5-31-1992

Computational aspects of a three dimensional non-intrusive particle motion tracking system

Avadhani S. Ashok
New Jersey Institute of Technology

ABSTRACT
Computational Aspects of a Three Dimensional Non-Intrusive Particle Motion Tracking System.

by
Avadhani S. Ashok

Development of a technique for non-intrusive particle motion tracking in three dimensions is considered. This technique is based on the principle of magnetic induction. In particular, the determination of the position and orientation of the particle from the information gathered is the principal focus of this thesis. The development of such a system is motivated by the need to understand the flow patterns of granular material. This is of critical importance in dealing with problems associated with bulk solids flows which occur in almost all industries and in natural geological events. A study of the current diagnostic techniques reveals the limitations in their ability to track the motion of an individual particle in a mass flow of other particles. These techniques fail when the particle must be tracked in three dimensions in a non-intrusive manner. The diagnostic technique we consider results in an unconstrained minimization problem of an overdetermined system of nonlinear equations. The Levenberg-Marquardt algorithm is used to solve such a system to predict the location of the particle. The viability of this technique is established through simulated and actual experimental results. Practical problems such as the effect of noise are considered. Directions for future work are provided.

COMPUTATIONAL ASPECTS OF A THREE DIMENSIONAL NON-INTRUSIVE PARTICLE MOTION TRACKING SYSTEM.

by
Avadhani S. Ashok

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science
Department of Mechanical and Industrial Engineering
May 1992

APPROVAL PAGE

Computational Aspects of a Three Dimensional Non-Intrusive Particle Motion Tracking System.

by
Avadhani S. Ashok

May 8, '92

Dr. Rajesh N. Dave, Thesis Advisor

Assistant Professor of Mechanical Engineering, NJIT

May 8, 1992

Dr. Ian S. Fischer, Committee Member

Associate Professor of Mechanical Engineering, NJIT

May 8, 1992

Dr. Anthony D. Rosato, Committee Member

Assistant Professor of Mechanical Engineering, NJIT

BIOGRAPHICAL SKETCH

Author: Avadhani S. Ashok

Degree: Master of Science in Mechanical Engineering

Date: May, 1992

Date of birth:

Place of Birth:

Undergraduate and Graduate Education:

- Master of Science in Mechanical Engineering, New Jersey Institute of Technology, Newark, NJ, 1992
- Bachelor of Engineering in Mechanical Engineering, A. C. College of Engineering, Karaikudi, India, 1989

Major: Mechanical Engineering

Presentations and Publications:

Ashok, Avadhani S. "Signal Processing in Non-Intrusive Tracking of a Particle in Three Dimensions." Presented at *Region-II Graduate Student Technical Conference*, Steven's Institute of Technology, April, 1992.

Dave, R. N., A. S. Ashok, and B. G. Bukiet. "On Development of a Three Dimensional Particle Motion Tracking System." *Proceedings of the ASME Winter Annual Meeting*, Anaheim, California, To be published (1992).

To My Dear Parents With Love

ACKNOWLEDGMENT

TABLE OF CONTENTS

Page

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Overview

Experimental studies are essential to validate theories for complex systems and may even be the only source to gain insight in cases where the theoretical models and solutions are unavailable. A diagnostic technique is considered to experimentally obtain data to validate developing theories of granular flow which has applications in bulk solids handling in many industries and in natural geological events.

1.2 Bulk Solids Handling

Today, the variety of materials being handled in bulk is almost endless. It ranges in size from fine dust to rocks introducing a high degree of complexity into the understanding of its flow. Examples of bulk solids handling can be found in almost every kind of industry and the problems associated with the design, installation and operation of plant for the storage and transport of materials in bulk are many and varied. Hence considerable attention has been devoted in the technical literature to the problem of bulk solids handling. Interest in the problem was aroused by the epidemic of flow stoppages in coal bunkers of steam power plants predominantly in the United States (Stepanoff 1969).

The principle concerns are those of maintaining an uninterrupted flow of bulk solids granular material by avoiding the phenomena of arching, doming and bridging without compromising the safety aspects under all possible conditions.

Bulk materials can take the form of granulated substances, crushed rock, powdered materials such as flours, sand and cement, or slurries and liquids. The problem of transport of solids in suspension is closely connected with the problem of storage of grain and

1

pulverized solids. Thus storage facilities for bulk solids represent an important part of the equipment for a number of major industries handling materials such as grain, coal, ore, sand, gravel, cement and flour. The storage of bulk material in the form of powder and other granular materials is usually done in tanks, bins, hoppers, bunkers, silos, etc. and is usually transported through chutes, channels, conveyors, pipes, etc.

A good understanding of the nature of bulk solids flow is an essential prerequisite to the design of virtually any system involving the storage or handling of such materials. There are innumerable examples in industry of problems that have been attributable to insufficient attention being paid to the properties of the bulk solid involved. Failure of a material to discharge from a storage hopper, blockage of a pneumatic conveying line and uncontrollable flushing of material through a weigh-feeder are typical of lesser problems. More serious problems include dust explosions and other mechanical damages. There are many records of serious mechanical damage occurring to bulk solids storage vessels, notably grain silos, as a result of an apparent physical weakness of the vertical walls. Investigations subsequently showed that the problem was basically due to the failure of the designer to appreciate that during discharge of the material from the bin or silo the lateral pressures developed could be considerably greater than existed with the material at rest.

Many bulk particulate solids, when dispersed in air to form a dust cloud, constitute a potentially explosive mixture which may be ignited by a naked flame, a hot surface or an electrical discharge. The range of products that are hazardous in this respect is quite wide and includes common foodstuffs such as sugar, flour and cocoa; synthetic materials such as plastics, chemicals and pharmaceuticals; metals such as aluminium and magnesium; and traditional fuels such as coal and wood.

Whilst there are certainly some similarities between the flow characteristics of bulk solids and liquids, it is more appropriate to model a bulk solid as a plastic solid than

as a fluid continuum. In general, it is much easier to handle liquids than dry materials. Variation of the properties of solids with time is the main cause of the difficulties.

Bulk solids at rest can transfer shearing stresses and, in many cases, possess sufficient cohesive strength after consolidation to retain their shape under pressure. Furthermore, when a bulk solid "flows" slowly, the shearing stresses within it are dependent upon the mean pressure to a much greater extent than the rate of shear. Thus, there are distinctive features of liquids which are not shared by bulk solids. The ability of a bulk solid to flow may be regarded as the summation of a number of different effects, but it is essentially concerned with the forces of attraction between constituent particles. Thus, when these forces of attraction are low, the bulk material can easily be made to flow under the influence of gravity with the particles moving as individuals relative to one another. Dry sand and granulated sugar are familiar examples of free - flowing materials. However, the high interparticle forces, which may be caused by such effects as moisture or electrostatic charging and are especially pronounced in very fine materials, result in a tendency for agglomerates to form so that the material flows in an erratic manner as "lumps", if indeed it flows at all. Examples of cohesive materials which usually exhibit this sort of behavior are flour and cocoa powder.

Even the prediction of the flow rate in various circumstances is somewhat of an intractable problem. Because of the complex nature of the gravity flow of bulk solids there is as yet no single convenient method that leads to a consistently reliable prediction of discharge rates for the full range of materials and various designs of bins and hoppers. Indeed, for materials of a fine cohesive nature no method has yet been developed that could be confidently recommended to the designer of storage vessels.

The assessment of the flow characteristics of a bulk solid is very much a matter of judgement based on experience assisted by evidence provided from various tests. These tests measure voidage and bulk density, particle density, size, shape, surface area, hard-

ness, cohesion and adhesion, angle of repose, wall friction, moisture content, explosiveness, etc.

At the present time our level of knowledge is not sufficient to enable a reliable prediction to be made of the behavior of a bulk solid solely from the characteristics of its constituent particles. Expert knowledge of conveying of solids is derived by building full scale models which is a very inefficient way to design the handling and storage systems. The theory available from textbooks is more useful as a guide than as a complete answer to a problem (Allegri 1984).

Thus there is clearly a need for developing a diagnostic technique to study the flow characteristics under various conditions. The present methods exhibit fundamental inadequacies in that they are either not non-intrusive or not three dimensional or both. The presently available techniques are discussed in detail in the next Chapter.

1.3 Statement of the Problem

Experimental investigation of the flow patterns of granular material undergoing flows and other mechanical disturbances poses many formidable difficulties. The most prominent of these is the problem of obtaining measurements of individual particle motion in a non-intrusive manner. Experimental studies are essential in order to gain insight into the behavior of systems involving flow of solids, liquids and/or gases for which exact theoretical models and solutions are not available for most system configurations. In the past, experimental studies for dry granular flows have employed various techniques in order to obtain pertinent data. Measurement of individual particle velocities, in particular angular velocities, have been very difficult using these techniques. Such measurements in full three dimensional flows have not been reported. There is obviously a need for developing a technique for measuring the three dimensional position of an individual particle in a mass of other particles.

In this thesis, the development of a non-intrusive tracking system is described. This system is based on the principle of radiosonde transmitters coupled to a set of receiving antennae through magnetic induction. Feasibility of such a system is established unequivocally. The proposed system has many crucial stages. But, the focus of this thesis is on the development of the techniques and algorithms to compute the three-dimensional position and orientation of the particle from the data obtained in the experiments.

1.4 Overview of the Remaining Chapters

Chapter 2 discusses the flow of bulk solids in chutes and the various diagnostic techniques that have been developed for similar experimental studies and provides the goals to be met by any new method to minimize, if not totally eliminate the limitations of the present techniques.

Chapter 3 presents the principles and the methodology of our experimental technique. Different methods to obtain the three-dimensional position and orientation are discussed.

Chapter 4 discusses the numerical methods employed to solve the nonlinear over-determined system which arises in our attempt to track the trajectory of a particle.

Chapter 5 provides the results obtained from simulated data and from actual experimental data. The effect of noise in the signals is considered and it is shown that the algorithms are stable and converge even under these conditions.

Appendix A presents briefly the important properties of bulk solids. The phenomena of arching or doming which is an important cause for flow stoppages is discussed in Appendix B. The solution to elliptic integrals which arise in model #2, discussed in Chapter 3, is provided in Appendix C. Homogeneous coordinates and transformations are used in generating the relative positions and orientations for each transmitter-

receiver pair and is discussed in Appendix D. In using the numerical methods of solution we discuss the convergence rates of various methods. The terms associated with the convergence is defined in Appendix E.

CHAPTER 2

TECHNICAL BACKGROUND

2.1 Introduction

In this chapter we discuss the flow of bulk solids in chutes briefly as our diagnostic technique can be directly used to study the flow pattern in such flows. We also discuss the currently available diagnostic techniques to study the flow parameters in the flow of dry granular material. This will provide the background for the development of our technique discussed in the next chapter.

2.2 Flow of Bulk Solids in Chutes

2.2.1 Introduction

There are many instances in bulk solids handling installations of gravity flow of a particulate or granular material along an inclined channel or chute. For example, where a bulk solid is to be discharged at a point below and to the side of a hopper outlet, it would be common practice to rely on gravity flow through a simple transfer chute. In such situations both straight and curved chutes are used.

Amongst common applications of transfer chutes for bulk materials, perhaps the most familiar occurs at the loading point of a belt conveyor. In this case it is important that the horizontal velocity component of the material leaving the chute is matched to the velocity of the belt in order to minimize the acceleration of this material and so effect reductions in power consumption and belt wear. Other situations may require that the exit velocity is as large as possible and of a direction to obtain the maximum possible

7

"throw" of the flowing material. Thus it is important that the design of gravity-flow chutes and channels is undertaken in the light of a clear appreciation of the characteristics of flow in such situations.

2.2.2 Flow Patterns in Straight Inclined Chutes

Bulk solids in gravity flow may be expected to exhibit two possible modes: varied flow or uniform flow. But these terms do not exactly correspond to "rapid" and "tranquil" flow used to describe liquid flow in channels.

If a particulate bulk solid is fed into a steeply inclined straight chute or channel of constant width, "fast flow" occurs, with the material accelerating, and consequently the depth of the flowing bed decreasing, until some steady condition is achieved, at which the downward component of the gravity force is balanced by the various drag forces on the particles (Figure 2.1a). If the slope of the channel is decreased, the rate of acceleration will also decrease, since the component of the gravity force on the material is smaller. As the slope of the channel approaches the angle of internal friction (Φ) of the bulk solid, the flow tends to become uniform. This condition of fully - developed flow at constant depth or "slow flow" is observed in straight chutes or channels at relatively shallow inclinations, normally only in the very restricted range between the angle of repose (α) and the angle of internal friction (Φ) of the bulk solid concerned (Figure 2.1b).

9

Decreasing depth

H

Increasing velocity (tending to reach a maximum or "terminal" value if the channel is long enough)

Slope θ greater than angle of internal friction

θ

Figure (2.1a) "Fast Flow"

Constant depth

H

Constant Velocity

Slope θ lies between the angle of repose and the angle of internal friction

θ

Figure (2.1b) "Slow Flow"

Figure (2.1) Modes of Flow in a Straight Inclined Chute

Placing an obstruction near the downstream end of a chute in which a bulk solid is flowing in the fast mode can cause a surge wave or stationary jump (sometimes called a "granular jump") to occur in much the same way as a hydraulic jump occurs in a flowing liquid (Woodcock & Mason 1987).

2.2.3 Flow Patterns in Curved Chutes

The general flow patterns that may be observed when a bulk solid flows through a curved, enclosed chute are illustrated in Figure 2.2. As with straight inclined chutes, two modes of flow have been observed to occur, termed "fast flow" and "slow flow" according to whether the stream of material is accelerating or travelling at a uniform slow velocity.

Figure 2.2a shows the general case of fast flow in which the particulate material first accelerates as it falls freely from a hopper into the chute, but then decelerates as a result of the curvature of the channel and the decreasing slope of the bottom surface. There exists an "optimum cut-off angle" (θ_{co} in Figure 2.2b) at which the velocity of the stream is a maximum and the stream thickness is a minimum. Ideally the chute should be terminated at this optimum cut-off angle, since any additional length of chute will result in an increase of stream thickness, frequently leading to an unstable flow condition (Figure 2.2a).

When the cut-off angle exceeds the optimum value, it is quite possible for the thickness of the stream at the lower end of the chute to increase to the point where the flowing material comes into contact with the top surface of the chute. The velocity of the stream of bulk material will then be considerably reduced and a surge wave travels upstream as illustrated in Figure 2.2c. This surge wave indicates a change from "fast" to "slow" flow as the channel becomes completely full of the bulk solid which is then in contact with all four internal surfaces (Figure 2.2d).

Even a temporary obstruction to fast flow in a chute having a cut-off angle greater than the optimum value can be sufficient to initiate a change to slow flow. Indeed, if the cut-off angle is too large (Figure 2.2e) the chute may become choked and flow cease altogether, or flood over the sides if the chute is not enclosed. In order to ensure that fast flow is maintained in a curved chute, the cut-off angle should not exceed some limiting value θ_f which depends upon the sliding friction between the particulate material and the internal surfaces of the chute.

Free fall zone due to "vena contracta" effect

θ_c: Chute cut-off angle
θ_{co}: Optimum value of cut-off

θ_{co}

θ_c

Tendency to form stagnant layer, leading to instability

Acceleration Zone

Deceleration Zone

Figure (2.2a) Fast Flow: General Case

Figure (2.2b) Fast Flow: Ideal Case

$\theta_c = \theta_{co}$

$\theta_c > \theta_f$

θ_f: limiting value of θ_c for "fast" flow

Surge wave

Figure (2.2c) Transition from Fast to Slow Flow

Figure (2.2d) Slow flow

Figure (2.2e) Choked Flow Condition

Figure (2.2) Modes of Flow in a Circularly Curved Chute.

2.3 Background on Existing Techniques

2.3.1 Classification

The experimental investigations (Tuzun *et al.* 1982)of the flow fields produced by granular solids employ a variety of techniques some of which are described below. The choice of a particular method seems to be decided by two primary considerations:

(1) the depth, i.e. the three-dimensionality of the flow field under investigation;

(2) the nature of the flow parameters to be measured.

The majority of the experimental work is carried out with two-dimensional, plane strain equipment which restricts observations to a single plane. Due to the much increased complexity of the three-dimensional experiments the complexity of the problem is reduced with the assumption of axial symmetry which reduces it to a 2-D problem despite the wide scale use of 3-D bunkers in industrial processes.

Aspects of the flow field which are of practical interest are given by the distributions with respect to time and position of:

(1) the components of particle velocity;

(2) the interstitial voidage and hence the bulk density;

(3) interstitial fluid pressure (not for dry granular flow);

(4) interparticle stresses.

The experimental techniques used to gather information are either "continuous flow" type or "stop-start" type and may be classified as:

- Tracer Techniques
  - Direct Detection
    - Visual and Photographic work
    - X-ray and γ-ray investigations
    - "Flow freezing" techniques

- Remote Detection
    - Radio-isotope
    - Radio-pill
    - Magnetic tracer
- Probe Techniques
    - Obstacle probe
    - Capacitance probe
    - Fibre optic probe

2.3.2 Tracer Techniques

Motion of distinct tracers within bulk materials are followed either directly or with the aid of remote detection equipment. The accuracy of the measurements is thus a direct function of the ability of the tracer to closely follow the bulk flow. It has to be ensured that there is no preferential segregation of the tracer from the bulk material. Otherwise irreproducible and hence inaccurate measurements will result. Preferential segregation of the tracer from the bulk material can result from differences in either size, specific gravity or frictional properties. The tracer movements may be followed continuously or by "stop-start" mode which is useful in the study of transient effects.

2.3.2.1 Visual and Photographic Work: Many investigators have made visual observations of the flow field next to a transparent wall plane using distinct dyed tracers. But this provides only the qualitative measurements. Hence, others have determined velocity fields with the use of high speed photography (Drake 1991). Long-exposure time photographs of the flow field have also been used to detect changes in the shape and position of the flow boundary during continuous discharge. An alternative way of analyzing photographic data is the "stereoscopic technique". Photography can also be

employed to reveal the voidage distribution within a plane-strain bin with transparent walls if the bulk material is also transparent.

2.3.2.2 X-Ray and γ-Ray Investigations: X-ray radiography (Bransby *et al.* 1973) is used in place of photography to study flow in bins of relatively large depth/ width ratios. An X-ray source is used to pass radiation through the material in the bunker and expose the radiographic film placed behind. The radiographs exposed at different stages of flow help to build a dynamic picture of the flow pattern. The voidage profiles within the flow fields may be measured using a γ-ray absorption technique. A caesium source is employed to beam γ-radiation through the silo contents and a detector placed at the back is used to measure that part of the beam not absorbed by the bulk material.

2.3.2.3 "Flow-Freezing Techniques": Layers of colored material are placed at various heights during the initial filling of the bunker. The flow fields corresponding to different time intervals after the start of discharge are obtained by closing the orifice at a given time and "freezing" the bunker contents in a cast of self-hardening substance. The flow patterns at different planes can then be investigated by cutting longitudinal slices of appropriate thickness of the resulting solid cake.

2.3.2.4 Radio-Isotope Tracer: McCabe (1974) has measured particle trajectories within the converging flow zone immediately above the discharge port from a cylindrical bunker. The trajectories were traced using gold isotopes. The tracer was followed using a mobile isotope scanner (Lin *et al.* 1985).

2.3.2.5 Radio-Pill: Radio-pills have been used extensively for measurements of internal pressures within a flowing medium as well as the velocity distributions. The

pills used for velocity measurements transmit radio signals at a set frequency thereby indicating the position of the pill within the flow field as a function of time. To measure the internal pressures within a flowing medium, Perry *et al.* (1970), Rao and Venkateswarlu (1975) have placed a pressure sensitive diaphragm inside the radio-pill. In this case, the frequency of the signals obtained from the pill is allowed to vary according to the external force exerted on the diaphragm.

