehicle to improve their levels of safety. Since motor manufacturers are well aware of these standards, very few vehicles ever fail the FMVSS tests. This prompted NHTSA to develop the New Car Assessment Program (NCAP). FMVSS basically involves crashing a vehicle at 30mph and confirming compliance with all safety regulations, while NCAP involves crashing the vehicle at 35mph and determining the level of safety in the vehicle. NCAP results provide a star rating for a vehicle that is based on the probably of the

driver or passenger sustaining serious injury. The crashes involve both frontal and side impacts and examine criteria like Head Injury Criteria (HIC), Chest deceleration, Femur load, Thoracic Trauma Index (TTI), and Lateral Pelvis Acceleration (LPA). The FMVSS are however far more extensive, and of particular interest is FMVSS 301, or the Fuel System Integrity standard. The standard involves three crash scenarios (front, rear, and side) to be covered by two future standards FMVSS 208 and FMVSS 214. After the crash event, the vehicle's fuel tank is inspected for integrity and fuel cut-off systems are evaluated. The actual regulation is as follows: FMVSS 301 [11] *"In the frontal impact test, a vehicle is driven forward into a fixed barrier at 48 km/h (30 mph), while in the side impact test, a 1,814 kg (4,000 lb) barrier moving at 32 km/h (20 mph) is guided into the side of a stationary vehicle, and in the rear impact test, a 1,814 kg (4,000 lb) barrier moving at 48 km/h (30 mph) is guided into the rear of a stationary vehicle. The standard limits fuel spillage from crash-tested vehicles to 28 grams (1 ounce) by weight during the time period beginning with the start of the impact and ending with the cessation of vehicle motion and to a total of 142 grams (5 ounces) by weight during the 5-minute period beginning with the cessation of motion. During the 25-minute period beginning with the end of the 5-minute period, fuel spillage during any 1-minute interval is limited to 28 grams (1 ounce) by weight."* Statistically fuel system integrity is of critical importance, according to the US Fatality Analysis Reporting System (FARS) in 1998, four percent (1,411) of light vehicle occupant fatalities occurred in crashes involving fire [11]. For all the above reasons experimental testing of vehicles, their components and of potential occupants (in the form of crash-test dummies) is of great importance in the vehicle design process.

The second group of experimental methods adopted by engineers is those used to validate mathematical models. With the increased use of Finite Element (FE) methods and Computational Fluid Dynamics (CFD) codes in the design process of vehicles and many other engineering structures and machinery, it is always important to maintain a reasonable level of understanding of the capabilities and limitations of the codes used. FE codes are typically verified by comparing stress, deflection, or acceleration data. An example of one such comparison can be seen in Marzougui, et al. [30], where acceleration data, both raw and processed (HIC, etc.), are compared for both a full

scale crash test and the LS-DYNA model that is presented. Typically the author will discuss the results with the intention of describing pitfalls and successes, followed by suggestions for improvement. CFD codes on the other hand are typically verified by comparing pressure, velocity, or temperature data. An example of one such comparison can be seen in an application brief presented by Fluent.Inc [19], Hadzic, et al. [31], where pressure data are compared for both an experimental setup as well as the presented model. A further method of comparison can also be seen in this particular study, one which is usually specific to transient analyses, and that is a visual comparison. In this study the behaviour of the free surface with respect to time is compared. It is interesting to note that this brief is in fact presented by the company that provides the CFD software package. The excellent agreement between experimental and simulated results is clearly a good marketing point.

2.6 Mathematical Optimisation

As this study combines CFD and mathematical optimisation, this section describes some of the methods and terms used in the field of mathematical optimisation.

Optimisation is a tool used in many engineering fields, both consciously and subconsciously. The most common form of design optimisation is usually based on a trial-and-error method and requires large amounts of experience and resources if any level of frequent success is to be achieved. If however the performance criteria in question can be quantified, the more precise method of mathematical optimisation may be employed. A branch of this method is numerical design optimisation that further involves the numerical modelling of the design to quantitatively evaluate the performance of the design. The flow chart shown below in Figure 2.12 provides an overview of the basic mathematical optimisation method.

