Nonlinear interactions of internal gravity waves in a continuously stratified fluid

A thesis submitted in partial
fulfilment of the requirements for
the degree of BSc. (Honours)

by

Andrew Elek Kiss

Supervisors:

Dr. Ross Griffiths,
Geophysical Fluid Dynamics Group,
Research School of Earth Sciences
Dr. Frank Houwing,
Department of Physics and Theoretical Physics

The Australian National University

November 14, 1994

Acknowledgments:

I would like to express my sincere thanks to Tony Beasley, Derek Corrigan and Ross Wilde-Brown for their patience and technical assistance, without whose superb craftsmanship and know-how none of this would have been possible. Many thanks also to Dr. Ross Griffiths for his expert supervision and many stimulating discussions throughout the year. I also owe thanks to Dr. Frank Houwing for his able co-supervision, and to Prof. Stuart Turner for his advice during the year. I am indebted to Dr. Dave Broutman for his greatly appreciated help with the theory. Thanks also to Dr. Ross Kerr and Prof. Brian Kennett for their helpful comments on my thesis.

I would like to thank my parents and brother for their support (and emergency food drops), and finally a big "thank-you" to Sarah who was always ready to listen and helped me retain more than a shred of sanity.

I gratefully acknowledge the support of an Anton Hales Honours scholarship granted by RSES.

Declaration:

Except where specified, the work contained herein is entirely my own.

Abstract

Internal gravity waves are buoyancy-driven oscillations which can arise in a density-stratified fluid. They exist throughout the oceans and atmosphere, the oceanic internal wavefield being sufficiently energetic for nonlinear effects to play an important role in the internal wave dynamics.

Numerical studies of oceanic internal waves (such as Broutman & Young, 1986) have suggested that under the right conditions a weak internal gravity wave can be strongly refracted and frequency-shifted by the time-varying shear of a large-amplitude internal wave.

My project aimed to experimentally observe this type of strongly nonlinear interaction by generating the required internal gravity waves in a continuously stratified aqueous solution. The waves were observed using a colour schlieren system, and power spectra of the internal wavefield were obtained using conductivity probes and polarimetry.

Several nonlinear phenomena were observed, including anharmonic waves and forced sum and difference frequencies, as well as second-harmonic generation from the wave sources. However a combination of wavelength limitations imposed by viscosity, inescapable restrictions on the strong wave amplitude and severe observational difficulties all conspired to prevent detection of the particular nonlinear interaction of interest. A proposed apparatus could overcome these difficulties, but its construction would be quite beyond the scope of an Honours project.

Contents

Chapter 1: Introduction 1

1.1 What are internal gravity waves? 1
1.2 Motivation: oceanic internal waves 1
1.3 Theoretical efforts 3
1.4 Project aims 3
1.5 Thesis overview 3

Chapter 2: Wave theory 4

2.1 Inviscid linear wave theory 4
2.2 Ray theory 9
2.3 The effects of viscosity 12
2.4 Weak nonlinear interactions 15
2.5 Strong nonlinear interactions 16
2.6 Adaptation of Broutman & Young's theory 22

Chapter 3: Experimental methods 26

3.1 Overview of the apparatus 26
3.2 Setting up a density gradient 29

Chapter 4: Wave observation methods 32

4.1 Overview 32
4.2 Shadowgraph 32
4.3 Moiré technique 34
4.4 Schlieren techniques 34
4.5 The conductivity probe 38
4.6 Polarimetry 40
4.7 Data acquisition 43

Chapter 5: Results 45

5.1 Observed wave properties 45
5.2 Schlieren results 47
5.3 Digital image subtraction 51
5.4 Point probe results 52

Chapter 6: Discussion 58

6.1 Why it didn't work 58
6.2 How it could be done (maybe) 60

Conclusion 64

Appendix 1: Derivation of the basic properties of internal gravity waves**......65**
A1.1 The equations of motion..65
A1.2 The hydrostatic equilibrium state..66
A1.3 Linearisation..67
A1.4 The dispersion relation..68
A1.5 The nature of the wavefield..69

Appendix 2: The effects of diffusion 71

Appendix 3: LabVIEW data acquisition program..75

Appendix 4: Second harmonic generation..77

References 79

List of figures

Fig. 1.2.1: The GM75 internal wave spectrum compared to observations. 2

Fig. 2.1.1: Surfaces of constant w in wavenumber space. 6
Fig. 2.1.2: The structure of a plane internal gravity wave.................... 7
Fig. 2.1.3: Internal waves spreading from a small source.................... 8
Fig. 2.2.1: Wave refraction and trapping due to a variation in N. 11
Fig. 2.3.1: Definition of coordinates aligned with the downgoing wave beam.................... 13
Fig. 2.3.2: Similarity displacement profiles at different times.................... 14
Fig. 2.3.3: Curves of constant η for my experiments. 14
Fig. 2.5.1: Coordinates aligned with the background wavefield.................... 17
Fig. 2.5.2: Test ray refraction by caustics. 19
Fig. 2.5.3: Contours of constant ω'' vs. background phase ξ. 21
Fig. 2.6.1: Normalised G vs. ω for various Ω. .... opp. 24
Fig. 2.6.2: Normalised m_c'/m_i' vs. ω for various Ω. .... opp. 24