2.3.2.6 Magnetic Tracer:
Smallwood and Thorpe (1980) have determined the velocity distributions in cylindrical bunkers by measuring the distribution of the residence times inside the bunker. The passage of the steel ball bearings was detected by the perturbations caused in the magnetic fields of the two coils placed around the entry tube. The residence time data thus obtained is inverted by numerical integration to yield the corresponding velocity distribution.

Nuclear magnetic resonance (NMR) and electron magnetic resonance (EMR) are closely related in principle. NMR senses the nuclei of a selected species contained within a material while EMR senses the free or unpaired electrons present. Both require that the sample material be exposed to a strong static magnetic field. Detection is then accomplished by sensing the interactions between an applied electromagnetic field and the magnetic moment of the subatomic particles of interest.

In the specific case of NMR (Altobelli *et al.* 1991) the flowing material enters the permanent magnetic field and undergoes nuclear magnetization. It then enters a second magnetic field where resonance occurs when the correct radio-frequency signal is applied. A modulating magnetic field creates demagnetized pockets and the modulating effects are detected by a final RF detector coil. The magnitude of the NMR response is proportional to the number of appropriate nuclei (EMR proportional to unpaired electrons) and the modulation provides velocity information (Beck *et al.* 1987).

2.3.3 Probe Techniques

A number of methods using specially designed probes are used for measuring flow properties in conveying media with low solids content.

2.3.3.1 Obstacle Probe: McCabe (1974) has measured the plug flow velocities in cylindrical bins using an obstacle probe. A 4 cm. diameter horizontal disc is attached to a fine steel wire which is allowed to slide freely in a flexible cable supported from the top of the silo. The velocity of descent during discharge is followed by observing the rate at which the steel wire enters the flexible cable at top. Another variation is the measurement of the drag force exerted on an obstacle suspended within the flowing material. The local velocity is then calculated from the measured drag force (Oki *et al.* 1977).

2.3.3.2 Capacitance Probe: Lockett (1967), Hancock (1970) and Burkett *et al.* (1971) have used capacitance plates to measure local values of the interstitial voidage in plane-strain silos. This method consists of measuring the change in capacitance between two fixed plates of a small capacitance probe due to a change in voidage of the material between the plates.

2.3.3.3 Fibre Optic Probe: Oki et al. (1977) report that with the use of a light-detecting circuit and an automatic signal correlator device, the velocities as well as the sizes of individual particles can be measured within a field of high solids flowrate. The probe consists of two pairs of small optical fibres which can be made of either glass or plastic. One fibre of each pair is used to illuminate individual particles and the other member of the pair detects the light reflected by the particles (Katsuya *et al.* 1975; Savage 1978; Ahn *et al.* 1989; Louge *et al.* 1991).

2.4 Limitations of Current Techniques

We now provide some of the limitations of the various methods discussed above. Visual observations are mostly qualitative and for 2-D only. Usually the walls and the bulk material has to be transparent. X-ray radiography requires increasing amount of radiation for a larger investigation area and bunker thickness. This can pose a health hazard. Also it provides no information as to the magnitude of flowing densities. Only a small section of the flow field can be covered using the γ-ray investigation. Moreover, the facilities required tend to be expensive besides requiring specially trained staff. Flow freezing technique is messy and time consuming as it involves using a self hardening substance. Secondly, it is rather difficult to distinguish the genuine flow deformation suffered by the tracer layers during discharge from further possible distortions during casting and cutting of the solid cake.

Radio-isotopes suffer from the problems of high cost, radiation safety and licensing problems. Radio-pills make assumptions as to the directions of the principal stresses and their variation with time and position which are not valid in reality. Magnetic tracers suffer from an accumulation of errors which become significant near the bin walls.

The probe techniques have the drawback at the outset itself because of their intrusive nature. The obstacle probe is suited for measurements in large-scale apparatus only. The parallel plate capacitors in the capacitance probe are subject to fringing electrostatic fields at their edges. Furthermore, the method can provide only one voidage reading over a relatively large cross-sectional distance at a given height. The fibre optic probe can give inaccuracies in the measurement of both the distance and the time of passage if the separation distance between the detecting points is small, while a given particle may not pass through both the points if the separation distance is large.

Thus the current techniques cannot track the trajectory of an individual particle within a mass flow in three dimensions non-intrusively. Moreover, none of the present

methods can detect the orientations of the particle or provide quantitative information in the whole region of the flow field.

CHAPTER 3

THE PROPOSED TRACKING SYSTEM

3.1 Introduction

In the previous chapter we discussed the various diagnostic techniques to study the flow parameters in the flow of dry granular material. We showed the limitations of the present techniques and clearly established the need for a new diagnostic technique to track the individual trajectory of a particle within a mass of particles non-intrusively in three dimensions. We also established the requirements to be met by the new method.

In this chapter we describe in detail our technique to track the motion of a particle in a three dimensional flow of a mass of particles. We discuss the method of determining the position and orientation of the particle from the information obtained.

3.2 Description of the Tracking System

The proposed tracking system is a novel system based on the principle of magnetic induction coupling (Dave *et al.* 1992). Radiosonde transmitters (Tug 1989) mounted inside the particle to be tracked, henceforth called the tracking sphere, are coupled to a set of receiving antennae through magnetic induction. One or more orthogonally placed transmitters, which are basically loops of wire carrying currents at different frequencies are mounted inside the tracking sphere with its own power source (battery) to act as the *"source"* or the *"emitter"*. Three or more large receiving antennae loops are orthogonally positioned around the experimental space and serve as the *"receivers"* in which the voltages are induced by the principle of magnetic induction.

The voltage signals in the receiving antennae are first amplified using amplifiers, and then pass through bandpass filters and a 12-bit A/D converter. The digital data is

recorded on the hard-disk of a IBM-PC AT486 compatible, which houses an AT-MIO-16 multi function card controlled through the data acquisition LabWindows software. The data is sampled and recorded 100 or more times per second and each data sample consists of a set of receiver voltages. The data is then transferred to a SUN SPARC2 workstation, where it is processed to compute the position corresponding to each set of receiver voltages. Figure (3.1) shows the block diagram of the complete system.

Figure 3.1: Block Diagram Showing the Complete System.
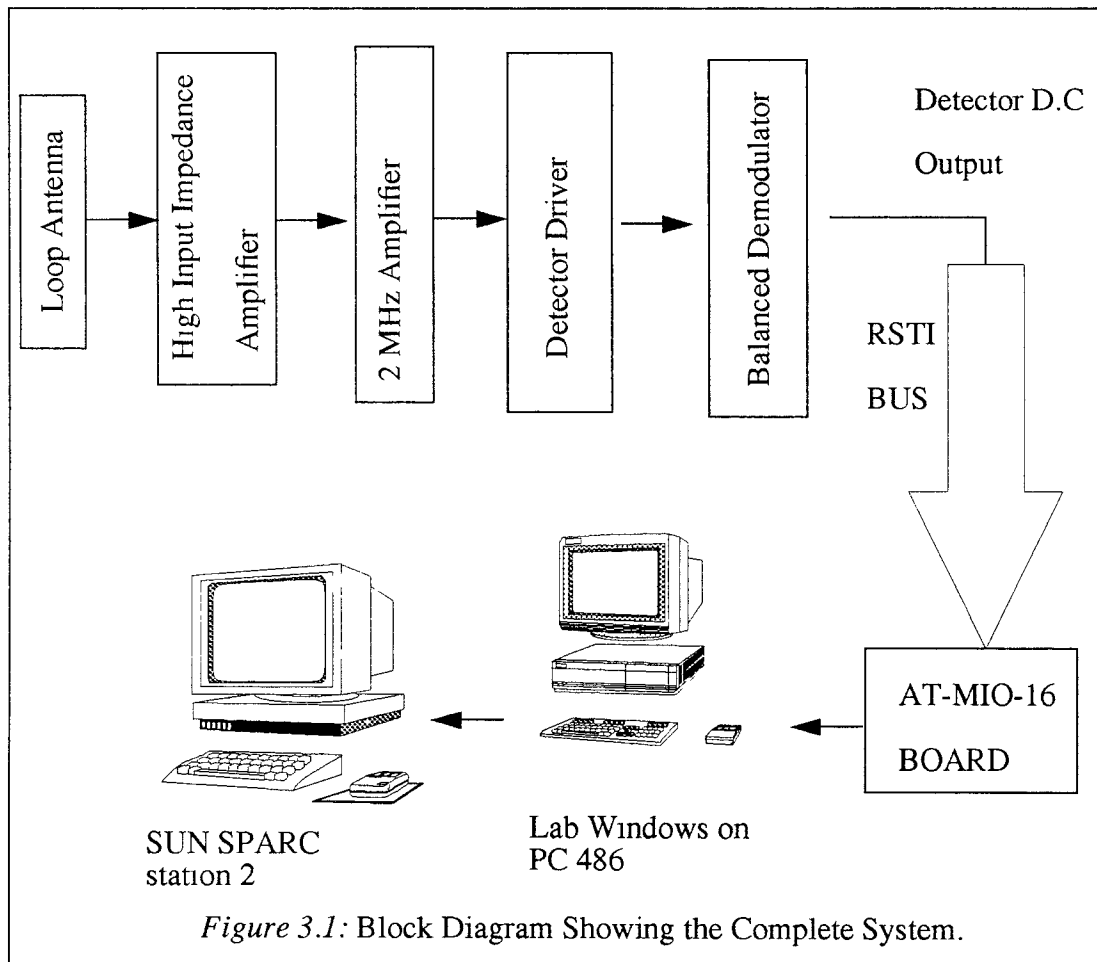
3.3 The Principle of Triangulation

The position of an *rf* electromagnetic source radiating at a fixed frequency may be computed by applying the triangulation principle either on amplitude measurements or on phase measurements made at several different receivers. Triangulation based on the phase shift measurements is similar to many radar based position tracking and homing systems. The accuracy of position measurement depends on the accuracy with which phase shift can be measured. Another critical requirement is that of a stable frequency of transmission. For our application, measurements requiring better than 1/4" resolution demands phase shift measurements to an accuracy of better than one nano-second. Such systems would be expensive and unrealistic. A system based on amplitude triangulation would ideally require an isotropic radiator. Unfortunately, all practical sources are anisotropic. Moreover, there are problems associated with reflected radio-waves from metallic surfaces and the design of receiving antennae adequately. Hence, for the spatial scales in our applications, it is best if the triangulation is based on purely magnetic coupling between the transmitter and the receiver, with the transmitter frequency selected in order to minimize the effect of standing waves and reflections. To illustrate this concept, consider a small loop of wire carrying a sinusoidal current $I=I_0 sin(t)$. This loop would behave like a small magnetic dipole, and would have a magnetic field around it. If another loop is placed anywhere in this magnetic field, then there would be a corresponding voltage induced in it. The strength of the induced voltage would depend on the relative position and orientation between the loops. If the space scale is comparable to the size of the dipole, then the induced voltage would be a function only of the magnetic field strength and the relative position and orientation.

To understand the principle of this system, consider for the time being that the motion of the tracking sphere in space is restricted such that its orientation is always the same, and is known. Also, assume that there is only one transmitter in the tracking

sphere. In that configuration, for a given value of the induced voltage in a receiver, the sphere must be on a hypothetical constant-voltage surface around the receiver. By finding the intersection of these surfaces of three or more receivers, one may find the location of the transmitting sphere. In general, there may be more than one solution to the intersection problem. One can use, however, the motion continuity condition to resolve this situation.

In practice, this approach is quite involved and becomes even more complex when the orientation also changes as in real experiments. Thus, the problem of determining the position and orientation from the measured voltages information, also called the "Inverse Solution problem", is not a trivial task. Several approaches to solve this problem are discussed in the next section.

3.4 The Problem of Inverse Solution

Every three dimensional location of the tracking sphere can be represented by six parameters, three for the position, and three for the orientation. Let x be the location vector, $x = (x_1, x_2, x_3, x_4, x_5, x_6)$, of a transmitter i inside the tracking sphere with respect to an external receiving antenna j. Then V_{ij}, the voltage induced in receiver j due to transmitter i, can be expressed as a function of x:

$$V_{ij} = g_{ij}(x) \tag{3.4.1}$$

If there are p transmitters and q receivers, then (3.4.1) represents $m = pq$ nonlinear equations in six positional variables x. Here, we adopt the following terminology: For a given value of x, determination of V_{ij} is called the *Forward Solution*, while determination of x given the values of V_{ij} is called the *Inverse Solution*.

One must have the Forward Solution available before it can be inverted. From the practical point of view, obtaining a good Forward Solution itself is a challenging problem. There are two cases for the Forward Solution; when a closed form solution exists -

that is, $g_{ij}(x)$ are known functions, and when a closed form solution does not exist or is very difficult to obtain. Closed form equations for the functions g_{ij} can be obtained for simple transmitter-receiver geometries using the governing laws for magnetic induction (Griffiths 1989). For complex geometries, closed form equations are difficult, if not impossible to obtain. In that case, the function map (fitting a function to discrete points) can be generated empirically by obtaining the experimental data. Therefore, the problem of Forward Solution can be solved either way.

The Inverse Solution to equation (3.4.1) can be obtained in several ways, at least in principle. We consider three different approaches to this problem. The first approach is based on a brute-force method, where an exhaustive search of the solution space is conducted in the vicinity of the expected solution. The advantage of this approach is that it would work for both the cases of Forward Solution, that is closed-form or empirical map. The second approach is a refinement of the first approach, where interpolation is used based on regression analysis. This approach is specifically suitable when closed-form equations are not available for the Forward Solution. Using regression analysis, the empirical map can be fit to a polynomial form of equations and inverted simultaneously (Rosato *et al* 1991). The third approach is based on numerically solving the system of equations represented by (3.4.1) with V_{ij} known, and x unknown. This approach is suitable when the Forward Solution is known in closed-form. Details of the first two approaches are considered next. The third approach is discussed in detail in the next Chapter.

3.4.1 The Look-Up Table Method

The first approach uses the empirical map of the Forward Solution. In general, it is a map from six dimension sphere parameters space to m dimension voltage space. In other words, it is a map of x onto $V = \{V_{ij}\}$ so that for each value of x, there is a corresponding

unique value of the voltage vector V. Let there be N locations in the map, each being a node point in the experimental space, so that the map can be described as $\{ (x_k | V_k), k=1,..., N \}$. If the map has a sufficiently high-resolution, then for a given value of the received voltages V_*, the Inverse Solution is $x_* = x_k$, such that the magnitude of (V_*-V_k) is a minimum. This technique would work well only if the search space is limited to the immediate vicinity of the expected value of x_*. In practice, one would know the location of the tracking sphere in the beginning of the experiment. Then, for the subsequent measurements, the sphere would not have moved too far, so that the search space can be limited. This approach is simple, but requires a large empirical map of equation (3.4.1). Consider for example, the experimental space of 12"x12"x72", which represents a typical chute flow experimental space. For an accuracy of 0.1" in position and 10 degrees in angle, one would require over 4.5x1011 nodes in the table. This is a gigantic table requiring a formidable amount of memory. Moreover, if this table, that is, the map, is to be created empirically, the amount of time required to build it is unrealistic.

3.4.2 Interpolation Using Regression Analysis

The above strategy can be improved by using interpolation techniques. A regression model is constructed by using a simulated empirical map of equation (3.4.1), and by treating voltages, that is, V_{1j} as independent variables, and x as dependent variables. By setting up regression functions this way, the inverted solution of (3.4.1) is directly obtained. As can be realized, even with the use of interpolation, the empirical map is still required. However, the resolution needed is lower. Several regression models were tried on a subset of the experimental space. One example is where every length variable was divided in steps of 3" and every angle variable was divided in steps of 60 degrees. Only one third of the space was considered. This resulted in a map having 27,648 nodes, each

node corresponding to a six-dimensional position storing m voltage values. The resulting regression model includes all these nodal values. The accuracy obtained was not satisfactory, for example, predicted location was off by almost two inches on an average. Improvements can be made by trying out many different regression models, but unfortunately, it is difficult to develop a systematic procedure for selecting different regression models.

3.4.3 The Numerical Solution Method

The Numerical Solution method is the third and most viable approach. The only issue is whether a numerical procedure can be developed to solve the system of m equations for six unknowns. In the case when the closed-form equations are not available, one may use "approximate" equations which may be empirically fine-tuned. This approach also tells that in order to solve for the six unknowns, m, the number of equations must be greater than or equal to six ($m \geq 6$). Since the system of equations is nonlinear, it is better if more equations are available than the number of unknowns. For example, one may use three transmitters and three receivers to obtain nine equations, or one may use two transmitters and six receivers to obtain twelve equations. The direct mathematical analysis to determine whether the system of equations is independent is not trivial. It is easier to use physical insight in positioning the transmitters and receivers so as to increase the probability of the equations being independent, or not completely redundant. For example, the transmitters can be positioned orthogonally to reduce redundancy of information. In the next chapter numerical procedures are described to solve the nonlinear equations represented by (3.4.1).

To solve the inverse problem by the numerical methods, there needs to be a "theoretical model" to relate the positions and orientations to the expected measured voltages at the receiving antennae. Hence, we present below three models which have been

developed for various geometries of the transmitting and receiving antennae (Parasar A. 1992; Rosato *et al.* 1992). A rigorous discussion of the derivation, the assumptions and the limitations of these models is beyond the scope of this thesis.

3.5 Transfer Function Between the Transmitters and the Receivers

3.5.1 Model #1

Euler angles: α, β, γ

Figure 3.2: Definition of Relative Location Parameters for Model #1 and Model #2.
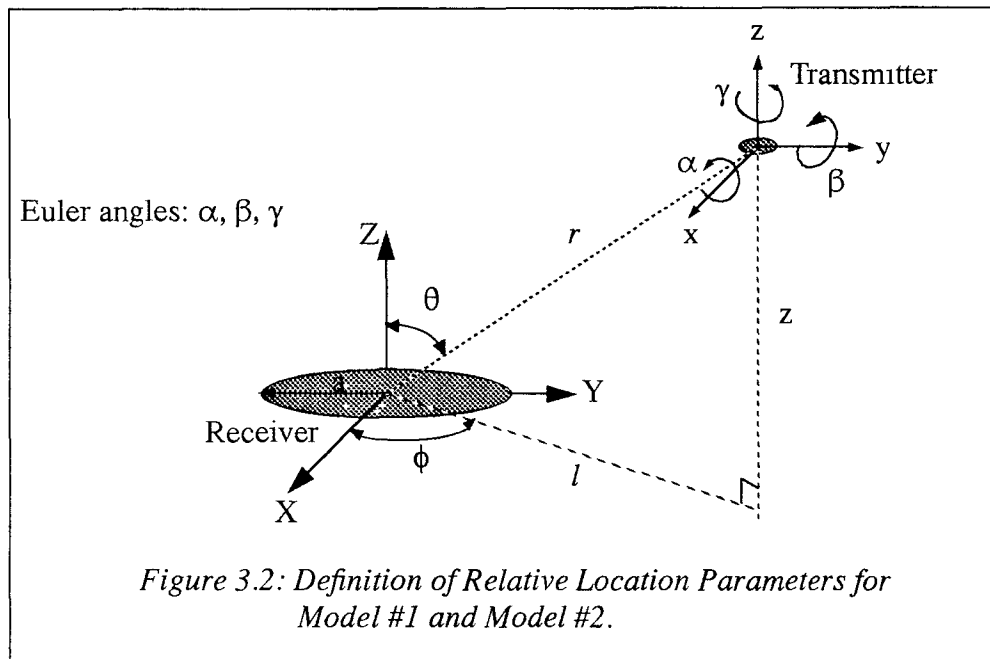
The model #1 is a simplified model to obtain a very crude functional form with which to begin. The assumptions made are that a point receiver is used, the transmitter is of finite size and of circular shape, and that the distance between the transmitter and the receiver is much greater than the dimensions of the transmitter or receiver. Figure (3.2) shows a transmitter-receiver pair with the various relative location (position and orientation) parameters defined for use in model #1 and model #2.