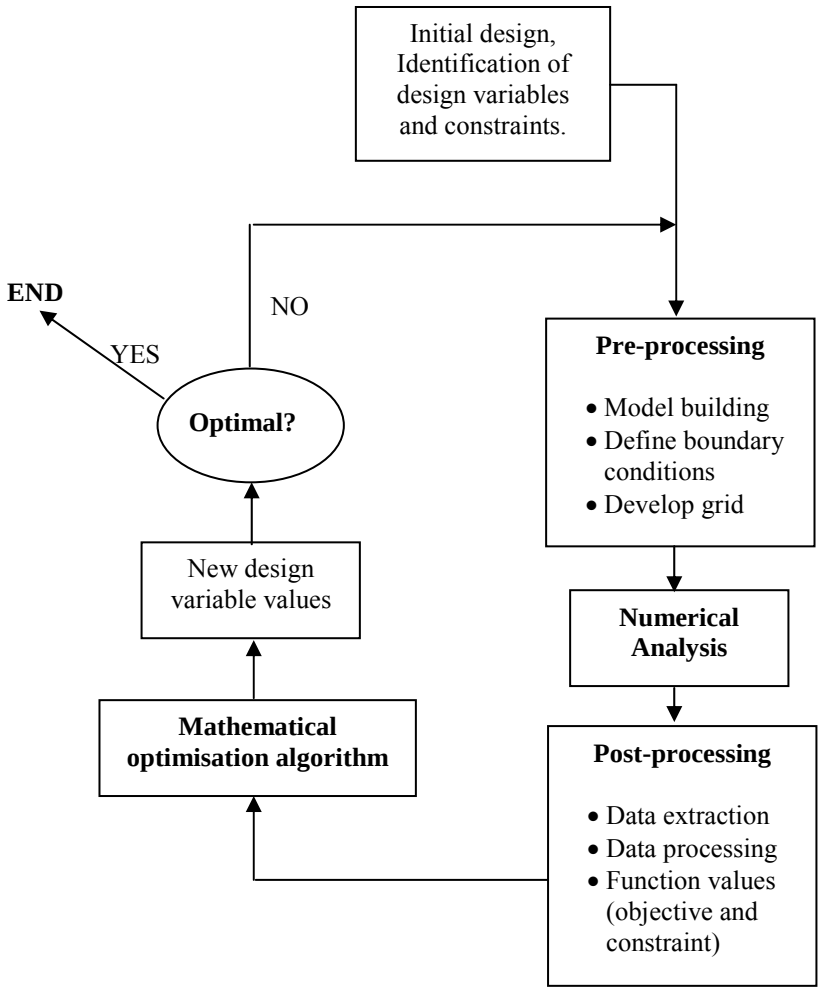

Figure 2.12: Basic mathematical optimisation flow-chart

The basic format of a mathematical optimisation method is as follows:

Minimise the objective function: $f(\mathbf{x})$

$\mathbf{x}=[x_1,x_2,\ldots x_n]^T \in R^n$

Subject to m inequality constraint functions: $g_j(\mathbf{x}) < 0 \qquad j = 1, 2,\ldots m$

and r equality constraints: $h_j(\mathbf{x})=0 \qquad j = 1,2,\ldots r$

where **x** is the vector of design variables in question that require optimisation under these criteria.

An established optimisation algorithm is the leap-frog method for constrained problems (LFOPC) (Snyman) [32]. The method is used internally in a number of design optimisation software packages, e.g., TDO [33], LS-OPT [34], to find the optimum on the approximated surface. The method has proven to be very appropriate for the optimisation of engineering problems where function values are time consuming to evaluate and may exhibit somewhat noisy results. Some of the major characteristics of the basic method are as follows [35]:

- Uses only gradient information for the function value.
- No explicit line searches are performed.
- Very robust, handles steep valleys, discontinuities, and noise in the objective function.
- It seeks relatively low local minima and thus provides a good basis for global optimisation.
- Less efficient than classical methods for smooth near quadratic functions.