Fig. 3.1.1: The experimental tank.................... 26
Fig. 3.1.2: The crank system for oscillating a wavemaker.................... 27
Fig. 3.1.3: Cross-sections of some of the wavemakers used.................... 28
Fig. 3.2.1: Tank filling apparatus.................... 29

Fig. 4.2.1: Refractive index of salt and sugar solutions vs. density.................... 33
Fig. 4.2.2: A shadowgraph system.................... 33
Fig. 4.4.1: A typical schlieren system.................... 35
Fig. 4.4.2: Shift in the slit image due to refraction.................... 35
Fig. 4.4.3: The colour schlieren system. 36
Fig. 4.5.1: The relationship between density and conductivity of a salt solution. 39
Fig. 4.5.2: The four-electrode conductivity probe (scale diagram). 39
Fig. 4.5.3: Calibration of the conductivity probe.................... 40
Fig. 4.6.1: Optical rotation at 589 nm vs. density of a sucrose solution.................... 41
Fig. 4.6.2: Schematic of the polarimeter.................... 42

Fig. 5.1.1: $\cos\Theta$ vs. Ω for N=1.65 rad/s. 47
Fig. 5.2.1: Schlieren image of tank at rest.................... opp. 47
Fig. 5.2.2: Schlieren image of test wave. .... opp. 47
Fig. 5.2.3: Schlieren image of background wave. .... opp. 47
Fig. 5.2.4: A colour schlieren slide (twice actual size). .... opp. 48
Fig. 5.2.5: Wave interaction geometry.................... 49
Fig. 5.2.6: Schlieren image of test waves with strong background waves.................... opp. 48
Fig. 5.3.1: Digital image subtraction.................... opp. 51
Fig. 5.4.1: A typical background wave power spectrum.................... 53
Fig. 5.4.2: Typical power spectrum with test and background waves.................... 56

Fig. 6.1.1: Normalised G vs. Ω/ω for various Ω/N.................... opp. 58
Fig. 6.1.2: Normalised $1/U'_c$ vs. ω for various Ω, with κ/M'=7/4.................... opp. 59
Fig. 6.2.1: Normalised $1/U'_c$ vs. ω for various Ω, with κ/M'=6.................... opp. 62
Fig. 6.2.2: Proposed wave-interaction apparatus.................... 61

Fig. A2.1: The effect of diffusion on a density gradient in salt water. 72
Fig. A2.2: The effect of diffusion on the buoyancy frequency N.................... 73
Fig. A2.3: Ray refraction in a diffused salinity gradient.................... 74

Fig. A4.1: Wavevectors for a resonant triad.................... 78

Glossary of symbols

Most capital letters refer to the background waves, and lowercase to the test wave or a generic wave. Dashed quantities are measured with respect to the rotated coordinates (x', z').
Subscript c refers to the value of a quantity at a caustic.

Roman

$\boldsymbol{c} = \dfrac{\omega \boldsymbol{k}}{\kappa^2}$	phase velocity
$\boldsymbol{c}_g = \nabla_{\boldsymbol{k}}\omega$	group velocity
$c_{gz'} \equiv \lvert\boldsymbol{C}\rvert^{-1}\left(\boldsymbol{c}_g \cdot \boldsymbol{C}\right)$	component of test wave group velocity parallel to background phase velocity
$G = c_{gz'}/C$	
$\boldsymbol{g} = -g\hat{z}$	acceleration due to gravity
$\boldsymbol{k} = (k,l,m)$	wavevector
M'	magnitude of background wavevector
m_c'	test wavenumber in y' direction at a caustic
m_i'	incident test wavenumber in y' direction
N	buoyancy frequency $N^2 = \dfrac{1}{\rho_o}\boldsymbol{g}\cdot\nabla\rho_o = -\dfrac{g}{\rho_o}\dfrac{\partial\rho_o}{\partial z}$
n	refractive index
$p = p_o + p'$	pressure (p_o is the hydrostatic equilibrium pressure)
$\boldsymbol{r} = (x,y,z)$	position vector
S	solute concentration
t	time
$\boldsymbol{u} = (u,v,w)$	velocity
U_c'	background flow speed at a caustic
$\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$	unit vectors of Cartesian coordinates; $\hat{z}$ points vertically upwards

Greek

θ	angle of $\boldsymbol{k}$ to the horizontal
κ	magnitude of Re($\boldsymbol{k}$)
$\lambda = 2\pi/\lvert\boldsymbol{k}\rvert$	wavelength
$\rho = \rho_o + \rho'$	density (ρ_o is the hydrostatic equilibrium density)
ω	angular frequency

Other

$\nabla_h = (\partial/\partial x, \partial/\partial y, 0)$	horizontal gradient operator
$\nabla_k = (\partial/\partial k, \partial/\partial l, \partial/\partial m)$	gradient acting in wavenumber space

Chapter 1: Introduction

1.1 *What are internal gravity waves?*

There are many phenomena in fluids which are driven by buoyancy forces arising from the action of gravity on density differences in that fluid (Turner, 1973) - the most familiar example would be the convection which arises in an unstable density stratification. In a fluid with a *stable* density stratification (that is, with density increasing with depth), it is possible to have wave motion as an oscillation about that stratification, since a fluid element displaced from its equilibrium depth will experience a gravitational restoring force tending to return it to its original depth. Such waves are called *internal gravity waves* as they can propagate through the interior of a stratified fluid and are driven by gravity. My project is concerned with the interaction of internal gravity waves in an aqueous solution whose density increases *continuously* with depth. More details on these waves and their interactions are revealed in Chapter 2 - for the moment it is sufficient to say that the waves involve oscillations of the fluid velocity, density and pressure about the hydrostatic equilibrium state.