The voltage induced in receiver j due to transmitter i according to the first model is given by

$$V_{ij} = \frac{\mu I_i A_i A_j N_i N_j \omega}{4 \pi r_{ij}^3} [2\cos\theta_{ij}\cos\beta_{ij} + \sin\theta_{ij}\sin\beta_{ij}]$$

where μ is the permeability of the transmission medium,

I_i is the current in the transmitter i,

A_i is the area of the transmitter i,

A_j is the area of the receiver j,

N_i is the number of turns of the ith transmitter coil,

N_j is the number of turns of the jth receiver coil,

ω is equal to $2\pi f$,

f is the frequency of transmitter i,

r_{ij} is the distance between transmitter i and receiver j,

θ_{ij} and β_{ij} are angles as defined in Figure (3.2).

Model #1 is developed for the case where the tracking sphere is to have three orthogonal transmitters. The transmitters were intended to be IC monolithic chips with integral loop antennae. Each antenna transmits at a different frequency at around 220MHz approximately. The Colpitts oscillator is used for the transmitting chips. Three or more orthogonally placed external circular receiving antennae are to be used. While the above model has the advantage of simplicity, the receiver loops are large, indicating that the assumption of point receiver is invalid. Mathematical difficulties associated with consideration of finite-sized receiver loops are eliminated by applying the *principle of reciprocity* which states that in a linear network the source and the effect are interchangeable. Therefore the receiver is assumed to be the "transmitter" and the transmitter the "receiver" in deriving the transfer function for Model #2.

3.5.2 Model #2

Model #2 assumes that we have a point transmitter and a circular receiver of finite diameter with the dimensions of the receiver being of the same order as of the distance between the transmitter and the receiver. The voltage induced in receiver j due to transmitter i according to the second model is given by

$$V_{ij} = \frac{N_i N_j \omega A_i I_i}{r_{ij}} \left[\sin\beta_{ij} \sin\gamma_{ij} \frac{\partial}{\partial r_{ij}}(r_{ij}A_\varphi) - \frac{\cos\beta_{ij}\cos\gamma_{ij}}{\sin\theta_{ij}} \frac{\partial}{\partial\theta_{ij}}(A_\varphi \sin\theta_{ij}) \right],$$

where N_1 is the number of turns of the ith transmitter coil,

N_j is the number of turns of the jth receiver coil,

ω is equal to $2\pi f$,

f is the frequency of transmitter i,

A_1 is the area of the transmitter i,

I_1 is the current in the transmitter i,

r_{1j} is the distance between transmitter i and receiver j,

θ_{1j}, β_{1j}, γ_{1j} are angles as defined in Figure (3.2).

A_φ is the magnetic vector potential given by

$$A_\varphi = \frac{\mu}{k\pi}\sqrt{\frac{a}{l}} \left\{ \left((\frac{2-k^2}{2k}) \left[\int_0^{\pi/2} \frac{d\gamma_{ij}}{\sqrt{1 - (k\sin\gamma_{ij})^2}} \right] \right) - \int_0^{\pi/2} \sqrt{(1 - (k\sin\gamma_{ij})^2)}\, d\gamma_{ij} \right\}$$

in which μ is the permeability of the transmission medium,

$$l = r_{ij}\sin\theta_{ij}, \; z = r_{ij}\cos\theta_{ij} \text{ and } k^2 = \frac{4al}{z^2 + (a+l)^2}$$

where the lengths a and l are as defined in Figure (3.2).

The magnetic vector potential, A_φ is the sum of two elliptic integrals of the first and the second kind. For the second model, three orthogonal transmitters and three or more orthogonally placed external receiving antennae are used.

3.5.3 Model #3

The use of a circular loop on a rectangular shaped chute results in a "dead space" with a resulting reduction in the signal to noise ratio. Therefore, model #3 uses rectangular shaped receiving loop which can follow the contours of the chute or the rectangular experimental space very closely. By proper positioning and size adjustments it can be made more sensitive to movements in one direction only.

Figure 3.3: Configuration for Model #3.
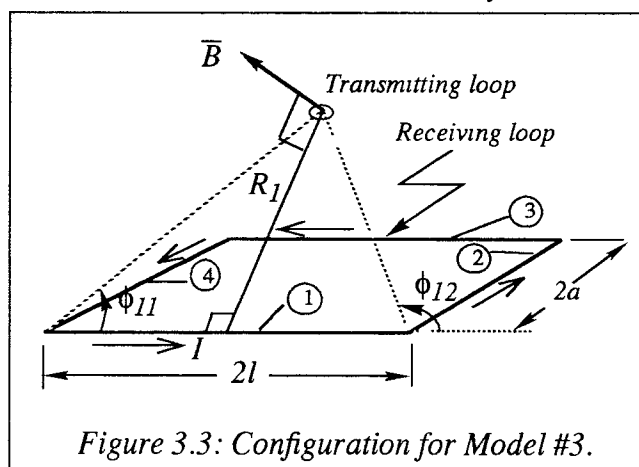
Since the monolithic transmitters did not provide adequate signal levels and since their operating frequency of 220 MHz was not suitable for measurements because of room resonances and standing waves, the operating frequency was changed to the lower frequency of 2Mhz where the wavelength is much larger than the dimensions of the room. Moreover, the "noise" due to other stray transmissions from various other sources is minimal at this frequency. But, in order to maintain the same signal strength the capacitance has to be much larger for the lower frequency. This results in building the transmitter using discrete components as such high capacitances cannot be supported on monolithic IC chips. Only one transmitter along with its battery can be accommodated inside the tracking sphere of 3/4" due to the use of discrete components which require more space than the IC chips.

With respect to computations the third model is simpler due to the absence of ellip-

tic integral terms of model #2. The rectangular receiving loop is treated as four independent filament sources and the Superposition Theorem is applied to get the composite response by summing up the individual responses. The voltage induced in receiver j due to transmitter i according to the third model is given by

$$V_{ij} = -N_i N_j \omega A_i \, [B_x \cos\alpha_{ij} + B_y \cos\beta_{ij} + B_z \cos\gamma_{ij}]$$

where N_i is the number of turns of the ith transmitter coil,

N_j is the number of turns of the jth receiver coil,

ω is equal to $2\pi f$,

f is the frequency of transmitter i,

A_i is the area of the transmitter i,

α_{ij}, β_{ij} and γ_{ij} are the direction cosine angles which the normal to the

plane of the transmitter coil makes with the x, y and z axes respectively

and are related by

$$\cos^2\alpha_{ij} + \cos^2\beta_{ij} + \cos^2\gamma_{ij} = 1$$

B_x, B_y, B_z are x, y and z components respectively of the magnetic field

density $\vec{B}$, given by

$$\vec{B} = \sum_{k=1}^{4} \left[\left(\frac{\mu I_i}{4\pi R_k} \right) (\cos\varphi_{k1} - \cos\varphi_{k2}) \right] \hat{\theta}_k$$

in which μ is the permeability of the transmission medium,

I_i is the current in the transmitter i,

R_k, $\cos\varphi_k$ and $\hat{\theta}_k$ are functions of the position (x_{ij}, y_{ij}, z_{ij}), denoted

below as (x, y, z), and the loop dimensions of the receiver (l_j, a_j), denoted

as (l, a) for simplicity:

$$\cos\varphi_{11} = \frac{x+l}{\sqrt{(x+l)^2 + (y+a)^2 + z^2}}$$

$$\cos\varphi_{12} = \frac{x-l}{\sqrt{(x-l)^2 + (y+a)^2 + z^2}}$$

$$\cos\varphi_{21} = \frac{y+a}{\sqrt{(x-l)^2 + (y+a)^2 + z^2}}$$

$$\cos\varphi_{22} = \frac{y-a}{\sqrt{(x-l)^2 + (y-a)^2 + z^2}}$$

$$\cos\varphi_{31} = \frac{-(x-l)}{\sqrt{(x-l)^2 + (y-a)^2 + z^2}}$$

$$\cos\varphi_{32} = \frac{-(x+l)}{\sqrt{(x+l)^2 + (y-a)^2 + z^2}}$$

$$\cos\varphi_{41} = \frac{-(y-a)}{\sqrt{(x+l)^2 + (y-a)^2 + z^2}}$$

$$\cos\varphi_{42} = \frac{-(y+a)}{\sqrt{(x+l)^2 + (y+a)^2 + z^2}}$$

$$R_1 = \sqrt{(y+a)^2 + z^2}$$

$$R_2 = \sqrt{(x-l)^2 + z^2}$$

$$R_3 = \sqrt{(y-a)^2 + z^2}$$

$$R_4 = \sqrt{(x+l)^2 + z^2}$$

$$\hat{\theta}_1 = \frac{-z}{\sqrt{z^2 + y^2 + a^2 + 2ya}}\hat{y} + \frac{(y+a)}{\sqrt{z^2 + y^2 + a^2 + 2ya}}\hat{z}$$

$$\hat{\theta}_2 = \frac{z}{\sqrt{z^2 + x^2 + l^2 - 2xl}}\hat{y} + \frac{(-x+l)}{\sqrt{z^2 + x^2 + l^2 - 2xl}}\hat{z}$$

$$\hat{\theta}_3 = \frac{z}{\sqrt{z^2 + y^2 + a^2 - 2ya}}\hat{y} + \frac{(-y+a)}{\sqrt{z^2 + y^2 + a^2 - 2ya}}\hat{z}$$

$$\hat{\theta}_4 = \frac{-z}{\sqrt{z^2 + x^2 + l^2 + 2lx}}\hat{y} + \frac{(x+l)}{\sqrt{z^2 + x^2 + l^2 + 2lx}}\hat{z}$$

3.6 Practical Considerations

There are many aspects to the development of this tracking system:

(1)   Design and construction of the miniature transmitters,

(2)   Derivation of the magnetic radiation pattern for these transmitters (Rosato *et al.*, 1991, 1992),

(3)   Design and construction of the receiving antennae,

(4)   Design and construction of the amplifiers, the bandpass filters and the detectors,

(5)   Measurement of the magnetically induced voltages in the receivers using the data acquisition system,

(6) Packaging the electronics and the power source (battery) inside a sphere of the smallest size possible. Indeed, the size of the battery is the limiting factor for the dimensions of the sphere. The sphere has to be dynamically balanced

(7) Design and construction of the experimental space (chute, hopper, etc.)

(8) Computing the position and the orientation of the particle from the induced voltages,

(9) Simulating the complete trajectory of the particle.

Each of the above mentioned tasks is a crucial part in the development of the complete system and is not trivial. The details of each task is not provided here and is beyond the scope of this thesis, since we focus on computing the location of the sphere from the induced voltage values.

CHAPTER 4

THE NONLINEAR OVERDETERMINED SYSTEM

4.1 Introduction

In the previous chapter a technique based on magnetic coupling to track the motion of a particle in three dimensions was discussed. Three approaches to the problem of inverse solution were described. In this chapter we consider the third and most viable method to be used for the problem of *inverse solution*. The methods of solution employed to numerically solve the system of nonlinear equations is discussed. We first discuss the general nonlinear least-squares problem which leads to the *Gauss-Newton* and *Levenberg-Marquardt* algorithms. Line searches and the model trust region approaches are described in terms of their ability to achieve global convergence (Byrd 1987). The theory and the implementation details of the algorithms are provided. The readers interested in the proofs are referred to the references provided.

Since the number of nonlinear equations is greater than the number of unknown parameters, the system is *over-determined*, and it is usually not possible to obtain an exact solution. Therefore a *least squares solution* is sought. The numerical methods of solution to the overdetermined system of nonlinear equations take advantage of the special structure of the objective function to be minimized (Fletcher 1987).

4.2 The Nonlinear Least-Squares Problem

The nonlinear least-squares problem is

$$\underset{x \in R^n}{\text{minimize}} \, f(x) \;=\; \frac{1}{2} R(x)^T R(x) \;=\; \frac{1}{2} \sum_{i=1}^{m} r_i(x)^2 \qquad (4.2.1)$$

where $m > n$, the nonlinear *residual function* $R : R^n \rightarrow R^m$ has components

$r_k(x)$ (k = 1, 2,. . . , m) which are given by

$$r_k(x) = g_k(x) - v_k, \quad k = 1, 2, \ldots, m$$

In the above equation, a single subscript k is used for g which is denoted as g_{ij} in equation (3.4.1). Here v_k is the experimentally measured voltage values. Subscript k goes from 1 to m, which is equivalent to i ranging from 1 to p and j ranging from 1 to q, since $m = p \times q$. An attempt is made to fit the data (v_i), $i = 1,\ldots, m$, with a model $M(x)$, previously discussed in Chapter 3, which is nonlinear in x. In this case $r_i(x) = M(x_i) - v_i$, and the nonlinear least-squares problem consists of choosing x so that the fit is as close as possible in the sense that the sum of the squares of the residuals - $r_i(x)$'s is minimized. In general, there are more equations m than unknowns n. In the application considered here m is greater than or equal to nine and n is equal to six.

When $m = n$, it includes as a special case solving a system of nonlinear equations, and for any value of m, it is just a special case of unconstrained minimization.

The first-derivative matrix of $R(x)$ is simply the Jacobian matrix $J(x) \in R^{m \times n}$, where $J(x)_{ij} = \partial r_i(x) / \partial x_j$. Thus an affine model or linear approximation of $R(x)$ around a point x_c is

$$M_c(x) = R(x_c) + J(x_c)(x - x_c),$$

The first derivative of $f(x) = \frac{1}{2}R(x)^T R(x)$ is given by its gradient, i.e.,

$$\nabla f(x) = \sum_{i=1}^{m} r_i(x) \cdot \nabla r_i(x) = J(x)^T R(x).$$

Similarly, the second derivative is

$$\nabla^2 f(x) = \sum_{i=1}^{m} (\nabla r_i(x) \cdot \nabla r_i(x)^T + r_i(x) \cdot \nabla^2 r_i(x)) = J(x)^T J(x) + S(x)$$

where

$$S(x) = \sum_{i=1}^{m} r_i(x) \cdot \nabla^2 r_i(x)$$

denotes the second-order information in $\nabla^2 f(x)$. Thus the quadratic model of $f(x)$ around x_c is

$$m_c(x) = f(x_c) + \nabla f(x_c)^T (x - x_c) + \frac{1}{2}(x - x_c)^T \nabla^2 f(x_c)(x - x_c)$$

$$= \frac{1}{2}R(x_c)^T R(x_c) + R(x_c)^T J(x_c)(x - x_c)$$

$$+ \frac{1}{2}(x - x_c)^T (J(x_c)^T J(x_c) + S(x_c))(x - x_c). \qquad (4.2.2)$$

This is a specialization of the Taylor series quadratic model for minimization of objective functions of form (4.2.1).

Using (4.2.2), Newton's method applied to (4.2.1) is

$$x_+ = x_c - [J(x_c)^T J(x_c) + S(x_c)]^{-1} J(x_c)^T R(x_c). \qquad (4.2.3)$$

where x_+ is the next iterate and x_c is the current iterate. (4.2.3) is locally q-quadratically convergent under standard assumptions (Dennis and Schnabel 1983). But, the problem with the full Newton approach is that $S(x)$ is usually either unavailable or inconvenient to obtain, and it is too computationally intensive to approximate by finite differences. Although frequently used in practice, a secant approximation to all of $\nabla^2 f(x)$ is undesirable, because the portion $J(x)^T J(x)$ of $\nabla^2 f(x)$ is already readily available, since $J(x)$ must be calculated analytically or by finite differences to calculate $\nabla f(x)$.

In discussing the various methods for nonlinear least-squares, we will want to distinguish between zero-residual, small-residual and large-residual problems. These terms refer to the value of $R(x)$ [or $f(x)$] at the minimizer x_* of (4.2.1). A problem for which $R(x_*) = 0$ is called a zero-residual problem; this means that the model $m(x)$ fits the data v_i exactly at each data point.

4.3 Gauss-Newton Type Methods

The first method for solving the nonlinear least-squares problem comes from using the affine model of $R(x)$ around x_c,

$$M_c(x) = R(x_c) + J(x_c)(x - x_c), \qquad (4.3.1)$$

where $M_c : R^n \to R^m$ and $m > n$.

We cannot in general expect to find an x_+ for which $M_c(x_+) = 0$, since this is an overdetermined system of linear equations. However, a logical way to use (4.3.1) to solve the nonlinear least-squares problem is to choose the next iterate x_+ as the solution to the linear least-squares problem

$$\text{minimize} \ \frac{1}{2} \| M_c(x) \|_2^2 = \hat{m}_c(x). \qquad (4.3.2)$$

Assuming that $J(x_c)$ has full column rank, then the solution to (4.3.2) is

$$x_+ = x_c - [J(x_c)^T J(x_c)]^{-1} J(x_c)^T R(x_c). \qquad (4.3.3)$$

The iterative method that consists of using (4.3.3) at each iteration is called the *Gauss-Newton method*.

Comparing (4.2.3) and (4.3.3), the two equations differ only in the term $S(x)$, included by Newton's method in the second-derivative matrix $J(x_c)^T J(x_c) + S(x_c)$ but omitted by the Gauss-Newton method. Equivalently, the only difference between the quadratic model $m_c(x)$, (4.2.2), from which the Newton's method was derived, and the quadratic model $\hat{m}_c(x)$, (4.3.2), that gives the Gauss-Newton method, is that the portion $S(x_c)$ of $\nabla^2 f(x_c)$ is omitted from $\hat{m}_c(x)$. Since Newton's method is locally q-quadratically convergent under standard assumptions, the success of the Gauss-Newton method depends on whether the omitted term $S(x_c)$ is important - that is, whether it is a large part of $\nabla^2 f(x_c) = J(x_c)^T J(x_c) + S(x_c)$. If $S(x_*) = 0$, then the Gauss-Newton method is also q-quadratically convergent. This occurs when $R(x)$ is linear, or when we have a zero-residual problem. If $S(x_*)$ is small relative to $J(x_*)^T J(x_*)$, the Gauss-Newton

method is locally *q*-linearly convergent. However, if $S(x_*)$ is too large, the Gauss-Newton method may not be locally convergent at all. Thus the speed of convergence of the Gauss-Newton method decreases as the relative nonlinearity or the relative residual size of the problem increases; if either of these is too large, the method may not converge at all. Alternatively, the larger $S(x_*)$ is in comparison to $J(x_*)^T J(x_*)$, the worse the Gauss-Newton method is likely to perform.

The Gauss-Newton is in a descent direction. But the Gauss-Newton method may take bad steps by taking steps that are too long, but in the correct direction. Two ways of improving the Gauss-Newton algorithm are by using it with a line search or with a trust region strategy. These two approaches lead to two algorithms that are used in practice.

The algorithm that uses the Gauss-Newton method with a line search is simply

$$x_+ = x_c - \lambda_c (J(x_c)^T J(x_c))^{-1} J(x_c)^T R(x_c), \qquad (4.3.4)$$

where λ_c is chosen by the Line search method.

Equation (4.3.4) is referred to as the *damped Gauss-Newton* method (Dennis & Schnabel 1983). Since the damped Gauss-Newton method always takes descent steps that satisfy the line-search criteria, it is locally convergent on almost all nonlinear least-squares problems, including large-residual or very nonlinear problems. However, it may still be very slowly convergent on the problems that the Gauss-Newton method had trouble with. Also, the damped Gauss-Newton algorithm still is not well defined if $J(x_c)$ doesn't have full column rank.

Another modification of the Gauss-Newton algorithm is to choose x_+ by the trust

region approach:

$$\text{minimize } \| R(x_c) + J(x_c)(x_+ - x_2) \|_2$$

$$\text{subject to } \| x_+ - x_c \|_2 \leq \delta_c. \tag{4.3.5}$$

The solution to (4.3.5) is

$$x_+ = x_c - [J(x_c)^T J(x_c) + \mu_c I]^{-1} J(x_c)^T R(x_c), \tag{4.3.6}$$

where $\mu_c = 0$ if $\delta_c \geq \| [J(x_c)^T J(x_c)]^{-1} J(x_c)^T R(x_c) \|_2$ and $\mu_c > 0$ otherwise.

Formula (4.3.6) is referred to as the *Levenberg-Marquardt* algorithm (Dennis & Schnabel 1983; Levenberg 1944; Marquardt 1963; More 1977).