LFOPC is a gradient method that generates a dynamic trajectory path from a given starting point towards a local optimum. The underlying analogical principal for this optimisation algorithm is that a particle's motion is traced as it moves over a surface, supposedly under the influence of a conservative force field. The higher the function value, the higher the particle's potential energy. So the point of lowest potential energy will be the optimum. Extending the analogy, as the kinetic energy of the particle increases, so the function value decreases. To ensure that the particle does not continue to oscillate in a valley, interfering strategies are imposed whenever the kinetic energy decreases. The method was extended to constrained problems (LFOP

to LFOPC) by the addition of a penalty function, which effectively artificially creates a steep gradient when a constraint is violated.

Global optimisation in typical engineering applications is an inherently difficult goal to achieve since knowledge of the objective is usually quite localised (also, most engineering functions are not strictly convex over **x**). However, certain methods can alleviate the difficulties experienced to some extent. The practice of multi-starting [36,37] is one that involves initiating the optimisation process from various randomly selected initial values for **x**. Further global optimisation techniques like Neural Network Metamodels and Kriging Interpolations (see section 2.6.2.3/4) can also go some way to alleviating these problems. These methods retain all information from previous function value evaluations in their approximation of the objective and constraint functions, thus maintaining a better global perspective of the problem.

A further important term in optimisation is *saddle points*. Figure 2.13 below illustrates an example of a function that has a saddle point. The saddle point is shown as a black star. Mathematically, a saddle point is described as follows [35]:

f(**x**) has a saddle point at $\mathbf{x'} = [\mathbf{x}^{\circ}, \mathbf{y}^{\circ}]^T$ if an $\varepsilon > 0$ $\exists$ such that for all **x**, $\| \mathbf{x} - \mathbf{x}^{\circ} \| < \varepsilon$ and **y**, $\| \mathbf{x} - \mathbf{x}^{\circ} \| < \varepsilon : f(x, y^{\circ}) \leq f(x^{\circ}, y^{\circ}) \leq f(x^{\circ}, y)$

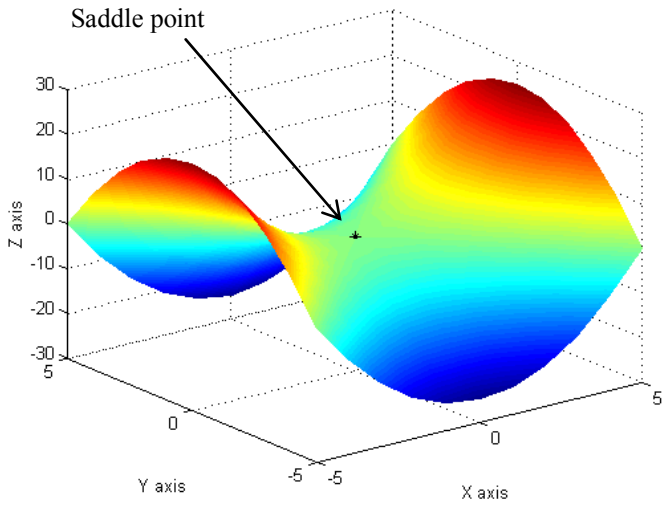

Figure 2.13: 3D function with saddle point ($Z = X^2 – Y^2$)

This concept will also prove to be of significance in this study, where a worst case needed to be found for one of the variables. One of the variables may represent an operational state for the design, in which case the design needs to perform best at its worst operational state.

The sections that follow describe some of the commercially available methods for numerical optimization, as well as methods that are under development.

2.6.1 Dynamic-Q

Snyman's DYNAMIC-Q method [38] is one of the optimisation methods investigated in this study. This method solves successive quadratic sub-problems using a gradient-based optimisation technique to find local minimums for the real or objective function. The successive sub-problems or approximations are generated by sampling the behaviour of the objective or exact function at specific points in the field. A quadratic sub-problem approximation is then generated based on the gradient at a specific point in the field. Constraints within the field of optimisation can be accommodated by altering the sub-problem to generate a penalty function that, as previously mentioned, effectively creates a steep gradient when the constraint is

violated. For each optimisation step, it is necessary to evaluate the objective function value (n+1) times, where n represents the number of variables in question. The DYNAMIC-Q method is also characterised by its use of 'move limits', as an aid for convergence. These represent pre-specified limits on the distance the algorithm may move the design from its current position. In general, this prevents the design from overshooting the optimum due to a quadratic sub-problem whose optimum lies far beyond the local minimum. DYNAMIC-Q is regarded as a very robust method for determining local minimums, and is considerably more economical than genetic or other stochastic type algorithms [39].