1.2 *Motivation: oceanic internal waves*

The Earth's oceans and atmosphere are stably stratified for much of their depth and internal gravity waves propagating through these regions have great oceanographic and meteorological importance (Müller *et al.*, 1986). Below the shallow well-mixed layer at the surface of the oceans, the water density increases with depth right down to the sea floor, allowing internal gravity waves to exist throughout the bulk of the oceans (Lighthill, 1978). Similarly, the atmosphere is stably stratified above a mixed layer near the ground. Density stratification in the oceans is due chiefly to changes in temperature and salinity, whilst that of the atmosphere is more the result of variation in *potential temperature* with height (that is, the temperature distribution which would result if the entire atmosphere were adiabatically brought to the same pressure) (Lighthill, 1978; Gill, 1982). This project was concerned with wave phenomena more relevant to oceanography than to meteorology.

Internal waves are found throughout the oceans, as a random superposition of different amplitudes, wavenumbers and frequencies. The range of vertical wavelengths is from around one meter to about one kilometre, whilst the horizontal wavelengths can be from a few metres up to tens of kilometres. (Müller, *et al.*, 1986). Typical velocities and vertical displacements are around 5 cm/s and 7 m, respectively, and wave periods can range from several minutes up to 12 hours (Lighthill, 1978; LeBlond & Mysak, 1978; Garrett & Munk, 1972, 1975). The wave periods are long because the restoring buoyancy force which drives the waves is very small in weakly stratified fluids such as the oceans.

Internal waves are very important in oceanography for many reasons, the main one being that they are thought to play a crucial role in the overall energy balance of the oceans by transporting energy from the planetary scales at which it is generated to the very small scales at which the energy is dissipated by molecular processes (Müller *et al.*, 1986).

Measurements of oceanic currents and temperature fluctuations due to internal wave motions have been taken since the 1920's (Briscoe, 1975). Since then, a great deal of data has been collected from field surveys throughout the oceans, which has given researchers a quite detailed picture of the oceanic internal wavefield. By 1970 essentially all the oceans had been surveyed in all seasons, and most of the data had been spectrally analysed. However it was not until 1972 that an attempt was made to characterise the essential features of the internal wavefield throughout the oceans by single model (Garrett & Munk, 1972). The Garrett & Munk (GM) spectra were essentially power-law fits to the available data, showing wavefield energy as a function of frequency and of vertical and horizontal wavenumber. As more detailed observations have been made, this empirical model has undergone a number of revisions, such as the 1975 version shown in Fig. 1.2.1 (Garrett & Munk, 1975).

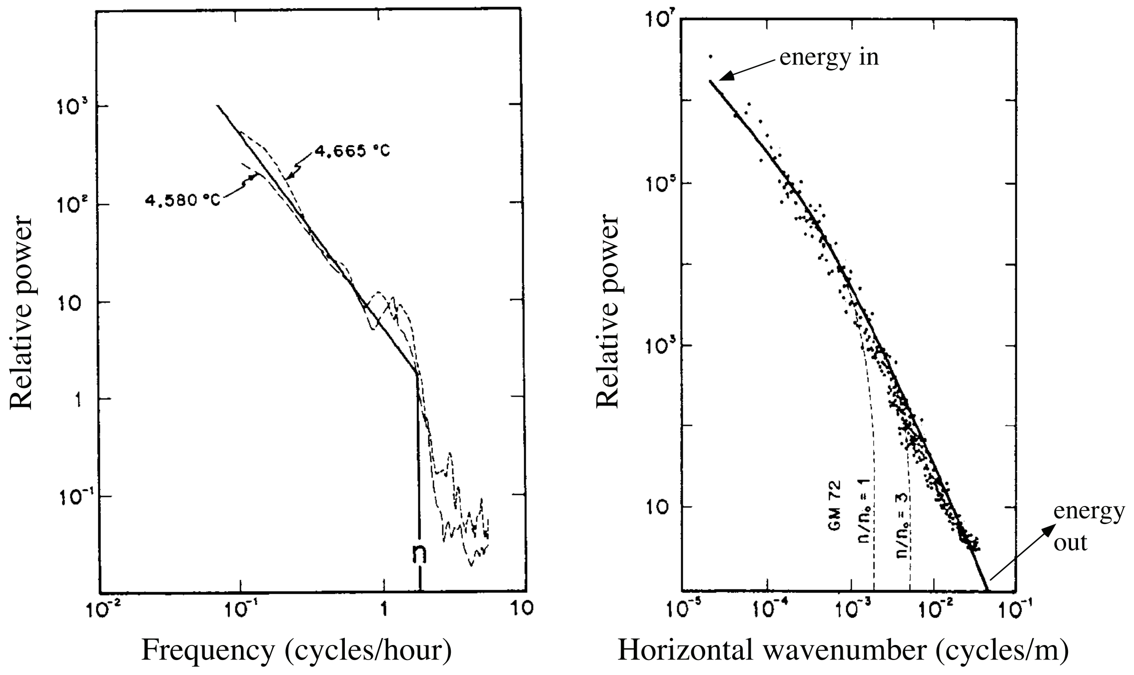


Fig. 1.2.1: The GM75 internal wave spectrum compared to observations (adapted from Garrett & Munk, 1975).