The local convergence properties of the Levenberg-Marquardt method are similar to those of the Gauss-Newton method. Several factors make the Levenberg-Marquardt algorithms preferable to the damped Gauss-Newton algorithms. One is that the Levenberg-Marquardt method is well defined even when $J(x_c)$ doesn't have full column rank. Another is that when the Gauss-Newton step is much too long, the Levenberg-Marquardt step is close to being in the steepest - descent direction $-J(x_c^T)R(x_c)$ and is often superior to the damped Gauss-Newton step.

4.4 Globally Convergent Modifications of Newton's Method

Newton's method is locally q-quadratically convergent. This means that when the current solution approximation is good enough, it will be improved rapidly and with relative ease. Unfortunately, it is not unusual to expend significant computational effort in getting close enough. In addition, the strategies for getting close constitute the major part of the program and the programming effort, and they can be sensitive to small differences in implementation.

Two major ideas for proceeding when the Newton step is unsatisfactory are used. The first major global approach, is modern versions of the traditional idea of backtrack-

ing along the Newton direction if a full Newton step is unsatisfactory. The second major approach is based on estimating the region in which the local model, underlying Newton's method can be trusted to adequately represent the function and taking a step to approximately minimize the model in this region.

4.4.1 The Quasi - Newton Framework

The basic idea in forming a successful nonlinear algorithm is to combine a globally convergent strategy with a fast local strategy in a way that derives the benefits of both. The most important point is to try Newton's method, or some modification of it, first at each iteration. If it seems to be taking a reasonable step - use it. If not, fall back on a step dictated by a global method. Such a strategy will always end up using Newton's method close to the solution and thus retain its fast local convergence rate. An algorithm that takes this approach is *quasi-Newton* (Dennis and More 1977).

4.4.2 Descent Directions

The basic idea of a global method is to choose a direction p from the current point x_c in which f decreases initially, and a new point x_+ in this direction from x_c such that $f(x_+) < f(x_c)$. Such a direction is called a descent direction. Mathematically, p is a descent direction from x_c if the directional derivative of f at x_c in the direction p is negative - that is,

$$\nabla f(x_c)^T p < 0.$$

Newton direction $s^N = -H_c^{-1} \nabla f(x_c)$, where H_c is either $\nabla^2 f(x_c)$ or an approximation to it, is in a descent direction if and only if,

$$\nabla f(x_c)^T s^N = -\nabla f(x_c)^T H_c^{-1} \nabla f(x_c) < 0,$$

which is true if H_c^{-1} or, equivalently, H_c is positive definite.

The quasi-Newton algorithm is far more efficient than the method of steepest descent. However, when the global strategy must take steps much smaller than the Newton step, it may take steps in, or close to, the steepest descent direction.

4.5 Line Searches

The first strategy for proceeding from a solution outside the convergence region of Newton's method is the method of line searches.

The line search method deals exclusively with the problem of finding an acceptable step length in a given direction of search with the assumptions that the direction would be the quasi-Newton direction, and that the full quasi-Newton step would always be the first trial step. If the full quasi-Newton step is unsatisfactory, it indicates that the quadratic model does not adequately model f in a region containing the full quasi-Newton step. The line search algorithms retain the same step direction and choose a shorter step length. This new length is determined by building a new one-dimensional quadratic or cubic model, based only on function and gradient information in the quasi-Newton direction.

The idea of a line-search algorithm is simple: given a descent direction p_k, we take a step in that direction that yields an "acceptable" x_{k+1}. That is,

at iteration k:

calculate a descent direction p_k,

set $x_{k+1} = x_k + \lambda_k p_k$ for some $\lambda_k > 0$ that makes x_{k+1}

an acceptable next iterate (4.5.1)

The term "line search" refers to a procedure for choosing λ_k in (4.5.1)

The common procedure is to try the full quasi-Newton step first and, if $\lambda_k = 1$ fails to satisfy the criterion in use, to backtrack in a systematic way along the direction defined by that step. Computational experience has shown the importance of taking a

full quasi-Newton step whenever possible. Failure to do so leads to forfeiture of the advantage of Newton's method near the solution.

While no step - acceptance rule is always optimal it is reasonable to expect that $f(x_{k+1}) < f(x_k)$. Armijo (1966) and Goldstein (1967) have suggested conditions to remove problems associated with very small decreases in f values relative to the lengths of the steps and with steps that are too small relative to the initial rate of decrease of f:

$$f(x_+) \leq f(x_c) + \alpha \nabla f(x_c)^T (x_+ - x_c) \tag{4.5.2}$$

Algorithm (4.5): Backtracking Line-Search Framework

Given $\alpha \in (0, \frac{1}{2})$, $0 < l < u < 1$

$\lambda_k = 1$;

while $f(x_k + \lambda_k p_k) > f(x_k) + \alpha \lambda_k \nabla f(x_k)^T p_k$, do

$\lambda_k = \rho \lambda_k$ for some $\rho \in [l, u]$;

(*ρ is chosen anew each time by the line search*)

$x_{k+1} = x_k + \lambda_k p_k$;

α is usually set quite small ($\alpha = 10^{-4}$) so that hardly more than a decrease in function value is required.

Strategy for reducing λ_k (choosing ρ)

$$\text{Let } \hat{f}(\lambda) = f(x_k + \lambda p_k),$$

the one dimensional restriction of f to the line through x_k in the direction p_k. If we need to backtrack we use our most current information about $\hat{f}$ to model it, and then take the value of λ that minimizes this model as our next value of λ_k in Algorithm (4.5). Initially, we have two pieces of information about $\hat{f}(\lambda)$,

$$\hat{f}(0) = f(x_k) \text{ and } \hat{f}'(0) = \nabla f(x_k)^T p_k. \tag{4.5.3}$$

After calculating $f(x_k + p_k)$, we also know that

$$\hat{f}(1) = f(x_k + p_k), \tag{4.5.4}$$

so if $\hat{f}(1) > \hat{f}(0) + \alpha f'(0)$, we model $\hat{f}(\lambda)$ by the one-dimensional quadratic satisfying (4.5.3) and (4.5.4).

$$\hat{m}(\lambda) = [\hat{f}(1) - \hat{f}(0) - \hat{f}'(0)] \lambda^2 + \hat{f}'(0) \lambda + \hat{f}(0),$$

and calculate the point

$$\hat{\lambda} = \frac{-\hat{f}'(0)}{2 [\hat{f}(1) - \hat{f}(0) - \hat{f}'(0)]}$$

for which $\hat{m}'_q(\hat{\lambda}) = 0$.

Thus $\hat{\lambda}$ minimizes $\hat{m}'_q(\hat{\lambda})$ and also $\hat{\lambda} > 0$

Therefore we take $\hat{\lambda}$ as our new value of λ_k. We impose a lower bound of $l = 1/10$ and a upper bound of $u = 1/2$ on the first value of ρ. This means that at the first backtrack at each iteration, if $\hat{\lambda} \leq 0.1$, then $\lambda_k = \frac{1}{10}$. If $\hat{f}(\lambda_k) = f(x_k + \lambda_k p_k)$, does not satisfy

$$f(x_k + \lambda_k p_k) \leq f(x_k) + \alpha \lambda_k \nabla f(x_k)^T p_k,$$

we use a cubic model of $\hat{f}$, $\hat{m}_{cu}(\lambda)$ and calculate the value of λ at which $\hat{m}_{cu}(\lambda)$ has its local minimizer. If (λ_{prev}) and (λ_{2prev}) are the last two previous values of λ_k, then

$$\hat{m}_{cu}(\lambda) = a\lambda^3 + b\lambda^2 + \hat{f}'(0) \lambda + \hat{f}(0),$$

where

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\lambda_{prev} - \lambda_{2prev}} \times \begin{bmatrix} \dfrac{1}{\lambda_{prev}^2} & \dfrac{-1}{\lambda_{2prev}^2} \\ \dfrac{-\lambda_{2prev}}{\lambda_{prev}^2} & \dfrac{\lambda_{prev}}{\lambda_{2prev}^2} \end{bmatrix} \begin{bmatrix} \hat{f}(\lambda_{prev}) - \hat{f}(0) - \hat{f}'(0) \lambda_{prev} \\ f(\lambda_{2prev}) - \hat{f}(0) - \hat{f}'(0) \lambda_{2prev} \end{bmatrix}$$

Its local minimizing point $\hat{\lambda}$ is

$$\frac{-b + \sqrt{b^2 - 3a\hat{f}'(0)}}{3a}$$

The lower bound $l = 1/10$ and the upper bound $u = 1/2$ are imposed. That is, if $\hat{\lambda} > \frac{1}{2} \lambda_{prev}$, then $\lambda_k = \frac{1}{2} \lambda_{prev}$ and if $\hat{\lambda} < \frac{1}{10} \lambda_{prev}$, then $\lambda_k = \frac{1}{10} \lambda_{prev}$

Thus a quadratic model is used for backtracking at the first iteration and if needed, a cubic model is used for backtracking for subsequent backtrack during the current iteration.

Also, a minimum step length is imposed called *minstep*. If condition (4.5.2) is not satisfied, but $\|\lambda_k p_k\|_2$ is smaller than *minstep*, then the line search is terminated. A maximum allowable step length is also imposed to prevent excessively long steps which could occur in practice when $p_k = s_k^N = -H_k^{-1}\nabla f(x_k)$ and H_k is nearly singular. This maximum step length prevents taking steps that would result in the algorithms leaving the domain of interest.

4.6 The Model Trust Region Approach

In the model trust region approach an attempt is made to achieve global convergence without sacrificing the local convergence properties of the quasi-Newton method by dropping the assumption that shortened steps must be in the quasi-Newton direction. In the line search algorithms, while the step length is calculated based on function and gradient information, it has the disadvantage that it makes no further use of the n-dimensional quadratic model, including the model Hessian. In the model trust region approach, when a shorter step is needed, a shorter step length is chosen and then the full n-dimensional quadratic model is used to choose the step direction.

Suppose that we have x_c and some estimate δ_c of the maximum length of a successful step we are likely to be able to take from x_c. The quasi-Newton step s_c^N is reasonable because it is the step from x_c to the global minimizer of the local quadratic model m_c (if the model Hessian H_c is positive definite). If we add the idea of bounding the maximal step length by $\delta_c > 0$, then the best way to select a step of maximal length

δ_c from x_c is to try the step s_c that solves

$$\text{minimize } m_c(x_c + s) = f(x_c) + \nabla f(x_c)^T s + \frac{1}{2} s^T H_c s,$$

$$\text{subject to } \|s\|_2 \leq \delta_c \qquad\qquad (4.6.1)$$

Problem (4.6.1) is the basis of the "model - trust region" approach to minimization. The name comes from viewing δ_c as providing a region in which we can trust m_c to adequately model f. The solution is given by

$$s(\mu) = -(H_c + \mu I)^{-1} \nabla f(x_c) \qquad\qquad (4.6.2)$$

for the unique $\mu \geq 0$ such that $\|s(\mu)\|_2 = \delta_c$, unless $\|s(0)\| \leq \delta_c$, in which case $s(0) = s_c^N$ is the solution, For any $\mu \geq 0$, $s(\mu)$ defines a descent direction for f from x_c. A complete step of a trust-region algorithm will have the following form:

Algorithm (4.6.1): A Global step by the Model-Trust Region approach

Given f: $R^n \to R$, $\delta_c > 0$, $(x_c \in R^n)$, $H_c \in R^{n \times n}$ symmetric and positive definite:

   repeat

   (1) $s_c :=$ approximate solution to (4.6.1),

   $x_+ := x_c + s_c$,

   (2) decide whether x_+ is acceptable, and calculate a new value of δ_c

      until x_+ is an acceptable next point;

  $\delta_+ := \delta_c$

4.6.1 The Locally Constrained Optimal ("Hook") Step

Hook step method is an algorithm for finding an approximate solution μ_c to the scalar equation

$$\Phi(\mu) = \| s(\mu) \|_2 - \delta_c = 0. \tag{4.6.3}$$

To solve (4.6.3) we use a local model of the form

$$m_c(\mu_l) = \frac{\alpha_{ic}}{\beta_{ic} + \mu_l} - \delta_c$$

with two free parameters α and β. The subscript i is used for the current values of the quantities α_{ic}, β_{ic}, μ_{ic} that are changing in the inner iteration on μ, and the normal type without the subscript i is used for the current value of $\delta_c, x_c, \nabla f(x_c), H_c$ that come from the outer iteration (the main x iteration), and are unchanged during the solution of (4.6.3). Thus μ_c is the inner iteration's last approximation μ_{ic} to μ_*, the exact solution of (4.6.3). Therefore, we define $x_+ = x_c + s(\mu_c)$.

Choose α_{ic} and β_{ic} to satisfy the two conditions:

$$m_c(\mu_{ic}) = \frac{\alpha_{ic}}{\beta_{ic} + \mu_{ic}} - \delta_c = \Phi(\mu_{ic}) = \| s(\mu_{ic}) \|_2 - \delta_c$$

and

$$m'_c(\mu_{ic}) = \frac{-\alpha_{ic}}{(\beta_{ic} + \mu_{ic})^2} = \Phi'(\mu_{ic}) = \frac{s(\mu_{ic})^T (H_c + \mu_{ic} I)^{-1} s(\mu_{ic})}{\| s(\mu_{ic}) \|_2}$$

This gives

$$\alpha_{ic} = -\frac{(\Phi(\mu_{ic}) + \delta_c)^2}{\Phi'(\mu_{ic})}, \tag{4.6.4}$$

$$\beta_{ic} = -\frac{(\Phi(\mu_{ic}) + \delta_c)}{\Phi'(\mu_{ic})} - \mu_{ic}. \tag{4.6.5}$$

Choose μ_{l+} such that $m_c(\mu_{l+}) = 0$ ----- that is,

$$\mu_{l+} = \frac{\alpha_{lc}}{\delta_c} - \beta_{lc}. \qquad (4.6.6)$$

Substituting (4.6.3), (4.6.4) and (4.6.5) into (4.6.6), gives

$$\mu_{l+} = \mu_{lc} - \left[\frac{\Phi(\mu_{lc}) + \delta_c}{\delta_c} \right] \left[\frac{\Phi(\mu_{lc})}{\Phi'(\mu_{lc})} \right] = \mu_{lc} - \frac{\| s(\mu_{lc}) \|}{\delta_c} \left[\frac{\Phi(\mu_{lc})}{\Phi'(\mu_{lc})} \right] \qquad (4.6.7)$$

4.6.2 Details of More's Implementation

We have used the implementation by More in MINPACK (More 1980). The following algorithm is based on More (1977) and Hebden (1973). μ_0, the starting value for μ is set to zero in solving (4.6.3) by Reinsch (1971). But, each iteration of (4.6.7) involves solving a linear system. Hence, to get a closer start More's implementation uses the approximate solution μ_- to the last instance of (4.6.3) to generate an initial guess for the current instance. If the current step bound δ_c is p times the last value of the step bound δ_-, then $\mu_0 = \frac{\mu_-}{p}$ is used to start (4.6.7).

4.6.3 Generation and Update of Lower and Upper Bounds

μ_{l+} is restricted to be in $[l_{l+}, u_{l+}]$, the lower and upper bounds on μ_{l+}. Take $l_{l0} = -\frac{\Phi(0)}{\Phi'(0)}$. Then calculate $\mu_{l+}^N = \mu_{lc} - \frac{\Phi(\mu_{lc})}{\Phi'(\mu_{lc})}$ along with each calculation of (4.6.7), and update the lower bound to $l_{l+} = \max \{l_{lc}, \mu_{l+}^N\}$, where l_{lc} is the current lower bound. Take $u_{l0} = \frac{\| \nabla f(x_c) \|_2}{\delta_c}$ and at each iteration, if $\Phi(\mu_{lc}) < 0$ update the upper bound to $u_{l+} = \min \{u_{lc}, \mu_{lc}\}$, where u_{lc} is the current upper bound.

If, at any iteration, μ_{l+} is not in $[l_{l+}, u_{l+}]$, More chooses μ_{l+} by $\mu_{l+} = \max \{ (l_{l+} \cdot u_{l+})^{1/2}, 10^{-3} u_{l+}\}$, the second term being a safeguard against near-zero values of l_{l+}.

More's implementation does not solve (4.6.3) to any great accuracy, settling

instead for $\|s(\mu)\|_2 \in [0.9\delta_c, 1.1\delta_c]$. Finally, More uses the scale trust region approach so that the steepest-descent step becomes

$$x_+ = x_c - \lambda D_x^{-2} \nabla f(x_c),$$

and the hook step becomes

$$s(\mu) = -(H_c + \mu D_x^2)^{-1} \nabla f(x_c).$$

where D_x is the positive diagonal scaling matrix.

4.6.4 The Double Dogleg Step

The Double Dogleg step is a modification of the trust region algorithm introduced by Powell (1970). It also finds an approximate solution to problem (4.6.1) by a piecewise linear function connecting the "Cauchy point", the minimizer of the quadratic model m_c in the steepest - descent direction, to the Newton direction for m_c. Then it chooses x_+ to be the point on this polygonal arc such that $\|x_+ - x_c\|_2 = \delta_c$, unless $\|H_c^{-1} \nabla f(x_c)\|_2 < \delta_c$, in which case x_+ is the Newton point. This strategy is a simple strategy for looking in the steepest - descent direction when δ_c is small and more toward the quasi - Newton direction as δ_c increases.

4.6.5 Updating the Trust Region

To complete the global step given in Algorithm (4.6.1), one needs to decide whether the point x_+ found by using the techniques of optimal "Hook" step or the double dogleg step, is a satisfactory next iterate. If x_+ is unacceptable, one reduces the size of the trust region and minimizes the same quadratic model on the smaller trust region. If x_+ is satisfactory, one must decide whether the trust region should be increased, decreased, or kept the same for the next step.

The condition for accepting x_+ is

$$f(x_+) \leq f(x_c) + \alpha g_c^T(x_+ - x_c),$$
(4.6.8)

where $g_c = \nabla f(x_c)$ or an approximation to it, and α is a constant set to 10^{-4}. If x_+ does not satisfy the above condition, reduce the trust region by a factor between $1/10$ and $1/2$ and return to the approximate solution of the locally constrained minimization problem by the locally constrained optimal step or double dogleg method. The reduction factor is determined by the same quadratic backtrack strategy used to decrease the line-search parameter in Algorithm (4.5). Model $f(x_c + \lambda(x_+ - x_c))$ by the quadratic model $m_q(\lambda)$ that fits $f(x_c)$, $f(x_+)$, and the directional derivative $g_c^T(x_+ - x_c)$ of f at x_c in the direction $x_+ - x_c$. The new trust radius is extended to the minimizer of this model, which occurs at

$$\lambda_* = \frac{-g_c^T(x_+ - x_c)}{2\,[f(x_+) - f(x_c) - g_c^T(x_+ - x_c)]}$$

Thus, $\delta_+ = \lambda_* \| x_+ - x_c \|_2$. If $\lambda_* \| x_+ - x_c \|_2 \notin \left[\frac{1}{10}\delta_c, \frac{1}{2}\delta_c\right]$, then δ_+ is set to the closer end-point of this interval. Suppose x_+ satisfies condition (4.6.8). If x_+ is a full Newton step from x_c, then make the step, update δ, form the new model, and go to the next iteration. However, if $(x_+ - x_c)$ isn't the Newton step then consider whether to try a larger step using the current model.

To decide whether to attempt a larger step from x_c, compare the actual reduction $\Delta f = f(x_+) - f(x_c)$ to the predicted reduction $\Delta f_{pred} = m_c(x_+) - f(x_c)$. And x_+ is the next iterate unless:

(1) $|\Delta f_{pred} - \Delta f| \leq 0.1|\Delta f|$ which implies that the agreement is so good that

δ_c is an underestimate of the radius in which m_c adequately represents f

(2) $f(x_+) \leq f(x_c) + \nabla f(x_c)^T(x_+ - x_c)$ which implies the presence of negative

curvature and thus a continuing rapid decrease in f because of the large actual

reduction in f

In both the above cases save (x_+) and $f(x_+)$, double δ_c and compute a new x_+ using the current model. If the condition (4.6.8) isn't satisfied for the new x_+, drop back to the last good step. But, if it is satisfied, consider doubling again.