For simplification of use, the DYNAMIC-Q optimisation method has been made available in a package (by the University of Pretoria) that has other optimisation methods as options as well as guiding the user through the steps using a graphical user interface (GUI). The package used for this study is the TOOLKIT for DESIGN OPTIMISATION® (TDO) [33] developed by the Multidisciplinary Design Optimisation Group (MDOG) at the University of Pretoria.

2.6.2 LS-OPT

LS-OPT [34] is a further optimisation "toolkit" which is currently packaged with a full purchase of LS-DYNA, and like LS-DYNA is also developed by LSTC [26]. LS-OPT utilises a number of techniques that can be collectively called Response Surface Methodology (RSM) [40], to construct smooth approximations to the actual objective functions in the appropriate number of dimensions, i.e., the dimensions represent the design variables. These response surfaces lead to the approximated sub-problem, and it is the sub-problem which is mathematically optimised. The main reason for the commitment to RSM is that it alleviates the effects of noise, something that can be a major issue when using gradient-based methods.

A further interesting feature of LS-OPT is its trade-off curve capabilities. This tool uses the constructed response surface and the principal of Pareto† optimality to establish a curve that represents the best possible result of one criterion if a compromise must be made on another. The curve is constructed by considering a number of values for the criteria to be compromised, reporting the best result, and joining the results in a piecewise linear fashion.

Within an LS-OPT reference frame, experimental design is the procedure through which the points required for the construction of the response surface are selected. These are the points that must be analysed using the solution code in question (e.g. FEM or CFD) to determine their effectiveness. There are a number of techniques available for selecting these points [40]. LS-OPT provides factorial, Koshal, composite, *D*-Optimal, and Latin Hypercube methods. However, only *D*-Optimal design [37, 40] is considered in this study due to its abilities in handling strict constraints, e.g., geometric constraints and irregular design spaces. *D*-Optimal design selects its points based on the solution of $\max|X^TX|$, where X is a matrix of possible experimental design points and X^TX occurs in the definition of the least-squares fit coefficients. In LS-OPT, a genetic algorithm is used to solve this maximisation problem. X is a subset of points selected from a larger group of basis points that are spread over the region of interest in either an organised or random manner. The number of points in X will depend on the chosen response surface. The idea behind *D*-Optimal design is to try to reduce the response surface's approximation error due to variance in the experimental results. This would seem a good idea when considering the variance one may get in practical experimental results, however with numerically calculated experimental results no variance will occur for the exact same set of design points. For this reason, Qu, et al. [42] suggest the minimisation of bias error due to the choice of response surface as a better criterion.

---

† V. Pareto, a prominent Italian economist, introduced the idea of Pareto optimality at the end of the 19th century. To define the notion of domination let $\mathbf{f}=(f_1,...,f_n)$ and $\mathbf{g}=(g_1,...,g_n)$ be two real-valued vectors of *n* elements; **f** is partially less than **g** ($\mathbf{f} <_p \mathbf{g}$) if:

$$\forall i \in (1,...,n); f_x \le g_i ... \ \& \ \exists i : f_i < g_i$$

If $\mathbf{f} <_p \mathbf{g}$, we say that **f** dominates **g.** Consequently, a feasible solution $\mathbf{x}^*$ is said to be a Pareto optimal if and only if another **x** does not exist such that $\mathbf{f}(\mathbf{x}) <_p \mathbf{f}(\mathbf{x}^*)$. [41]

The following four sections give a brief description of the available response surfaces in LS-OPT. (NB. Kriging Interpolation is available in alpha phase only)

2.6.2.1 Linear Approximations

Linear approximation is inherently the most economical approximation method within a single optimisation iteration. However, the method can be less stable and will often require more optimisation steps to converge, which could ultimately mean more function evaluations are required. The number of points required per optimisation step when using the *D*-Optimal design approach, can be calculated by "1.5*(n+1)+1" for 'n' variables. This may seem like too many points at first, but the multiplication by 1.5, or 50% over sampling, goes a long way to providing one with more certainty about the appropriateness of your response surface. Furthermore, the over sampling filters some of the noise due to the numerical evaluation of the responses surface.