Observations of internal wave spectra throughout the oceans show a remarkable agreement with the "universal" Garrett & Munk model spectrum of 1975 (see Fig. 1.2.1), and distorted spectra rapidly relax to this universal form (Müller *et al.*, 1986). These observations suggest that the GM spectrum is close to a steady state, with energy transfers between different parts of the spectrum maintaining the equilibrium spectral shape. The sources and sinks of internal wave energy also have a bearing on the spectral shape, but since these are likely to be more localised and sporadic in nature, and occur at quite different parts of the spectrum, some redistribution process in the wavefield is also needed to account for the equilibrium. To oversimplify, the hypothesis is that wave energy is transported from the large planetary scales at which it is generated to the microscales where it is dissipated, and this energy cascade from large to small scales comes about through interactions between internal waves of different wavenumbers (LeBlond & Mysak, 1978; Müller *et al.*, 1986).

1.3 Theoretical efforts

Since energy transport through the internal wave spectrum is forbidden by linear theory, such transport is an intrinsically *nonlinear* process. The problem of explaining the origin and stability of the empirical GM spectrum has inspired a great deal of theoretical work on nonlinear interactions between internal gravity waves. Such investigations have often involved statistical methods to model the transport of energy through the continuum of waves in the spectrum via the "resonant triads" which arise in weakly nonlinear interactions. These allow groups of three waves to interact strongly with each other only if they satisfy certain selection criteria, and thus restricts the number of interactions which need to be taken into account.

However Holloway (1980, 1982) and others pointed out that oceanic internal waves are about 100 times too energetic for weakly nonlinear theory to be applicable. This meant that a full understanding of the oceanic internal wavefield would probably require a theory incorporating *strong* nonlinear interactions, which allow much more indiscriminate energy transport through the spectrum than is possible with resonant triads. Theoretically describing the spectral evolution under the influence of such indiscriminate energy exchange is a daunting problem. As a result, several researchers have investigated strongly nonlinear processes which can occur when only a few waves are present (rather than a continuous spectrum), in order to understand some of the basic processes which can occur.

1.4 Project aims

In this project I attempted to observe experimentally some of the strongly nonlinear interactions predicted by the theoretical investigations of Broutman and others (Broutman, 1984, 1986; Broutman & Grimshaw, 1988; Broutman & Young, 1986; Macaskill & Broutman, 1988; Broutman & McIntyre, 1994; Thorpe, 1989). Waves were generated in a stratified salt or sugar solution and observed chiefly by optical methods, providing both flow visualisation and frequency spectra.

1.5 Thesis overview

The following chapter presents the theoretical issues which needed to be addressed in designing and conducting my experiments. The details of strongly nonlinear theory and both inviscid and viscous linear theory were of crucial importance to my project.

Chapter 3 presents the apparatus used for producing both internal gravity waves and stratified aqueous solutions, whilst Chapter 4 discusses the numerous methods used for their observation.

The experimental results are presented in Chapter 5 and discussed in Chapter 6.

Chapter 2: Wave theory

2.1 Inviscid linear wave theory

To understand the behaviour of nonlinear internal gravity waves we must first grasp the very unusual properties of these waves as predicted by linear theory, which applies in the case of infinitesimal wave amplitude. A fairly rigorous derivation of the properties of infinitesimal plane internal gravity waves is given in Appendix 1. This section will summarise the most important results of that derivation, and provide them with a physical interpretation. The results are all quite standard and can be found in any good physical oceanography text, such as Gill (1982) or LeBlond & Mysak (1978).

The fluid properties are described as functions of time t and of position $\boldsymbol{r}=(x,y,z)$ relative to right-handed Cartesian coordinate axes which are fixed to an inertial reference frame. The coordinates are chosen so that the z axis points vertically upwards. The fluid is characterised by the density $\rho(\boldsymbol{r},t)$, pressure $p(\boldsymbol{r},t)$ and velocity $\boldsymbol{u}(\boldsymbol{r},t)$, and is assumed to be inviscid. Furthermore the conduction of heat, diffusion of solute and compressibility of the fluid are neglected (as these are seen in retrospect to be insignificant on the time and distance scales of the waves), so the density of a fluid element is constant. Gravity is assumed to be the sole external force, so the inviscid fluid is driven only by buoyancy forces and pressure gradients.