If x_+ is acceptable as the next iterate, then update δ_c to δ_+ according to the following rules. If the current quadratic model predicts the function well, increase the trust region, but if it predicts poorly, decrease the trust region.

(1)   If the quadratic model has predicted the actual function reduction sufficiently well, $\Delta f \leq 0.75 \Delta f_{pred}$ then $\delta_+ = 2\,\delta_c$.

(2)   If the model has greatly overestimated the decrease in $f(x)$, $\Delta f > 0.1\Delta f_{pred}$ then

$$\delta_+ = \delta_c/2.$$

(3)   otherwise $\delta_+ = \delta_c$.

4.7 The Homogeneous Transformation

Our objective is to determine the position and orientation of the sphere with respect to a base co-ordinate frame defined conveniently by the user. Hence, we attach a co-ordinate frame at the centre of the sphere. We also attach co-ordinate frames to each transmitting and receiving antennae with the z - axis of each co-ordinate system parallel to the axis of the loop antennae.

We know the position and orientation of each transmitter with respect to each other and with respect to the centre of the sphere. We also know the position and orientation of each receiver with respect to each other and with respect to the base co-ordinate frame. This information is represented in the form of Homogeneous transformation matrices. Therefore, we know ${}^{c}T_{t_i}$ i=1,...,p and ${}^{base}T_{r_j}$ j=1,...,q, where ${}^{c}T_{t_i}$ is the transformation from transmitter i to the centre of the sphere and ${}^{base}T_{r_j}$ is the transformation from the receiver j to the base co-ordinate frame.

The transfer function between the transmitter and the receiver includes parameters

for the relative position and orientation of the transmitter with respect to the receiver. Hence, to calculate the voltages induced in each receiving antenna, we need to compute the relative position and orientation of each transmitter with respect to the corresponding receiving antenna. This is done by computing the transformation between each transmitter - receiver pair:

$$^{r_j}T_{t_i} = {}^{r_j}T_{base} \, {}^{base}T_c \, {}^cT_{t_i}$$

where $^{r_j}T_{base} = [{}^{base}T_{r_j}]^{-1}$, the transformation from the base co-ordinate frame to the receiver j, and $^{base}T_c$ is the transformation from the centre of the sphere to the base co-ordinate frame. From $^{r_j}T_{t_i}$, we compute the x necessary to compute the voltage V_{ij} in equation (3.4.1).

CHAPTER 5

RESULTS AND CONCLUSIONS

5.1 Introduction

In this Chapter we present the results obtained by using the numerical methods discussed in Chapter 4. We show from the results that the numerical method of solution to the problem at hand is a viable and, perhaps the only, efficient method. We show that a particle's trajectory can be tracked accurately within limits of experimental error thereby establishing the feasibility of our technique to track a particle in a mass flow of granular material non-intrusively. Results from simulated and actual experimental data are provided. The effect of and remedies to signal noise are discussed.

5.2 The Forward Solution

The forward solution involves computing the voltages induced in the different receiving antennae for a given location of the sphere using the models discussed in Chapter 3. The forward solution is essential to compute the residuals needed in the inverse solution. Moreover, the forward solution helps in validating and "fine tuning" the models by comparing them with the actual experimental radiation pattern of the transmitter. It also helps in generating the voltage values for the simulated experiments in the absence of the experimental set-up to conduct the experiments. Thus a concurrent approach to various phases of the diagnostic technique can be adopted with the method of solution for the inverse problem decided even before the actual experimental voltage values are available. The forward solution was obtained for all the three models discussed previously.

To verify the cosine behavior of the model #3, the transmitter was oriented at various angles with the X and Y axes. Measurements of the induced voltage was made by

(1)   Fixing angle α = 90° and varying γ between 0° and 90°, and

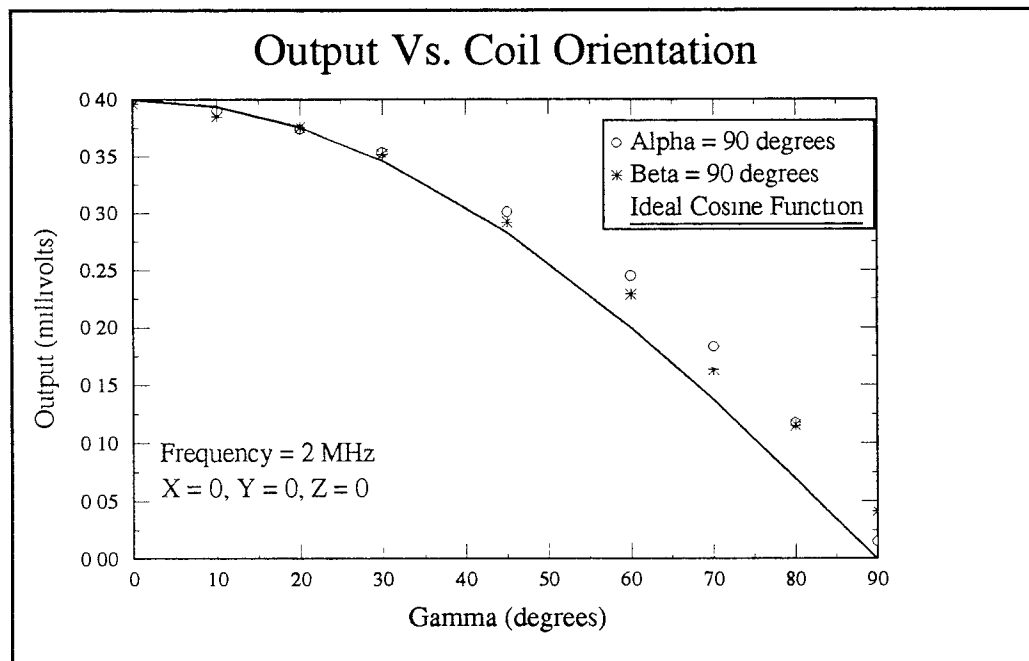(2)   Fixing angle β = 90° and varying γ between 0° and 90°.

Figure 5.1: Output versus Coil Orientation.

Figure (5.1) shows the predicted output voltages and the actual cosine function as the coil orientation is varied. The trend shown in Figure (5.1) agrees well with the cosine function.
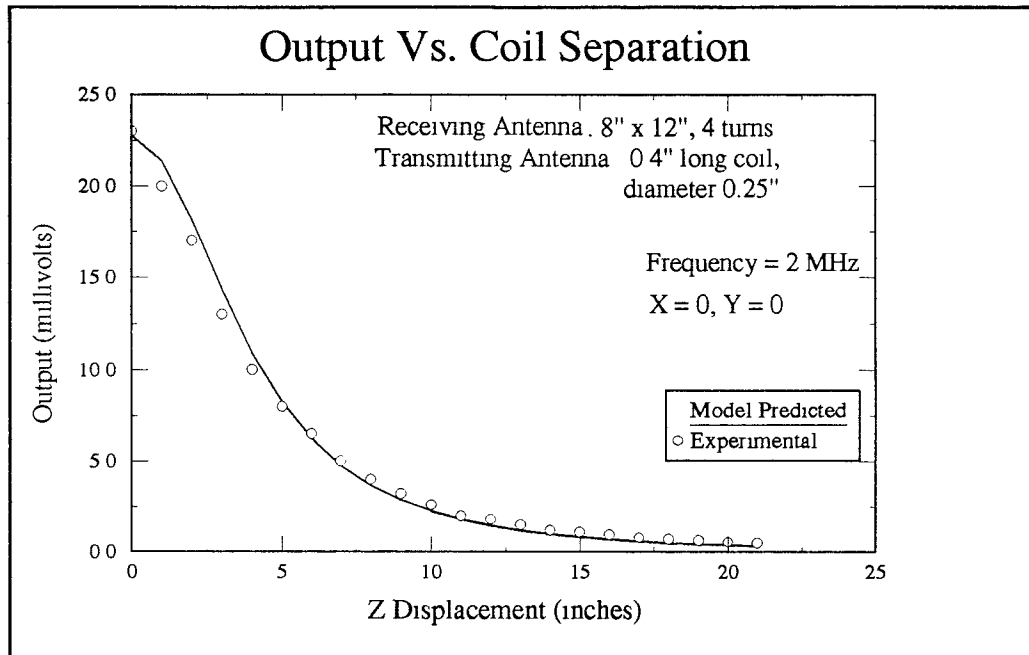
Figure 5.2: Output versus Coil Separation.

Figure (5.2) shows the plots of the predicted and the actual experimental measured voltages as the transmitter is moved along the Z axis. There is a strong correlation between the measured and the predicted voltages for displacement of the transmitter along the Z axis. It should be noted that a scale factor related to attenuation was applied to the theoretical values.

Figure (5.3) shows the plots of the predicted and the actual experimental measured voltages as the transmitter is moved along the X axis. There is a strong correlation between the measured and the predicted values of the induced voltage for variation of the transmitter along the X axis. The experimental as well as the predicted voltage curves display a minimum at X = 0" and a maxima at X = -4.7" and 4.7".
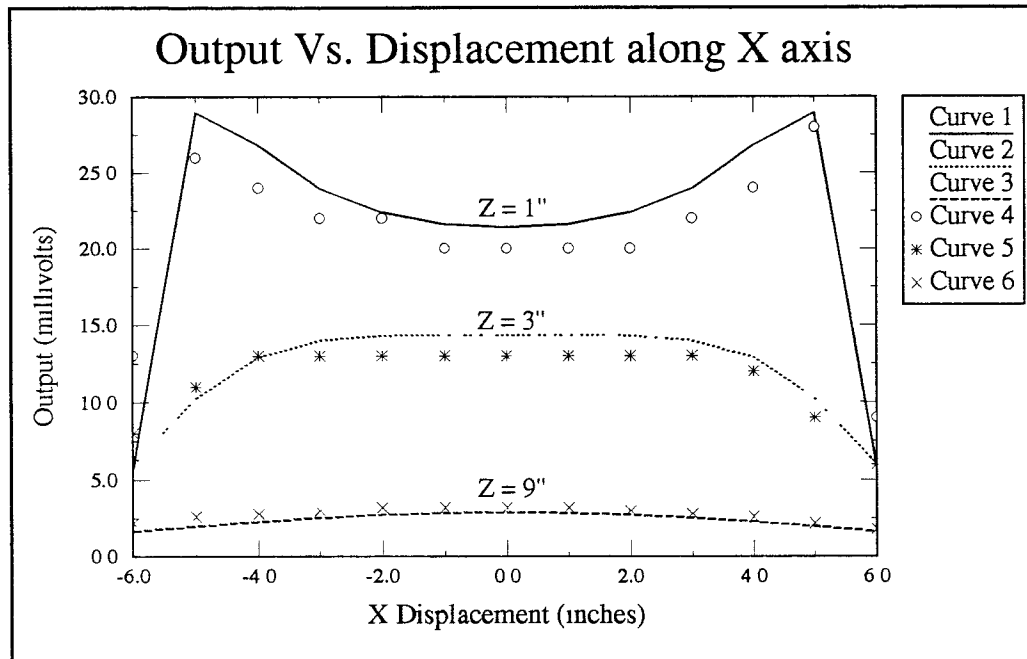
Output Vs. Displacement along X axis

	Curve 1
	Curve 2
	Curve 3
o	Curve 4
*	Curve 5
x	Curve 6

Z = 1"

Z = 3"

Z = 9"

Output (millivolts)

X Displacement (inches)

Figure 5.3: Output versus X Axis Displacement.

5.3 The Inverse Solution

The form of the equations (3.4.1) is dependent on the assumed geometric shapes of the transmitters and the receivers. As discussed in Chapter 2, the transmitters are assumed to be small circular loops while the receivers are modeled in three different geometric shapes. All the three models were considered for inverse solution using the numerical techniques described in Chapter 4. For the first and second models there are three orthogonal transmitters inside the tracking sphere with three orthogonal receivers around the experimental space which is shown in Figure (5.4). The figure is not shown to scale in order to clearly show the arrangement of the transmitters within the sphere. Thus there are nine equations for the six location variables. For the third model there is a single transmitter with six receivers. Hence, there are six equations for the five independent variables in model #3. Thus there is an overdetermined system in each case and

the method of solution remains the same. The results were excellent in each case.

Figure 5.4: Arrangement of Three Orthogonal Transmitters and Receivers.
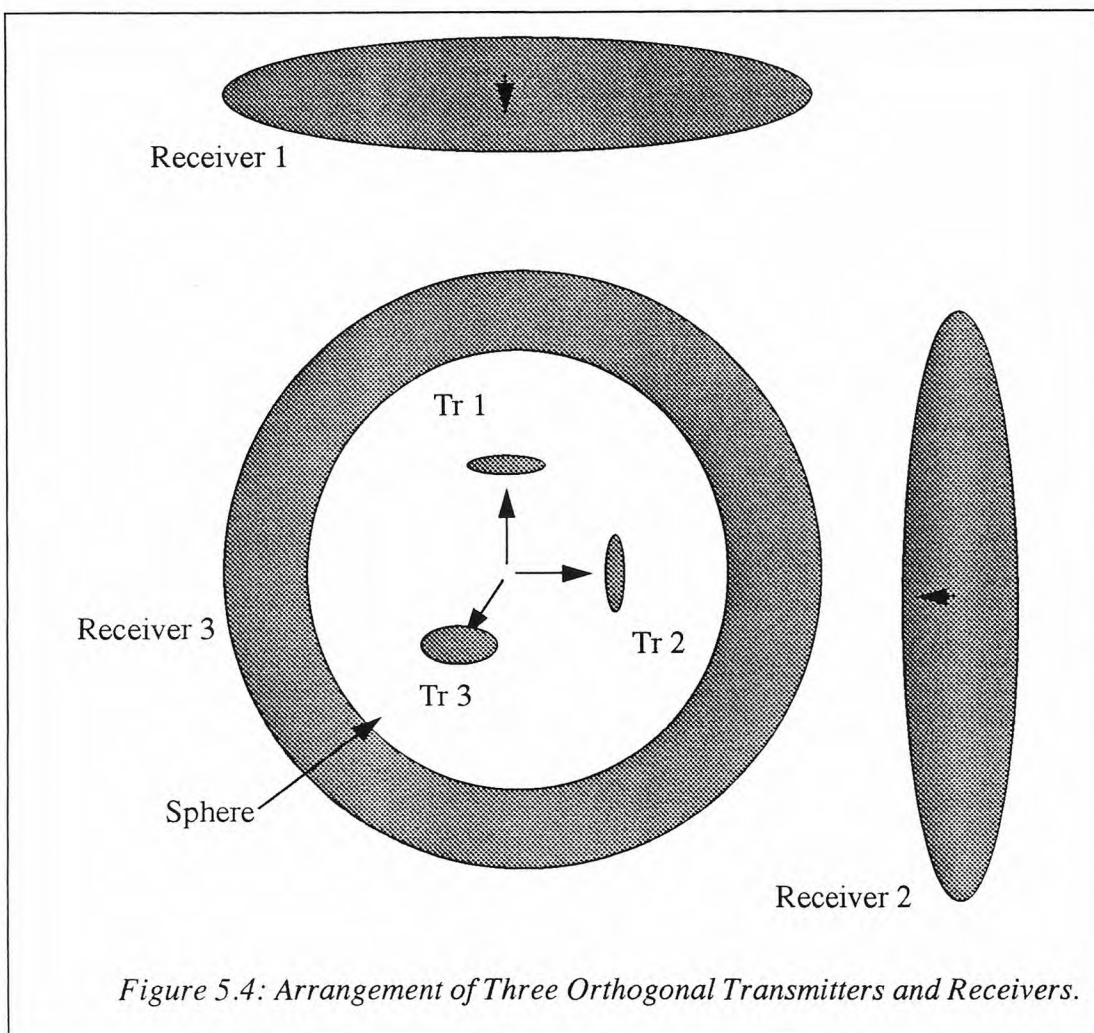
The reason for using only one or two transmitters is due to the size limitation of the tracking sphere. More transmitters take up more space and forces a higher drain on the power supply. Model #3 is the most realistic as the configuration closely corresponds to the experimental set-up. It should be noted that only the direction cosine angles are present in model #3 which does not uniquely determine the orientation. Moreover, the voltage induced in the receiving antenna is independent of the rotation of the transmitter about its own axis because of the symmetric geometry of the transmitter. Thus at least

two transmitters are needed to uniquely determine the orientation. The model in this case shows that the induced voltages are functions of five location variables, and not six. There are multiple solutions in the case of orientation for model #3. But the 3-D position is solved unambiguously. Model #3 is used because experimental results are available only for this case.

Initially the simple Gauss-Newton method was used for the inverse solution. Even though the algorithm was able to converge to the correct solution in individual cases, it suffered from its inability to achieve global convergence. The algorithm would converge only to the local minima. Hence the initial guess, as required by all these methods, had to be close to the actual solution. It had problems especially in the regions of high non-linearity and would get entrapped in "troughs" associated with local minima.

Hence, the damped Gauss-Newton method incorporating the line search to achieve global convergence was tried next. It was successful in achieving global convergence over the domain of the problem at hand. But the rate of convergence was unsatisfactory. Figure (5.5) shows the slow convergence rate of the damped Gauss-Newton method. The residual voltages tend to zero as convergence is achieved with the damped Gauss-Newton algorithm.