2.6.2.2 Quadratic Approximations

Since many phenomena behave quadratically with respect to their variables, it is often far more sensible to use quadratic approximation techniques. In a *D*-Optimal design, the number of points required per optimisation iteration is, "1.5*(n+1)(n+2)/2 +1". Intuitively one would require more points to construct a quadratic surface than a linear surface, but if the bias error is reduced by using a quadratic response surface, one can expect that convergence may occur sooner than with a linear approximation. In turn, one may require fewer function evaluations over the span of the optimisation.

Both Linear and Quadratic approximation methods in LS-OPT make use of the Successive Response Surface Method (SRSM) [43]. This method effectively relocates and resizes the subspace of interest for each optimisation iteration, in an attempt to "zoom-in" to the region of the optimum design. This results in a more effective use of

the function value evaluation to gain knowledge only near the region of interest. However, since a new response surface is constructed, previous data collected in prior optimisation iterations are lost and the new surface loses any previous global validity.

2.6.2.3 Neural Network Metamodels

One of the main points of interest when considering neural networks (NN) and Kriging methods as compared with polynomial response surfaces (e.g., linear and quadratic) is the ability to update the response surface with appended data from subsequent optimisation iterations. This gives improved global validity as well as higher definition in the region of interest when combined with the domain reduction approach of SRSM.

On a basic level, a NN consists of input and outputs that are linked by the neural network's neurons which compute the outputs from the inputs. The NN is defined by parameters like the inter-neuron connection strengths and biases on both the input and output side of the neurons. These parameters are determined in the *learning* process, in which a training algorithm directs the parameters to a state where the error in approximation is minimised. For a more definitive description of Neural Networks see e.g. reference 44.

2.6.2.3 Kriging Interpolation

Kriging is a spatial interpolation technique originally developed by the South African geostatistition D G Krige, in an attempt to more accurately predict subterranean ore reserves. Kriging interpolation behaves similarly to NN in that it has the ability to update itself once more data is appended. The method uses a stochastic correlation to relate a known polynomial function to the unknown function of interest. A complete description can be seen in Simpson [45]. The method's main advantage is its accuracy, since it interpolates the available data points. In comparative studies,

Kriging interpolation has demonstrated similar performance characteristics to the previously mentioned metamodelling techniques, but has shown to be less robust [46] and more sensitive to noise [47]. The method is however still under development by LSTC and is only available indirectly and in an alpha phase within LS-OPT.

2.7 Multidisciplinary Design Optimisation

Throughout this chapter, references are made to journal publication and other sources that describe the work that has been done within that specific field. However, this study aims to examine a number of the fields simultaneously.

Due to the diversity of the problem in question, it is felt that a design technique be analysed that encompasses as many as possible of the most suitable methods. This idea leads the study to the field of Multidisciplinary Design Optimisation (MDO). MDO, as the name suggests, is the optimisation of a given design where multiple disciplines are considered for the purposes of establishing design performance.

Within the MDO framework a number of formulations exist. The appropriate formulation depends on factors like the degree of coupling between the disciplines and the ratio of shared to total design variables [48]. A paper by Giesing and Barthelemy [49] provides an industry perspective on MDO applications and needs. The standard method is to evaluate all disciplines simultaneously in one integrated objective and constraint set by applying an optimiser to the multidisciplinary analysis (MDA), similar to single-discipline optimisation. Figure 2.14 below is an adaptation of Figure 2.12 above that shows the similarity with single-discipline optimisation (SDO). The major difference between SDO and MDO is that two or more solvers may exist in parallel. Furthermore, all variables are not necessarily used by both solvers, this may be since certain variables do not form part of the model used for a discipline or that the influence the variable has on the results used from the discipline are insignificant. For this reason, variables are isolated into sub-sets used only by the specific discipline. This method of MDO is numerically the most correct, but in

industry it may not always be possible to analyse a set of design with multiple disciplines simultaneously or even at the same physical location. It is for this reason that other formulations exist. Fortunately for the purposes of this study, both disciplines, CFD and FEM, will be available simultaneously.