When at rest (ie. $\boldsymbol{u}(\boldsymbol{r},t)=\boldsymbol{0}\ \forall \boldsymbol{r},t$), the pressure and density of the fluid are denoted p_o and ρ_o, respectively and are functions of z only. A very important parameter governing the motion of a density-stratified fluid under gravity is the *buoyancy frequency N* defined by

$$N^2 = \frac{1}{\rho_o}\boldsymbol{g}\cdot\nabla\rho_o = \frac{-g}{\rho_o}\frac{\partial\rho_o}{\partial z} \qquad (2.1.1),$$

where $\boldsymbol{g}$ is the acceleration due to gravity and g is its magnitude. The buoyancy frequency is a measure of the stability of a stratified fluid: it is real for stably stratified fluids, and imaginary if the stratification is unstable. If an element of fluid in a stably stratified fluid is given a small vertical displacement from its equilibrium depth, it will oscillate vertically about that equilibrium depth with a frequency N. This frequency is larger for steeper density gradients, as this increases the restoring force arising from a given vertical displacement. The imaginary value of N in an unstable stratification indicates that arbitrarily small displacements from equilibrium (initially) increase exponentially with time.

When the fluid is in motion, the *perturbations* p', ρ' to the resting state are defined by

$$p = p_o + p', \quad \rho = \rho_o + \rho' \qquad (2.1.2).$$

If the perturbations $\boldsymbol{u}$, p' and ρ' to the resting state are infinitesimally small, products of these terms may be neglected from the equations of motion, a procedure known as *linearisation.*

The linearised equations of motion admit plane wave solutions when N is constant (ie. the fluid density increases exponentially with depth), with sinusoidal variations in $\boldsymbol{u}$, p' and ρ' given by

$$\boldsymbol{u} = \left(\frac{-km}{k^2 + l^2}, \frac{-lm}{k^2 + l^2}, 1 \right) w(\boldsymbol{r}, t) \qquad (2.1.3),$$

$$p' = \frac{-\omega m \rho_o}{k^2 + l^2} w(\boldsymbol{r}, t) \qquad (2.1.4)$$

and

$$\rho' = \frac{i \rho_o N^2}{\omega g} w(\boldsymbol{r}, t) \qquad (2.1.5)$$

respectively. Here

$$w(\boldsymbol{r}, t) = W \exp\left[i(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t) \right] \qquad (2.1.6)$$

is the vertical component of the velocity field, $\boldsymbol{k}=(k,l,m)$ is the wavevector, ω the wave frequency and W is an arbitrary small constant. We use the usual convention whereby the actual physical quantity is the real part of any complex expression.

The expression for $\boldsymbol{u}$ shows that the waves are plane polarised, and it is easy to show that $\boldsymbol{u}$ and $\boldsymbol{k}$ are perpendicular, that is, the waves are *transverse*. The perturbation density field ρ' arises simply from advection of the background density ρ_o by the vertical velocity field w. Note that the factor of i in (2.1.5) indicates a $+\pi/2$ phase shift in ρ' relative to p' and w. A representation of these fields is shown in Fig. 2.1.2, and will be explained more fully later.

The wavevector $\boldsymbol{k}$ and frequency ω are connected by the *dispersion relation* which, for short waves such that $|\boldsymbol{k}| >> N^2/2g$, may be written

$$\omega = N \cos\theta \qquad (2.1.7),$$

where θ is the angle of $\boldsymbol{k}$ to the horizontal (see Figs. 2.1.1 & 2.1.2). Note that N forms an upper limit on the possible frequencies of infinitesimal internal gravity waves. The dispersion relation shows that short internal gravity waves are extremely anisotropic: the frequency depends only on the *direction* of the wavevector, *not* its magnitude. Thus the wavelength $\lambda=2\pi/|\boldsymbol{k}|$ is *independent* of ω - a disturbance at a given frequency can produce waves of *any* wavelength, provided they are sufficiently short. Surfaces of constant ω in wavenumber space are surfaces of constant θ, ie. *cones*. This is illustrated in Fig. 2.1.1.