The Levenberg-Marquardt algorithm was tried to achieve faster convergence rate. We used the implementation by More (1980) in MINPACK. A modified Levenberg-Marquardt algorithm is used in this implementation. More uses the scaled trust region approach using Hook step method as discussed in Chapter 4. The ability to converge to a solution was good throughout the domain of the experiment. The rate of convergence proved to be satisfactory. Figure (5.6) shows the relatively quick convergence rate of the Levenberg-Marquardt method. The residual voltages tend to zero as convergence is achieved with the Levenberg-Marquardt algorithm.
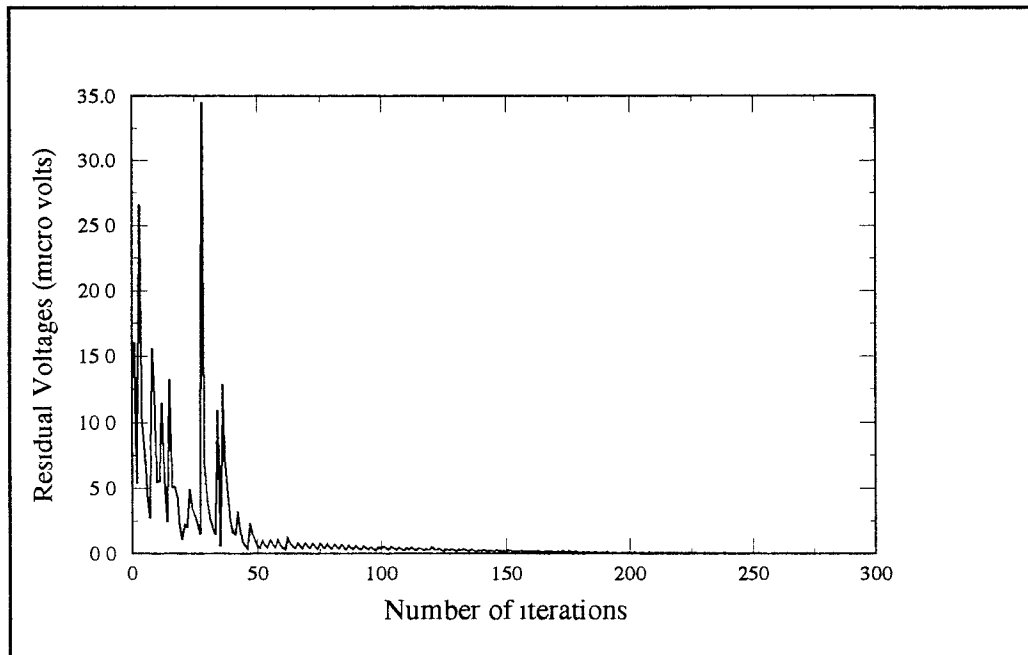

Figure 5.5: Slow Convergence Rate of Damped Gauss-Newton Algorithm.
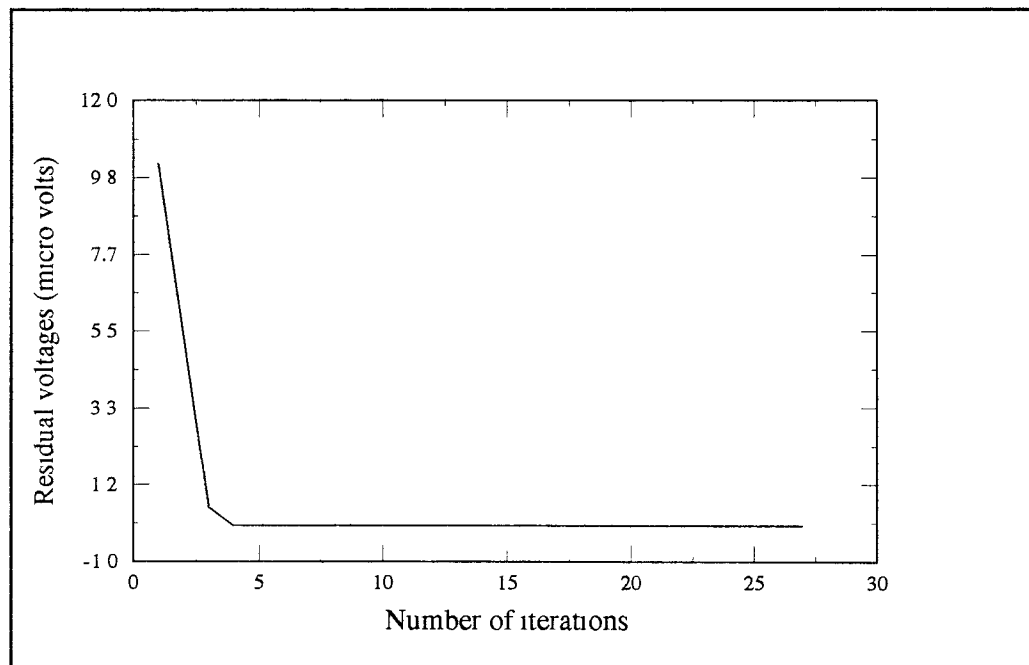

Figure 5.6: Faster Convergence Rate of Levenberg-Marquardt Algorithm.

Thus the Levenberg-Marquardt algorithm was successfully used for the inverse solution to predict the trajectory of the tracking sphere. Simulated experiments are considered first. Assuming that the induced voltage values are available, we were able to track the trajectory of the sphere. The voltage values are generated at suitable intervals of the trajectory using the forward solution. This is then used in the inverse solution in place of the actual experimental voltage values. Figure (5.7) shows the arrangement of the six receiving antennae around the experimental space.

Figure 5.7: Configuration of Receiving Antennae around the Experimental Space.
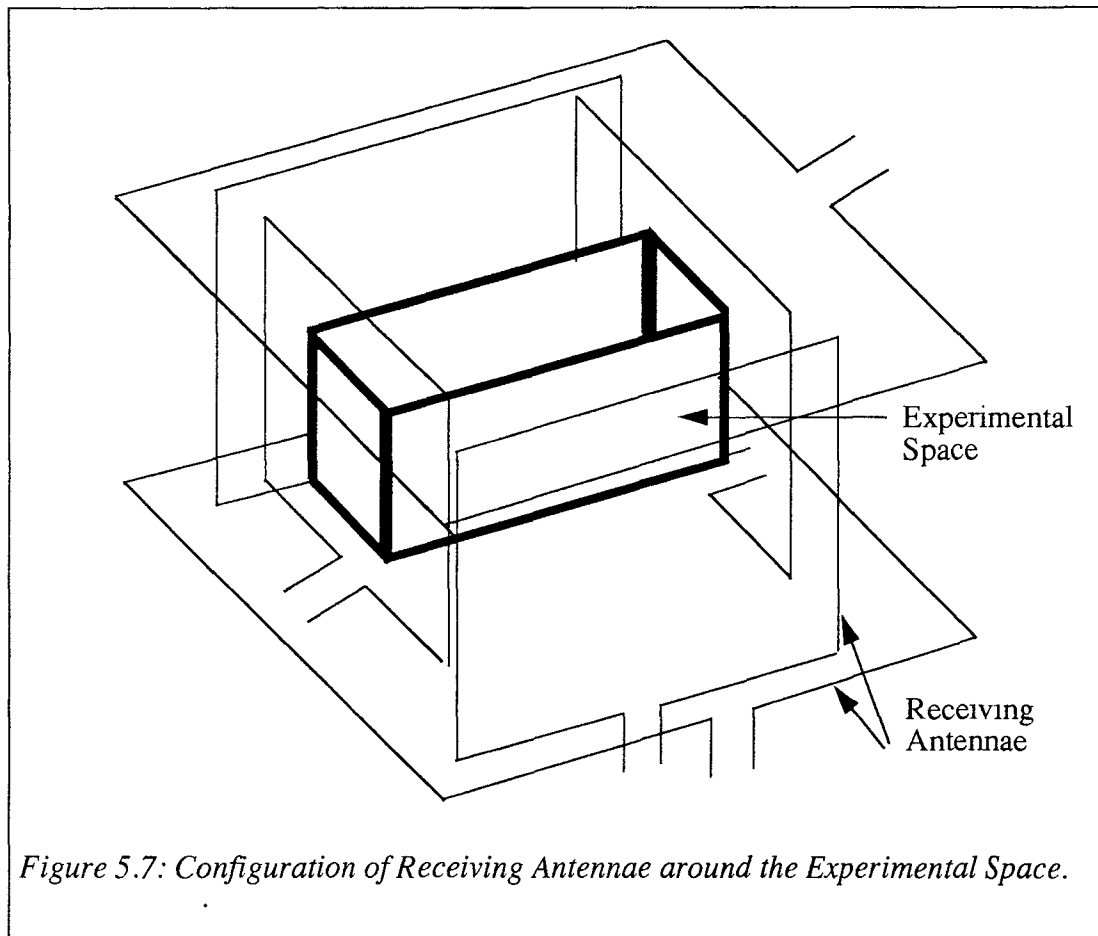
Two hypothetical trajectories are considered, a straight line trajectory and a sinusoidal trajectory. These trajectories have no special significance in that they were used for the sole reason that the voltages for these trajectories are easy to generate using the forward solution. Figure (5.8) shows the results of the simulated experiments without noise. The solid curves are actual trajectories while the markers are the predicted trajectories. The results are excellent as is to be expected because these are the ideal situations with no noise or experimental errors. Thus, in actuality it is a zero-residual problem.
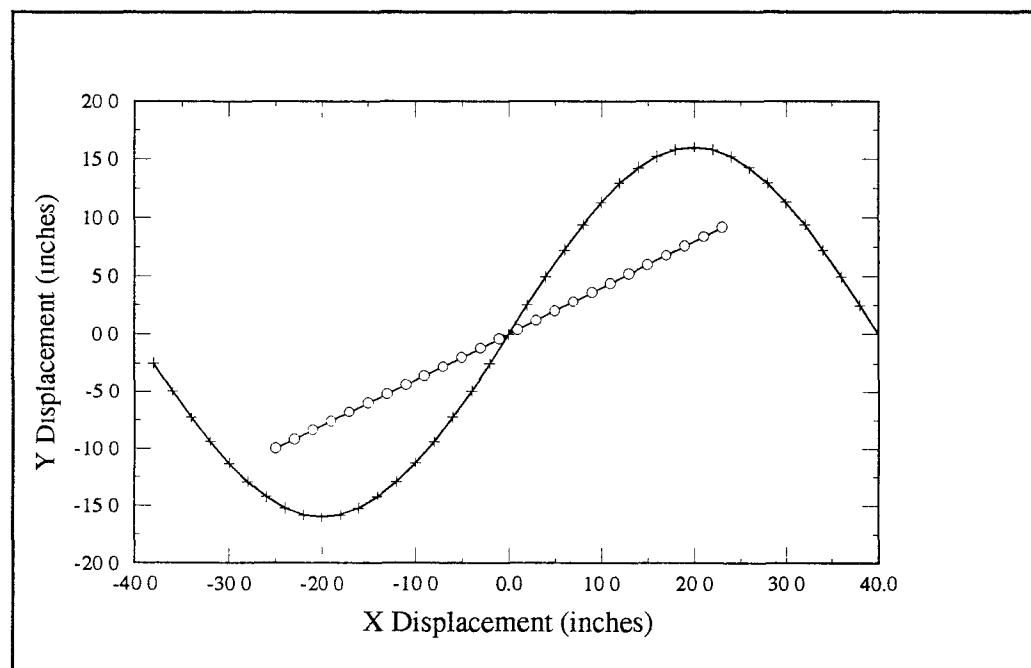

Figure 5.8a: Projection of the Trajectories onto the XY Plane.

Figure (5.8a) and (5.8b) shows the projection of the trajectories onto the XY plane and the XZ plane respectively.
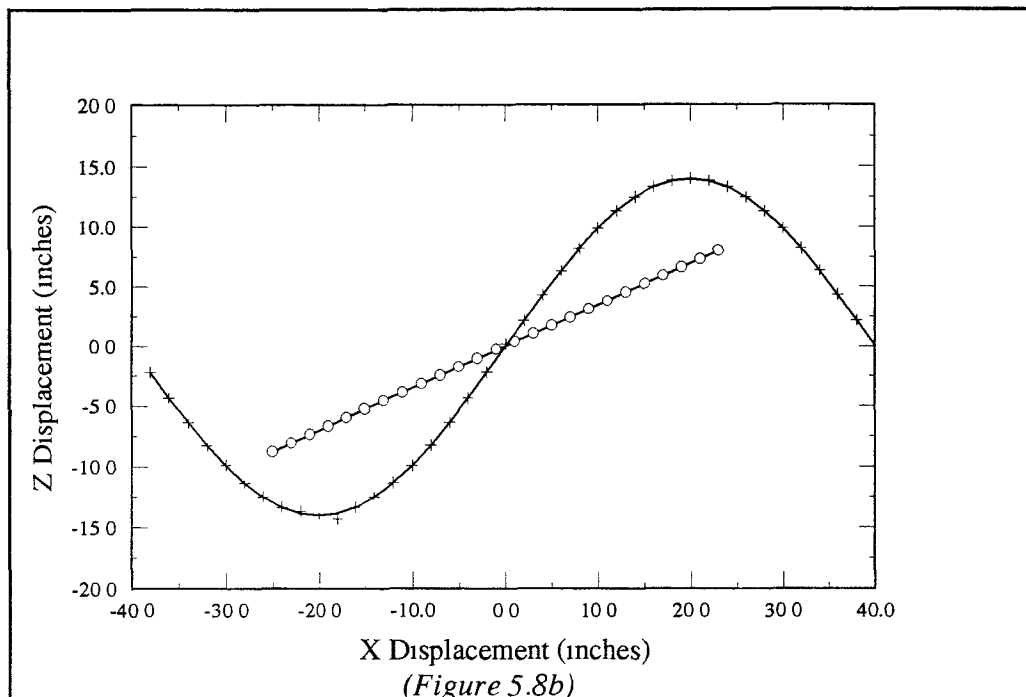
Figure 5.8b: Projection of the Trajectories onto the XZ Plane.

Figure 5.8: Simulated Experiment Without Noise.

To study the prediction of the orientation angles a simulated experiment was carried out with the angles varying linearly. Again, the linear variation of the angles was selected only because it is easy to generate the voltage values for the inverse solution using the forward solution. The angles may vary in any manner and still the inverse solution would predict the orientation correctly. Figure (5.9) shows the results obtained for the simulated experiment for calculating the sphere orientation using the inverse solution. As can be seen, the predicted orientation is in close agreement with the actual angles. The solid curves represent the actual orientation angles while the markers denote the predicted values.
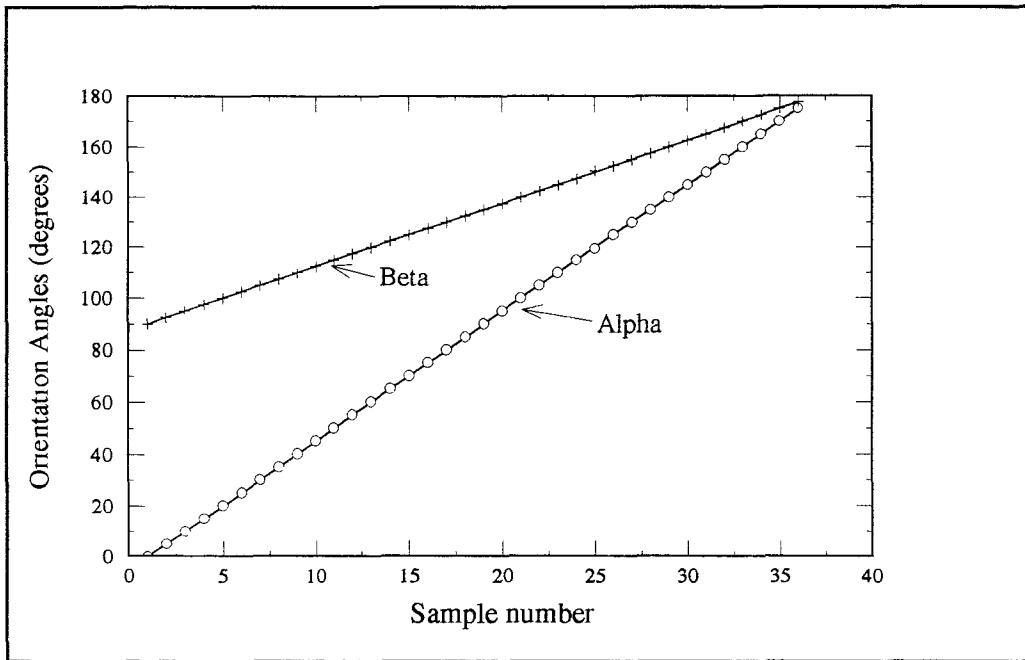
Figure 5.9: Simulated Experiment to Predict Sphere Orientation.

At the time of writing this thesis, the construction of the experimental set-up is not complete. Hence, tracking the sphere in one dimension is considered. Using two receivers and one transmitter, one can track the trajectory in one dimension perpendicular to the plane of the receivers. An experiment is conducted where the tracking sphere is moved along the lines AB and CD shown in Figure (5.10) with the other positions and orientations fixed. Figure (5.10) also shows the experimental set-up with two parallel receiving antennae.

Figure 5.10: Experimental Set-Up with Two Receiving Antennae.
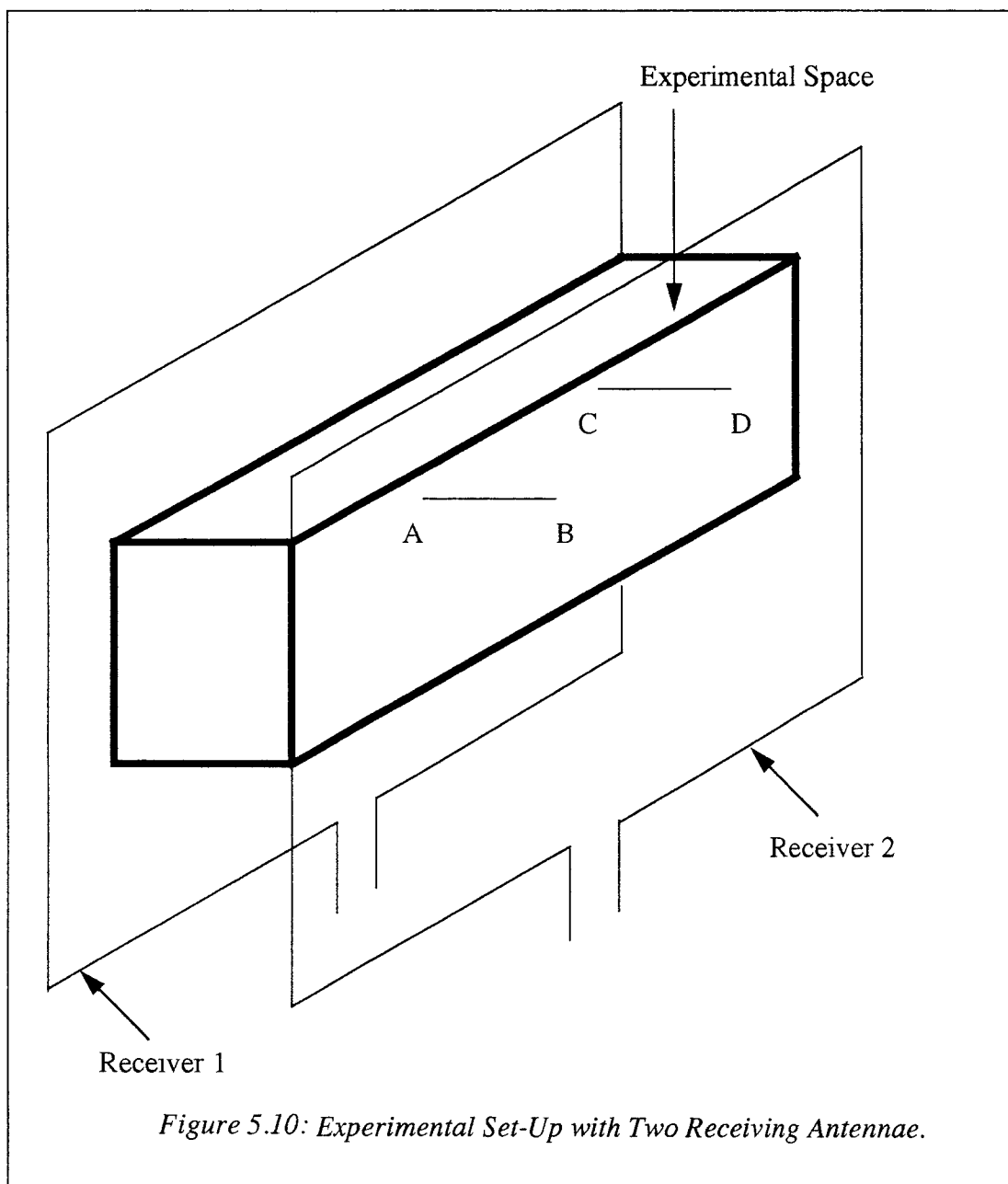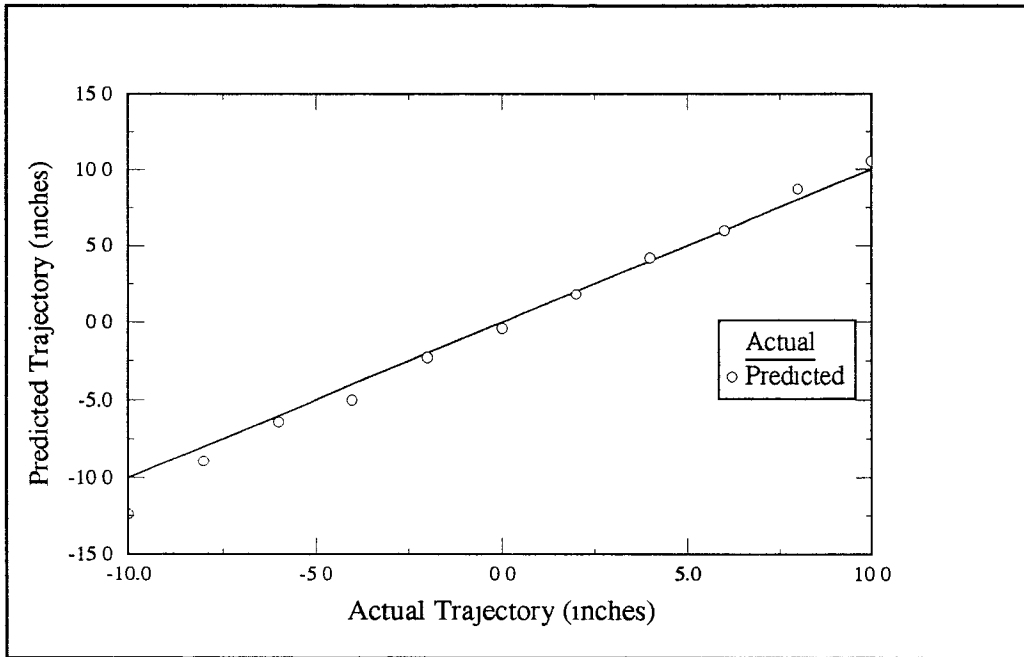
Figure 5.11 a: Predicted Trajectory versus Actual Trajectory for Sphere Movement: Along Line AB in Actual Experiment.
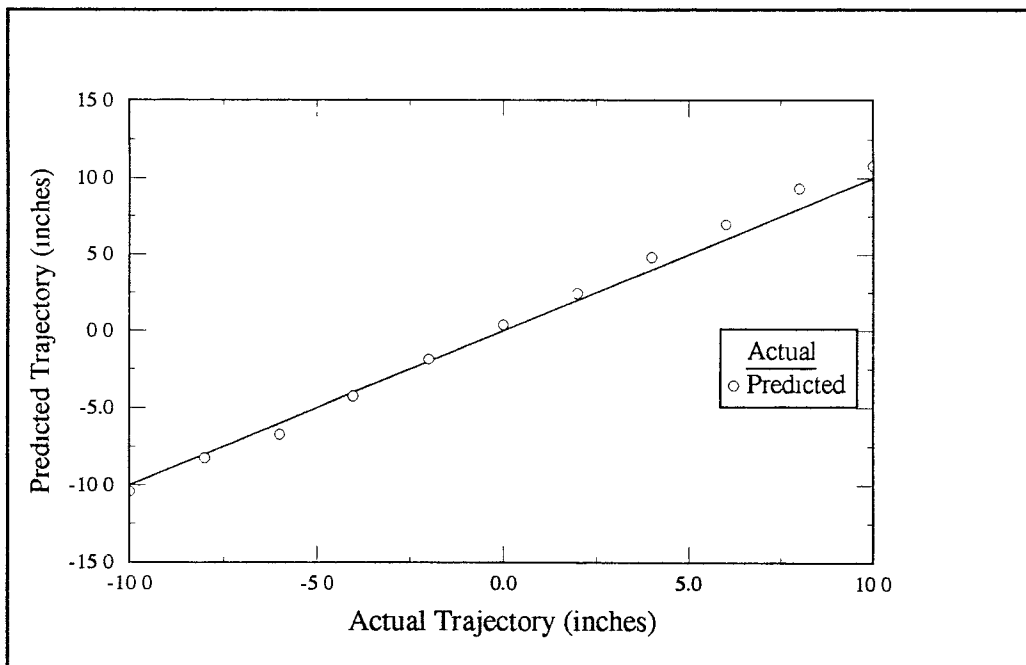