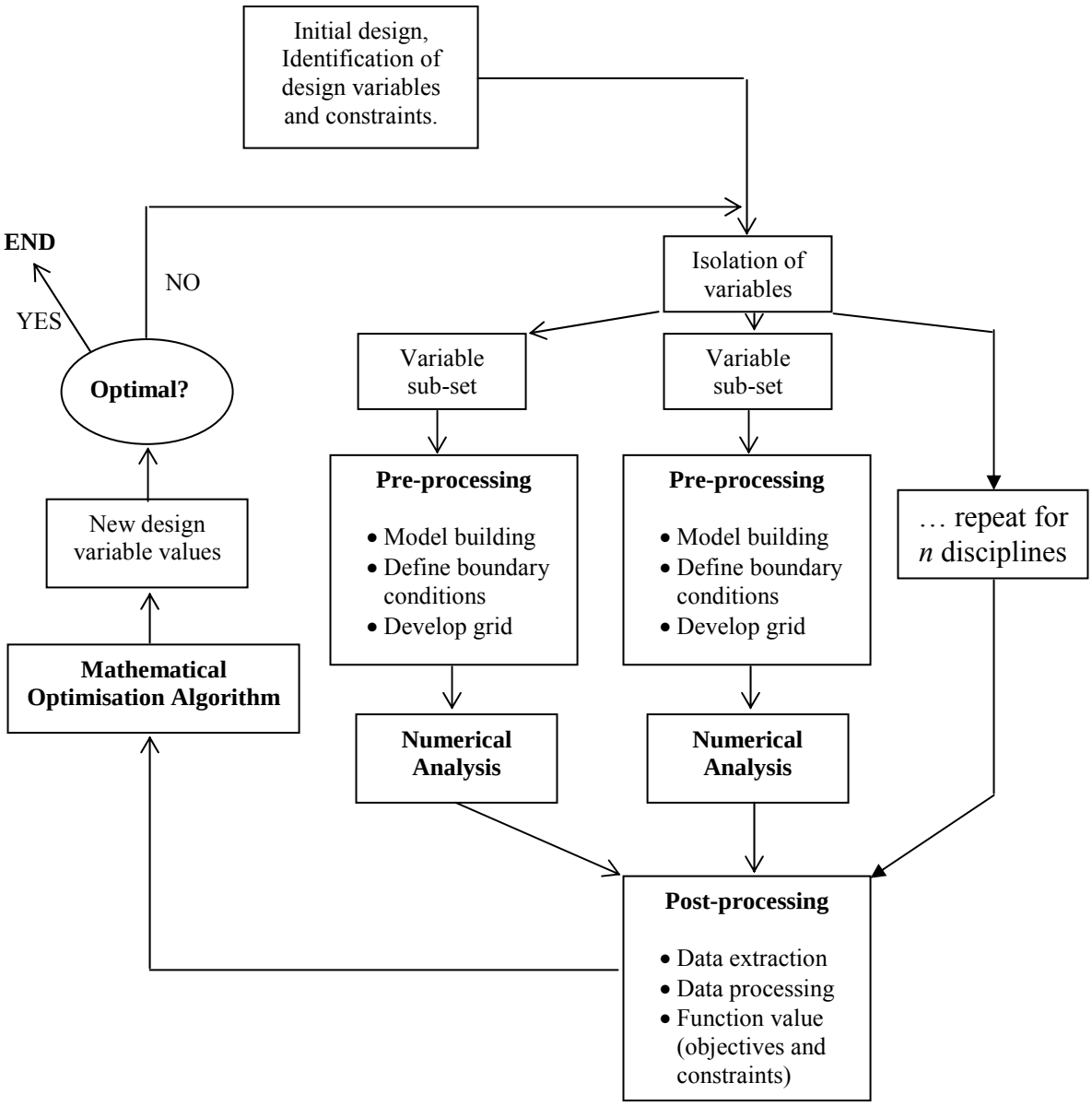

Figure 2.14: MDO cycle for *n* disciplines

2.8 Conclusion

Sloshing is a real and present challenge, one which has troubled engineers over the last century in one form or another. This chapter has one outstanding message; there are a wide number of techniques that are available for use in the design of liquid tanks, a number of which are interdependent. Depending on the application, it may be required of the engineer to consider a range of design criteria and standards that are pertinent. This study will examine a number of techniques and issues surrounding liquid container design, both from an optimisation and validation perspective.