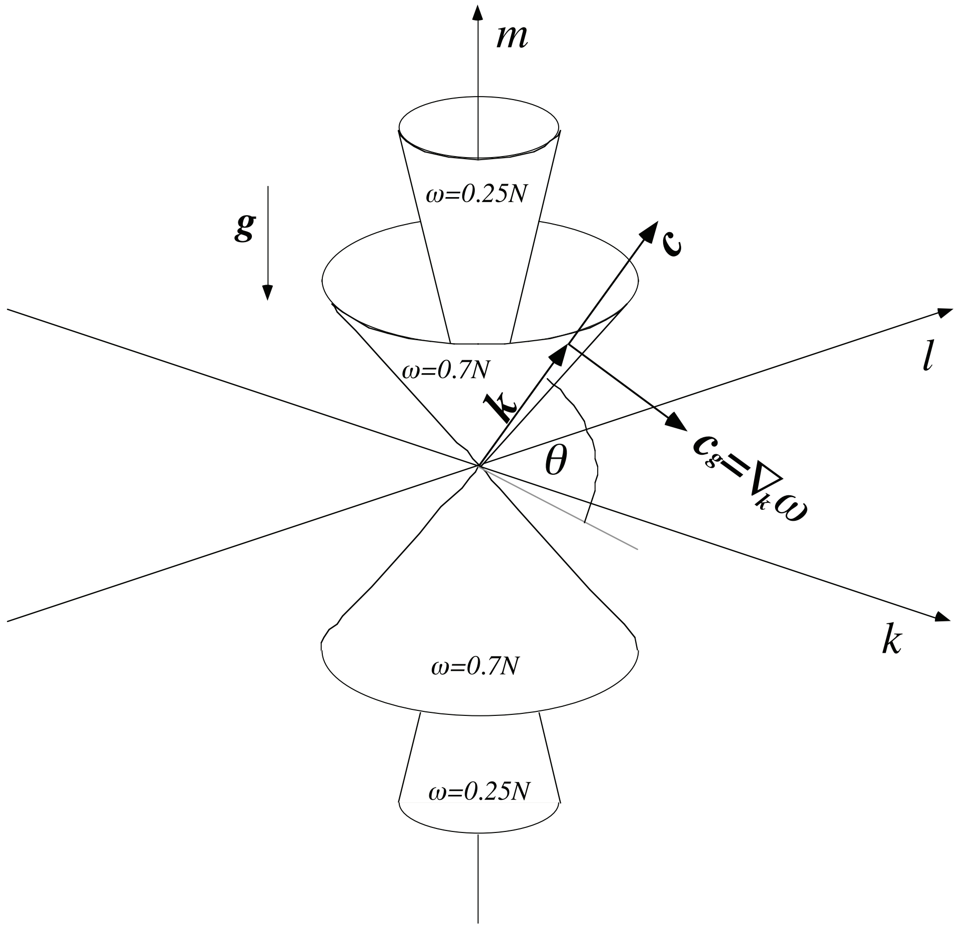


Fig. 2.1.1: Surfaces of constant ω in wavenumber space.

The *phase velocity* **c** (the propagation velocity of surfaces of constant phase) is defined by

$$\boldsymbol{c} = \frac{\omega \boldsymbol{k}}{\kappa^2} \tag{2.1.8}$$

(where κ is the magnitude of **k**) and is thus parallel to **k**. The *group velocity* $\boldsymbol{c}_g$ is defined as the gradient of ω in wavenumber space, and is thus perpendicular to the conical surfaces of constant ω. Hence *the group and phase velocities are perpendicular*, as shown in Fig. 2.1.1.

Analytically, taking the gradient with respect to **k** of the dispersion relation gives

$$\boldsymbol{c}_g = \frac{Nm\sqrt{k^2+l^2}}{\kappa^3}\left(\frac{km}{k^2+l^2}, \frac{lm}{k^2+l^2}, -1\right) \tag{2.1.9},$$

showing (by comparison with (2.1.3)) that $\boldsymbol{c}_g$ is parallel to **u**, and that

$$\boldsymbol{c}_g \cdot \boldsymbol{c} = 0 \tag{2.1.10}.$$

In contrast, an *isotropic* dispersion relation would imply surfaces of constant frequency which are *spheres* centred on the origin. The group velocity is normal to these surfaces and is thus parallel to the phase velocity - a much more familiar situation.

A vertical slice though a portion of a plane internal gravity wave at an instant in time is shown in Fig. 2.1.2. Wavefronts (lines of constant phase) extend from the bottom right to the top left and propagate towards the lower left with the phase velocity $\boldsymbol{c}$. The variation of the three fields along lines perpendicular to the wavefronts is shown. Although the wavefronts are moving to the lower left, the wave energy is propagating to the upper left with the group velocity $\boldsymbol{c}_g$.

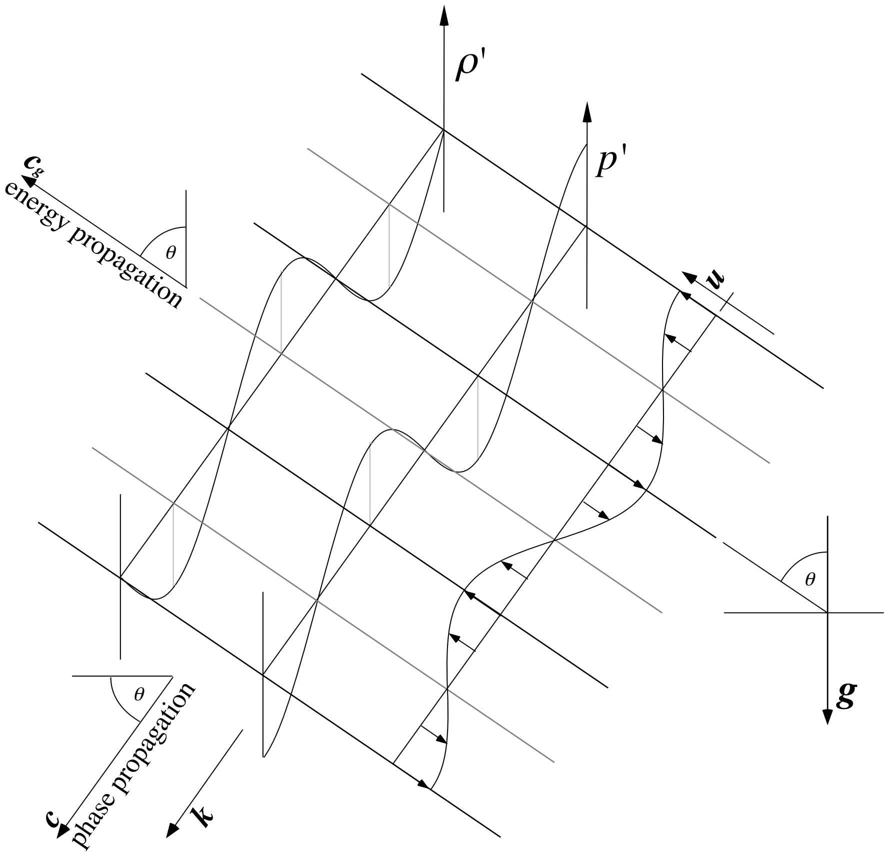


Fig. 2.1.2: The structure of a plane internal gravity wave.