Figure 5.11 b: Predicted Trajectory versus Actual Trajectory for Sphere . Movement: Along Line CD in Actual Experiment.

The voltages induced in the two receivers are used to predict the trajectory of the sphere. Figure (5.11) shows the predicted versus the actual trajectory for sphere movements along lines AB and CD.

The theoretical forward model has not been fine tuned using the empirical data. The deviations are due to the combined effects of the discrepancy between the theoretical model and the actual radiation pattern and the noise in the signals and measuring instruments. Thus the problem is no longer a zero residual problem. Considering that, the results are excellent and are highly promising.

5.4 Effect of Noise in Signal Voltages

In reality there is always noise along with the signals. The noise may be due to the instrument errors or because of some spurious signals. To study the noise pattern an experiment was conducted by placing the transmitter at a stationary position. The induced voltage was measured for a second at a rate of 10000 samples per second. This high rate of measurement ensures that the signal level does not change appreciably between individual readings. In the ideal case all the 10000 readings should give the same signal level because the transmitter is fixed at a location. Due to the errors associated with the measuring instrument, it is reasonable to expect a Gaussian distribution around the actual signal level. But, because of other noise sources we do not get a true Gaussian distribution. Figure (5.12) shows the plot of the signal level versus the frequency of its occurrence. As can be seen there is a second smaller "peak" associated with noise of a lower frequency.
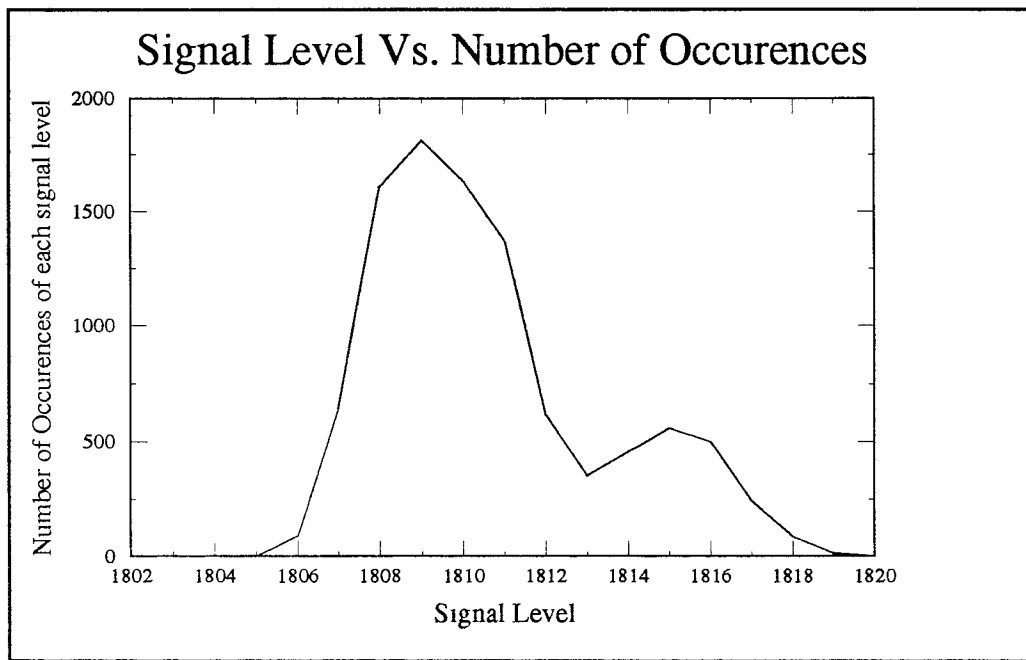
Figure 5.12: Signal Level versus its Frequency of Occurrence.

To get rid of this low frequency noise, a fast fourier transform (*fft*) was performed on the signal. It was found that the noise was at the low frequency of 120 Hz., 240 Hz. and its higher harmonics and beat frequencies. This can be removed by setting the *fft* values of the noisy signal corresponding to 120 Hz., 240 Hz., etc. to zero and then performing an inverse *fft* to get back to the signals. But the low frequency noise was removed in the hardware itself by incorporating certain electronics (Troiano, A. 1992). Figure (5.13) shows the plot of the power spectral density versus the frequency before the noise removal and after the noise removal. The Power spectral density values are high before noise removal and comparatively very low after noise removal.
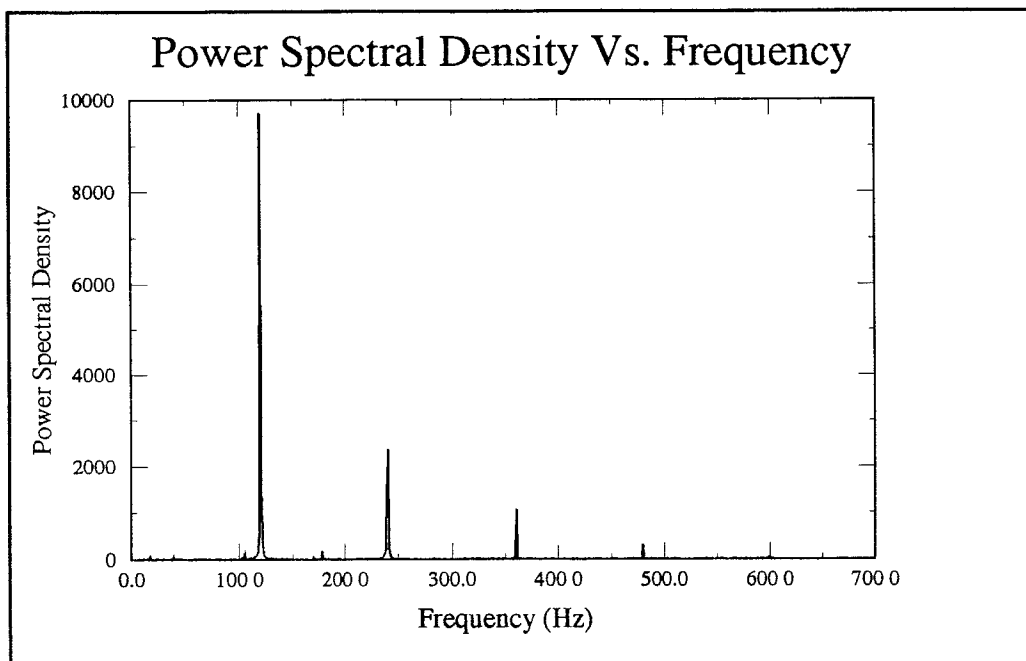
Power Spectral Density Vs. Frequency

Figure 5.13a: Power Spectral Density versus Frequency: Before Noise Removal.
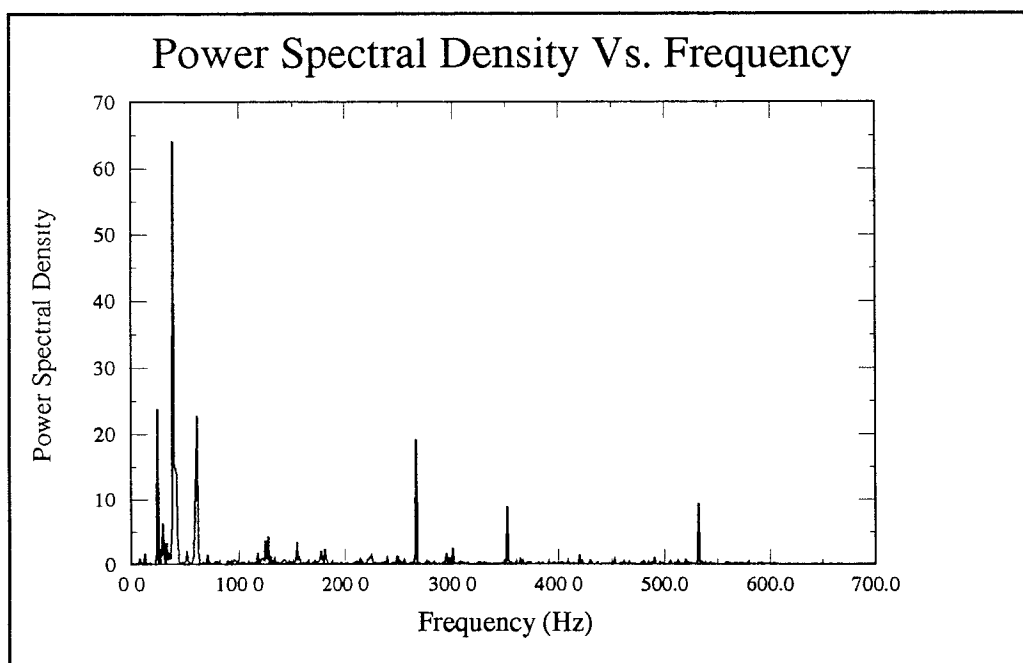
Power Spectral Density Vs. Frequency

Figure 5.13b: Power Spectral Density versus Frequency: After Noise Removal.
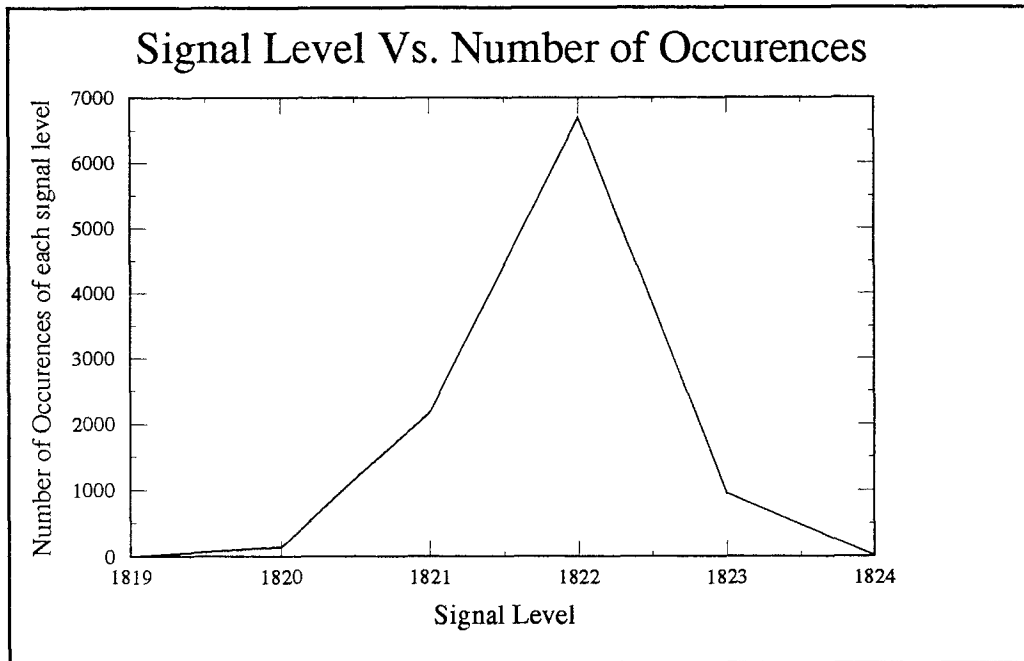
Figure 5.14: Gaussian Instrument Error Function Obtained after Removal of Spurious Noise.

After the noise removal in the hardware we find that we have Gaussian instrument error function (Figure 5.14). This can be further removed by using any data smoothing routine (Hayden 1987). To study the effect of noise on the predicted trajectory, a uniformly distributed noise with an amplitude of 5% is added to the set of voltage values given by the forward model. The noise added voltage values are then used to obtain the sphere trajectories. Figure (5.15) shows trajectories obtained from noise added voltages. Again, the solid curves are the actual trajectories while the markers are the predicted trajectories. The error in prediction of the trajectory is because of the introduction of noise which makes the problem no longer one of zero-residual. But the algorithm is robust and does not completely fail and is still able to converge to a solution.
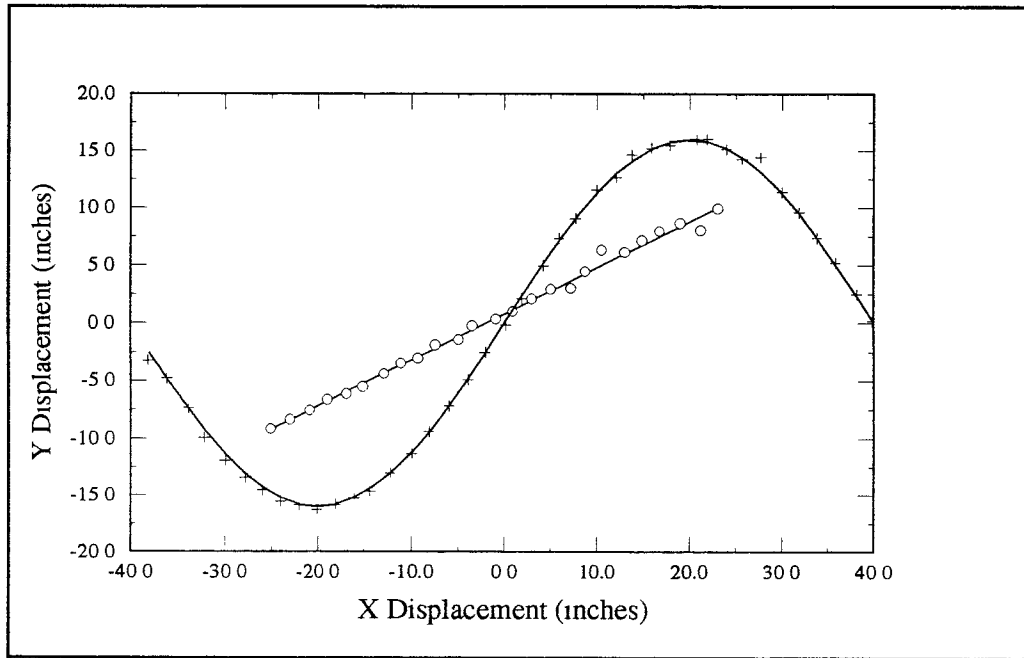
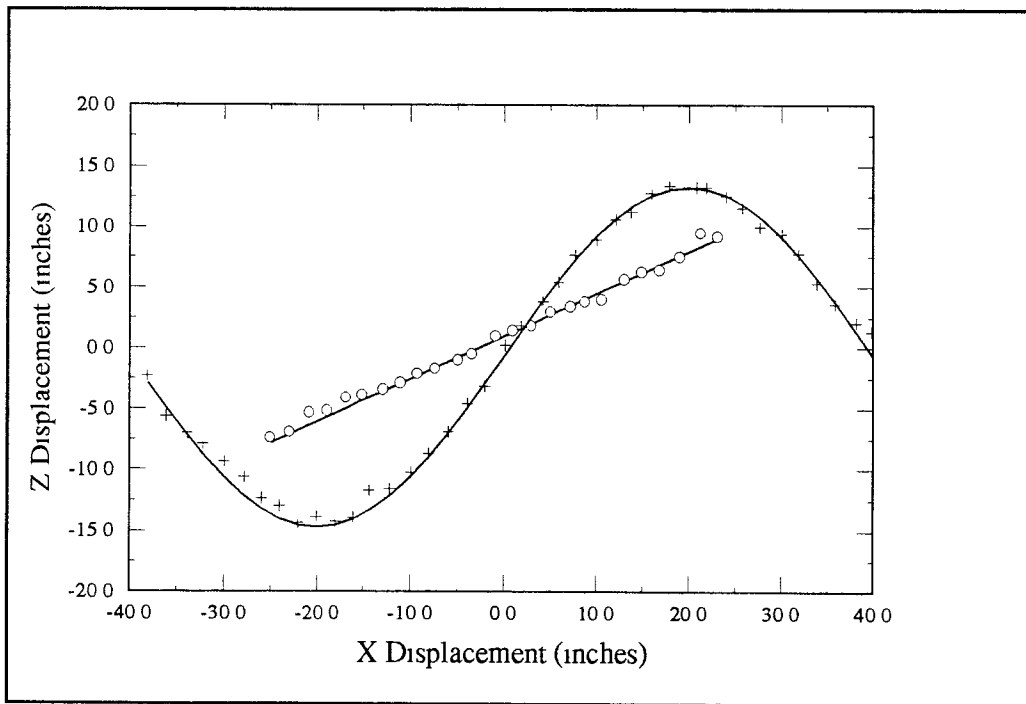*Figure 5.15a: Projection of the Trajectories onto the XY Plane.*

Figure 5.15b: Projection of the Trajectories onto the XZ Plane.
Figure 5.15: Simulated Experiment with Added Noise.

After processing the signals to remove the noise the voltage values are closer to those predicted by the model. So, it gives better results. Figure (5.16) shows the plot of error in the predicted positions before and after noise smoothing.
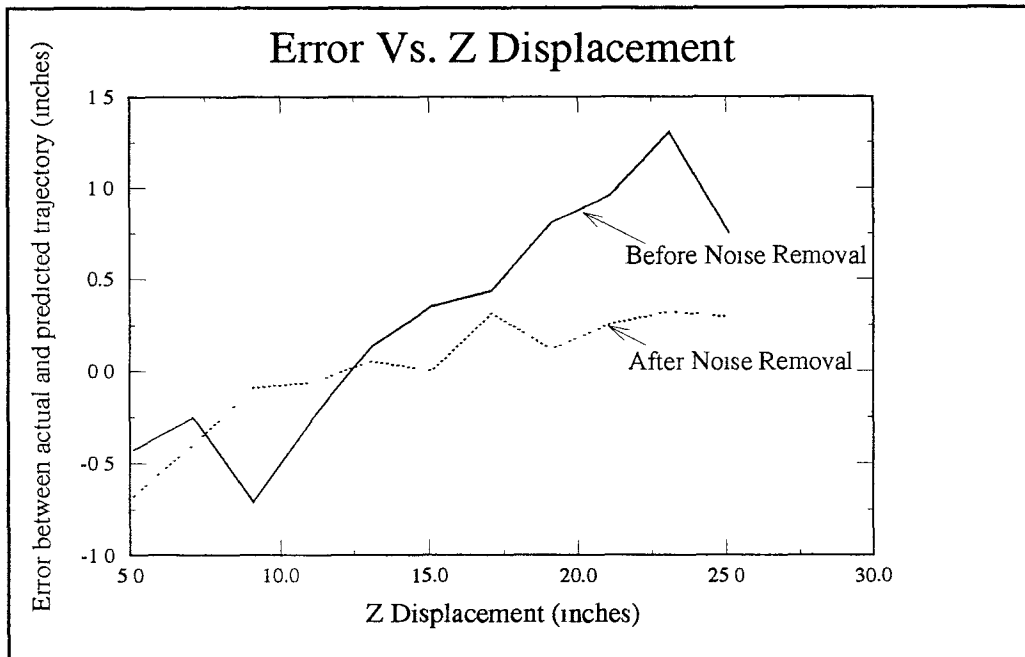
Figure 5.16: Reduced Error in Prediction of Trajectories after Noise Removal as Compared to Before Noise Removal.

5.5 Conclusions

A new system based on the principle of magnetic induction is presented. The inverse solution techniques to predict the location of the particle is presented. We have shown that the Levenberg-Marquardt algorithm is an appropriate choice for this problem. Simulated experiments for different transmitter-receiver geometry and dimensions were successfully completed. The solution technique is the same and is proved to be adequate for all the different cases. Convergence is achieved even if the initial guesses are far away. This is significant as the predicted location becomes the initial guess for the next

trajectory point. Hence, convergence from a relatively far off initial guess ensures that there is no accumulation or propagation of errors from a single bad point. The bad point could always be eliminated by a simple heuristic check based on the principle of continuity of motion.

The dynamic range of the predicted voltage values is still not exactly the same as that of the experimental values. Hence the model predicts voltage values which are different from the actual values particularly near the plane of the receiving antenna. Hence the error in the prediction of the trajectory is the greatest near to the antenna. A simple remedy to this problem could be to use sufficiently large dimensions of the receiving antenna so that the experimental space is in the relatively linear region of the voltage space. But increasing the distance of the transmitter from the receiver results in lower signal levels and hence inferior signal to noise ratios. A more desirable solution to this problem would be to fine tune the model using the empirical map.

Actual experimental results support our premises based on simulated experiments. Hence, the feasibility of the technique is established beyond any reasonable doubt. If the functions are highly nonlinear, the solution technique may not be adequately robust against signal noise and bad initial guesses. More experimental data in full three dimensions is necessary to study the robustness of the solution technique. Noise smoothing to remove the Gaussian instrument error can be used to increase the accuracy and the robustness of the solution technique.

In summary a new experimental technique to track the particle motion non-intrusively is presented. The feasibility of the method is established using computational and actual experimental results. It is an extremely useful technique having applications in many fields. The technique developed will directly be used in the study of inclined chute granular flows.

APPENDIX A

PROPERTIES OF BULK SOLIDS

Some of the important properties of bulk solids are as given below.

(1)   Granular and pulverized material sustain shear forces.

(2)   The pressure is not uniformly transmitted throughout the grain mass.

(3)   Coarse bulk solids (lumps of coal or ore) can have voids and thus are not
continuous, not even homogeneous.

(4)   Archimedes' principle of buoyancy does not hold for granular substances.

(5)   The angle of repose for granular materials is that which the material will assume
when piled on itself. The coefficient of static friction m of grain against grain is

$\mu = \tan\phi$

(6)   Ketchum (1919) found that the amount of grain discharged through an orifice in
the center of the bottom of the bin under gravity forces varies as the cube of the
opening

diameter. For fluids, the discharge varies as the square of the opening diameter.

(7)   The volume of bulk solids discharged through an orifice in the side of a bin near
the bottom varies as the diameter of the opening.

(8)   In both cases, the volume of flow is independent of the head above the orifice.

APPENDIX B

ARCHING PHENOMENA

The terms "arching", "bridging", and "doming" are used interchangeably to designate the cause of the stoppages of flow of bulk goods from bunkers, silos and hoppers. The terms are very descriptive since the free surface of the blocked material has the form of an arch, a bridge, or a dome. Thus one of the most important practical consequences of the cohesiveness of a bulk solid is that the material can develop sufficient "strength" to form a stable "arch" (or "bridge" or "dome") over an opening, even when the opening is very large in comparison to the particle size of the bulk solid concerned. The mechanical strength of arches is usually such that flow cannot be promoted by such means as poking by a crowbar from the top, bottom or side.

The main factors contributing to the tendency of a bulk solid to form a stable arch across an opening are the presence of very fine particles or of moisture, both of which increase the cohesiveness of the material. Compaction during storage also tends to increase the strength of the material and so aggravate the flow situation. Thus solids which run satisfactorily in the process of continuous operation may not flow after a period at rest owing to arching.

Two forms of stable arch can occur across an opening. A simple "mechanical arch" can develop directly as a result of interlocking of particles that are of large size compared with the opening. However, this problem can usually be avoided by ensuring that the hopper outlet is at least ten times the largest particle size. A "cohesive arch" is somewhat more difficult to predict as it forms as a result of the consolidation and strength of a cohesive bulk solid and can therefore occur even with materials of very fine particle size. The shear strength of a particulate bulk solid is a function of the consolidating pressure. Of special significance to the ability of a bulk solid to form a cohesive arch is the

"unconfined yield stress" which represents the strength of the material at the free surface. For a free-flowing (non-cohesive) material, such as dry sand, the unconfined yield stress is zero and therefore a cohesive arch could not occur.

The internal solids friction (grain against grain) is the cause of the pressure distribution (Stepanoff 1969). Internal friction is also the cause of arching. Since average grain velocities in the bin are negligible, the distinction between static and kinetic coefficients is negligible. All properties of solids affect the value of the coefficient of friction: nature of solid, grain size, grain screen test, grain shape, moisture, temperature, adhesion and cohesion. It should be noted that cohesion is independent of pressure. Also, the mobility of grain is reduced at higher temperatures, which is contrary to the behavior of Newtonian fluids. It has been established that solids graded as to size are more fluid than the mixture of sizes.

It is for these reasons that a knowledge of the flow behavior of bulk solids is essential when designing storage containers and other components of bulk handling installations.

APPENDIX C

SOLUTION OF ELLIPTIC INTEGRALS

The solution to the elliptic integrals which are present in model #2 discussed in chapter 3 is given below.

$$\int_0^{\pi/2} \frac{d\gamma}{\sqrt{1 - (k\sin\gamma)^2}} = \frac{\pi}{2}\left(1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1 \times 3}{2 \times 4}\right)^2 k^4 + \left(\frac{1 \times 3 \times 5}{2 \times 4 \times 6}\right)^2 k^6 + \ldots\right)$$

$$\int_0^{\pi/2} \sqrt{(1 - (k\sin\gamma)^2)}\, d\gamma = \frac{\pi}{2}\left(1 - \left(\frac{1}{2^2}\right) k^2 - \left(\frac{1^2 \times 3}{2^2 \times 4^2}\right) k^4 - \left(\frac{1^2 \times 3^2 \times 5}{2^2 \times 4^2 \times 6^2}\right) k^6 - \ldots\right)$$

where $k^2 < 1$.

APPENDIX D

HOMOGENEOUS COORDINATES AND TRANSFORMATION MATRIX

Since a 3 x 3 rotation matrix does not give us any provision for translation and scaling, a fourth coordinate or component is introduced to a position vector $p = (p_x, p_y, p_z)^T$ in a three-dimensional space which makes it $\hat{p} = (wp_x, wp_y, wp_z, w)^T$. The position vector $\hat{p}$ is said to be expressed in *homogeneous coordinates*. The concept of a homogeneous-coordinate representation of points in a three-dimensional euclidean space is useful in developing matrix transformations that include rotation, translation, scaling and perspective transformation. In general, the representation of an N-component position vector by an $(N+1)$ - component vector is called *homogeneous coordinate representation* (Fu, Gonzalez & Lee 1987). In a homogeneous coordinate representation, the transformation of an N-dimensional vector is performed in the $(N+1)$-dimensional space, and the physical N-dimensional vector is obtained by dividing the homogeneous coordinates by the $(N+1)$th coordinate, w. Thus, in a three-dimensional space, a position vector $p = (p_x, p_y, p_z)^T$ is represented by an augmented vector $(wp_x, wp_y, wp_z, w)^T$ in the homogeneous coordinate representation. The physical coordinates are related to the homogeneous coordinates as follows:

$$p_x = \frac{wp_x}{w}, \quad p_y = \frac{wp_y}{w}, \quad p_z = \frac{wp_z}{w}$$

There is no unique homogeneous coordinates representation for a position vector in the three-dimensional space. Thus, the fourth component of the homogeneous coordinates, w, can be viewed as a scale factor.

The homogeneous transformation matrix is a 4 x 4 matrix which maps a position vector expressed in homogeneous coordinates from one coordinate system to another coordinate system. A homogeneous transformation matrix can be considered to consist

of four submatrices:

$$T = \begin{bmatrix} R_{3 \times 3} & P_{3 \times 1} \\ \hline f_{1 \times 3} & 1 \times 1 \end{bmatrix} = \begin{bmatrix} \text{Rotation Matrix} & \text{Position Vector} \\ \hline \text{Perspective Transformation} & \text{Scaling} \end{bmatrix}$$

The upper left 3 x 3 submatrix represents the rotation matrix; the upper right 3 x 1 submatrix represents the position vector of the origin of the rotated coordinate system with respect to the reference system; the lower left 1 x 3 submatrix represents perspective transformation; and the fourth diagonal element is the global scaling factor.

A homogeneous transformation matrix geometrically represents the location of a rotated coordinate system (position and orientation) with respect to a reference coordinate system.

APPENDIX E

CONVERGENCE OF SEQUENCES OF REAL NUMBERS

Given an iterative method that produces a sequence of points $x_1, x_2,...,$ from a starting guess x_0, it is useful to know if the iterates converge to a solution x_*, and if so, how quickly. If we assume that we know what it means to write

$$\lim_{k \to \infty} a_k = 0$$

for a real sequence $\{a_k\}$, then the following definition characterizes the properties we need.

Let $x_* \in R$, $x_k \in R$, $k = 0, 1, 2,....$ Then the sequence $\{x_k\} = \{x_0, x_1, x_2,...\}$ is said to *converge* to x* if

$$\lim_{k \to \infty} |x_k - x_*| = 0.$$

If in addition, there exists a constant $c \in [0, 1)$ and an integer $\hat{k} \geq 0$ such that for all $k \geq \hat{k}$,

$$|x_{k+1} - x_*| \leq c|x_k - x_*|$$

then $\{x_k\}$ is said to be *q-linearly convergent* to x_*. If for some sequence $\{c_k\}$ that converges to 0,

$$|x_{k+1} - x_*| \leq c_k|x_k - x_*|,$$

then $\{x_k\}$ is said to converge *q-superlinearly* to x_*. If there exist constants $p > 1$, $c \geq 0$, and $\hat{k} \geq 0$ such that $\{x_k\}$ converges to x_* and for all $k \geq \hat{k}$,

$$|x_{k+1} - x_*| \leq c|x_k - x_*|^p,$$

then $\{x_k\}$ is said to converge to x_* with *q-order at least p*. If $p = 2$ or 3, the convergence is said to be *q-quadratic* or *q-cubic*, respectively (Dennis & Schnabel 1983).

REFERENCES

Ahn, H., C. E. Brennen, and R. H. Sabersky. 1989. "Measurements of Velocity, Veloc ity Fluctuations, Density and Stresses in Chute Flows of Granular Materials." *Geophysical Grain Flows*. Scripps Institution of Oceanography. University of California. San Diego. July 13-15.

Allegri, T. H. 1984. *"Materials Handling-Principles and Practice."* Van Nostrand Reinhold Company. New York.

Altobelli, S. A., A. Caprihan, E. Fukushima, M. Nakagawa, and R. C. Givler. 1991. "Multiphase Flow Studies by NMR with Application to Granular Flows." *Joint DOE/NSF Workshop on Flow of Particles and Fluids*. Worcester. Mass. October 22-24. 34-40.

Armijo, L. 1966. "Minimization of functions having Lipschitz-continuous first partial derivatives." *Pacific J. Math.* 16: 1-3.

Beck, M. S., R. G. Green, and R. Thorns. 1987. "Non-intrusive measurement of solids mass flow in pneumatic conveying." *Journal of Physics. E. Scientific Instruments*. July. 20: 835-40.

Bransby, P. L., P. M. Blair-Fish, and R. G. James. 1973. "An investigation of the flow of granular materials." *Powder Technology*. 8: 197-206.

Burkett, R. J., C. Dixon, P. J. Morris, and D. L. Pyle. 1971. *Chem Engg. Sci.* 26: 405.

Byrd, R. H., J. Nocedal and Y. X. Yuan. 1987. " Global convergence of a class of quasi-Newton methods on convex problems." *SIAM J. Numer. Anal.* October. No. 5. 24: 1171-90.

Dave, R. N., A. S. Ashok, and B. G. Bukiet. 1992. "On Development of a Three Dimensional Particle Motion Tracking System." submitted to *ASME Winter Annual Meeting*. Nov.

Dennis, J. E., and J. J. More. 1977. "Quasi-Newton methods, motivation and theory." *SIAM Rev.* 19: 46-89.

Dennis, J. E., and R. B. Schnabel. 1983. *Numerical Methods for Unconstrained Opti mization and Nonlinear Equations*. Prentice-Hall.

Drake, T. G. 1991. "Granular flow: physical experiments and their implications for microstructural theories." *J. Fluid Mech.* 225: 121-152.

Fletcher, R. 1987. *"Practical Methods of Optimization."* 2nd ed. John Wiley & Sons. New York.

Fu, K. S., R. C. Gonzalez, and C. S. G. Lee. 1987. *"Robotics: Control, Sensing, Vision And Intelligence."* 1st ed. McGraw-Hill. New York.

Goldstein, A. A. 1967. *"Constructive Real Analysis."* Harper & Row. New York.

Griffiths, D.J. 1989. *"Introduction to Electrodynamics."* Prentice-Hall. Englewood Cliffs. New Jersey. Second Edition.

Hancock, A. W. 1970. Ph.D Thesis. University of Cambridge.

Handley, M. F., and M. G. Perry. 1967. *Powder Technology.* 1: 245.

Hayden, H. C. 1987. "Data Smoothing Routine." *Computers in Physics.* Nov/Dec. 74-75.

Hebden, M. D. 1973. "An algorithm for minimization using exact second derivatives." Rept. TP515. A. E. R. E. Harwell. England.

Katsuya, O., T. Akehata, and S. Takashi. 1975. "A New Method for Evaluating the size of Moving Particles With a Fiber Optic Probe." *Powder Technology.* 11: 51-57.

Ketchum. 1919. *"Walls, Bins and Grain Elevators."* 3rd ed. McGraw-Hill. New York.

Levenberg, K. 1944. "A method for the Solution of Certain Problems in Least Squares." *Quart. Appl. Math.* 2: 164-168.

Lin, J. S., M. M. Chen, and B. T. Chao. 1985. "A novel radioactive particle tracking facility for measurement of solids motion in gas fluidized beds." *AIChE Journal.* March. 31: 465-73.

Locket, M. J. 1967. Ph.D Thesis. University of Cambridge.

Louge, M., D. J. Lischer, and H. Chang. 1990. "Measurements of Voidage near the Wall of a Circulating Fluidized Bed Riser," *Proceedings of NAFIPS'90.* 121-124.

Louge, M., S. A. Iyer, E. P. Giannelis, D. J. Lischer, and H. Chang. 1991. "Optical Fiber Measurements of Particle Velocity Using Laser-Induced-Phosphorescence." *Research Report.* Cornell University. Ithaca. NY.

McCabe, R. P. 1974. *Geotechnique.* 1: 45.

Marquardt, D. 1963. "An algorithm for Least Squares Estimation of Nonlinear Parame ters." *SIAM J. Appl. Math.* 11: 431-441.

More, J. J. 1977. "The Levenberg-Marquardt Algorithm: Implementation and Theory." in *Numerical Analysis.* G. A. Watson. ed. Lecture Notes in Math. 630. Springer Verlag. Berlin. 105-106.

More, J. J., B. S. Garbow, and K. E. Hillstrom. 1980. "User Guide for MINPACK-1." Argonne National Lab Report ANL-80-74.

Oki, K., W. P. Walawender, and L. T. Fan. 1992. *Powder Technology.* 18: 171.

Parasar, A. 1992. Personal communications.

Perry, M. G., H. A. S. Jangda. 1970. *Powder Technology.* 4: 89.

Powell, M. J. D. 1970. "A hybrid method for nonlinear equations." in *Numerical Meth ods for Nonlinear Algebraic Equations.* P. Rabinowitz. ed. Gordon and Breach. London. 87-114.

Rao, V. L., and D. Venkateswarlu. 1975. "Internal Pressures in Flowing Granular Mate rials from Mass Flow Hoppers." *Powder Technology.* 11: 133-146.

Reinsch, C. 1971. "Smoothing by spline functions, II." *Numer. Math.* 16: 451-454.

Riley, A. C., and M. Louge. 1989. "Quantitative Capacitance Measurements of Voidage in Gas-Solid Flows." *Particulate Science and Technology.* 7: 51-59.

Rosato, A. D., R. N. Dave, I. S. Fischer, and W. N. Carr. 1991. "Methodology of a Radiosonde System for the Non-Intrusive Tracking of a Particle in a Flow." *Joint DOE/NSF Workshop on Flow of Particulates and Fluids.* Worcester. Mass. Octo ber 22-24. 94-108.

-----. 1992. "Development of a Non-Intrusive Particle Tracing Technique for Granular Chute Flows." *Quarterly Progress Report.* USDOE contract DE-AC22-91PC90181. Pittsburgh Energy Technology Center. April.

Savage, S. B. 1978. "Experiments on Shear Flows of Cohesionless Granular Materi als." *Proceedings of the US-Japan Seminar on Continuum-Mechanical and Statis tical Approaches in the Mechanics of Granular Materials* (eds. S. Cowin and M. Satake). 241-254.

Smallwood, M. A., and R. B. Thorpe. 1980. *Tripos Part 2.* Chemical Engineering Research Project. University of Cambridge.

Stepanoff, A. J. 1969. *"Gravity flow of bulk solids and transportation of solids in sus pension."* 1st ed. John Wiley & Sons. Inc. New York.

Troiano, A. 1992. Personal communications.

Tug, H. "Position measurements with Radiosondes. " IEEE Journal of Oceanic Engi neering. April. No.2. 14: 208-10.

Tuzun U., G. T. Houlsby, R. M. Nedderman, and S. B. Savage. 1982. "The flow of Granular Materials - II." *Chem. Engg. Sci.* (12). 37: 1691-1709.

Walton, O. R. 1988. "LLNL Granular Solids Flow Project Summary Report." UCID-20297-88-1. Lawrence Livermore Labs. May.

Woodcock, C. R., and J. S. Mason. 1987. *"Bulk Solids Handling-An Introduction to the Practice and Technology."* Blackie & Son. London.