Physically, the anisotropy of the dispersion relation (which implies that $\boldsymbol{c}$ and $\boldsymbol{c}_g$ are not parallel) can be understood as follows. A vertical buoyancy force acts on the density perturbations, directed upwards for negative perturbations and downwards for positive. For short waves, the pressure gradient (directed normal to the wavefronts) exactly cancels the component of the buoyancy force normal to the wavefronts. Thus the only restoring force is parallel to the wavefronts, resulting in a transverse wave. This restoring force is proportional to $\cos\theta$, giving the anisotropic dispersion relation (Lighthill, 1978).

In terms of the wavelength $\lambda=2\pi/|\boldsymbol{k}|$, the magnitudes of the phase and group velocities are

$$|\boldsymbol{c}| = \frac{\lambda\omega}{2\pi} = \frac{\lambda N\cos\theta}{2\pi} \tag{2.1.11}$$

and

$$|\boldsymbol{c}_g| = \frac{\lambda N\sin\theta}{2\pi} \tag{2.1.12}.$$

Thus both velocities are proportional to wavelength, but $|\boldsymbol{c}|$ increases with increasing frequency, whilst $|\boldsymbol{c}_g|$ decreases.

Wave energy generated by an infinitesimal simple-harmonic disturbance at a frequency less than N will travel away from the source region with the group velocity $\boldsymbol{c}_g$, that is, along cones whose angle to the vertical is θ. This results in the cross pattern of waves shown in Fig. 2.1.3, where the wave energy is confined to narrow beams extending away from the source (the small dotted rectangle corresponds to the portion of a plane wave shown in Fig. 2.1.2). Wavefronts extend *along* these beams and propagate across them towards the horizontal plane containing the source, as shown by the fine arrows in the figure. The wavefronts appear "from nothing" at the insides of the cones, and disappear "into nothing" at their outsides. This appearing and disappearing trick is possible because the wavefronts carry no energy in their propagation direction. Note that the vertical components of the group and phase velocities are opposite in sign - this allows one to tell the direction of the group velocity by observing that of the phase velocity. If the source is two-dimensional, the wave energy is confined to travel one-dimensionally and thus suffers no loss of intensity with distance (ignoring the effects of viscosity). The contrast between the behaviour of internal gravity waves and the spherical wavefronts seen in the familiar case of isotropic dispersion could hardly be more marked.

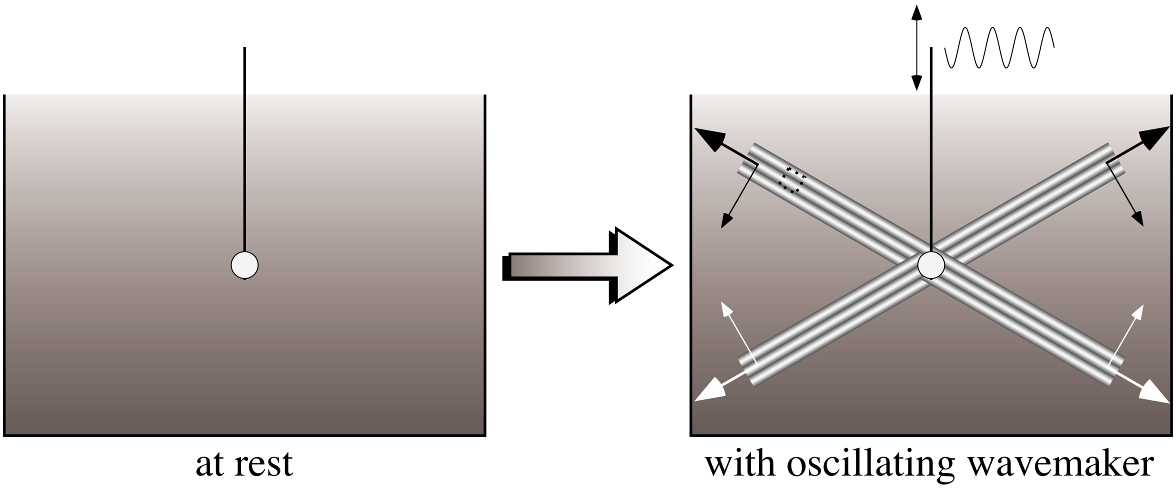


Fig. 2.1.3: Internal waves spreading from a small source.

In addition, if the periodic oscillation of the source is not simple-harmonic, wave beams will form for each Fourier component of the oscillation whose frequency lies below N. Each beam has a different angle θ given by the dispersion relation for the frequency of the Fourier component which produced it, and as a result the different frequencies in the wavefield will separate out as they spread from the source. Thus the wave motion at a sufficient distance from a small source will be simple-harmonic, no matter how complex the disturbance which produced it. Furthermore, a periodic source oscillation which is not simple-harmonic will *not* produce higher frequencies in the wavefield if the extra harmonics lie above the buoyancy frequency. Thus the (propagating) wavefield is band-limited by N, even if the source oscillation is not. This is so because source motions with $\omega > N$ produce evanescent waves which die off exponentially with distance from the source rather than forming free waves. These conclusions based on linear theory are strictly valid only for infinitesimal waves, and begin to break down as the wave amplitude becomes large (Samodurov, 1974; Kistovich *et al.,* 1990).

These predictions of linear theory were confirmed experimentally by Mowbray and Rarity (1967) using a modified schlieren technique to observe waves produced in a stratified salt solution. Although they were far from infinitesimal, the behaviour of the waves generated in my project was to a large degree dictated by the constraints predicted by linear theory.

2.2 Ray theory

We have seen in the previous section the properties of linear plane internal gravity waves in a medium with a constant buoyancy frequency N. Although plane waves are unrealistic in that they are infinite in both spatial and temporal extent, they still form a useful starting point in understanding the properties of more localised linear waves in inhomogeneous media.

To be more specific, a plane wave in some property ϕ (such as the perturbation density of an internal gravity wave) is of the form

$$\phi(\boldsymbol{r},t) = A\exp\left[i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)\right] \qquad (2.2.1),$$

where the amplitude A, wavevector $\boldsymbol{k}$ and frequency ω are constant. A more general wave motion is of the form

$$\phi(\boldsymbol{r},t) = A(\boldsymbol{r},t)\exp\left[iP(\boldsymbol{r},t)\right] \qquad (2.2.2),$$

with a variable amplitude $A(\boldsymbol{r},t)$ and a phase function $P(\boldsymbol{r},t)$. If the amplitude function is slowly varying in space and time compared with the phase function, the waves are “nearly plane”, and can be described by way of *ray theory.* This is identical to the approach of geometrical optics as an approximation to wave optics, and is also known as the WKB approximation.

By analogy with plane waves, we define the local frequency and wavevector of such a nearly-plane wave by

$$\omega = -\frac{\partial P}{\partial t} \qquad (2.2.3)$$

and

$$\boldsymbol{k} = \nabla P \qquad (2.2.4),$$

respectively. These immediately imply that

$$\frac{\partial \boldsymbol{k}}{\partial t} + \nabla\omega = \boldsymbol{0} \qquad (2.2.5).$$

This is often called the conservation of crests equation, as it implies that wavecrests can be neither created nor destroyed. The applicability of ray theory is closely linked to the validity of (2.2.5), which breaks down when $A(\boldsymbol{r},t)$ varies on similar time or distance scales to $P(\boldsymbol{r},t)$.

In a medium whose properties vary in space or time, the dispersion relation connecting ω and $\boldsymbol{k}$ depends on these variations. Thus the dispersion relation is of the form

$$\omega(\boldsymbol{r},t) = \omega\big(\boldsymbol{k}(\boldsymbol{r},t);\chi(\boldsymbol{r},t)\big) \qquad (2.2.6),$$

where the parameter $\chi(\boldsymbol{r},t)$ represents the influence of the properties of the medium, such as variations in the buoyancy frequency N.

We can define the local group velocity by

$$\boldsymbol{c}_g = \nabla_{\boldsymbol{k}}\omega \qquad (2.2.7).$$

This is analogous to the plane wave case, but now $\boldsymbol{c}_g$ is a function of $\boldsymbol{r}$ and t. The wave energy from a point propagates with a velocity $\boldsymbol{c}_g$ which varies with position and time, along a curved path known as a *ray*. Ray theory treats a wave like a particle (a wave packet) moving with velocity $\boldsymbol{c}_g$ and with properties ω and $\boldsymbol{k}$ which can vary as the "particle" propagates. In a frame moving with the ray, these variations are given by (LeBlond & Mysak, 1978)

$$\frac{d\boldsymbol{k}}{dt} = -\frac{\partial\omega}{\partial\chi}\nabla\chi \qquad (2.2.8)$$

and

$$\frac{d\omega}{dt} = \frac{\partial\omega}{\partial\chi}\frac{\partial\chi}{\partial t} \qquad (2.2.9),$$

which follow from the conservation of crests equation (2.2.5).

If the medium has no variation with respect to time, it is clear from (2.2.9) that the frequency remains constant along a ray. The short-wave dispersion relation (2.1.7) remains applicable when the buoyancy frequency N is constant in time and varies on length scales much greater than a wavelength (Lighthill, 1978). Since ω is constant for a given ray, this results in a relationship between the local propagation angle θ and the local value of N:

$$\cos\big(\theta(\boldsymbol{r})\big) = \frac{\omega}{N(\boldsymbol{r})} \qquad (2.2.10).$$

Thus internal gravity waves are *refracted* by variations in N, that is, by deviations of the density profile from an exponential increase with depth. As shown in Appendix 1.3, N can only be a function of the vertical coordinate z for a fluid in hydrostatic equilibrium, so if the wavefield involves negligible density variations, the refraction will be a function of depth alone.

This refraction serves to confine waves within a depth range in which their frequency is exceeded by the buoyancy frequency. This is illustrated in Fig. 2.2.1.

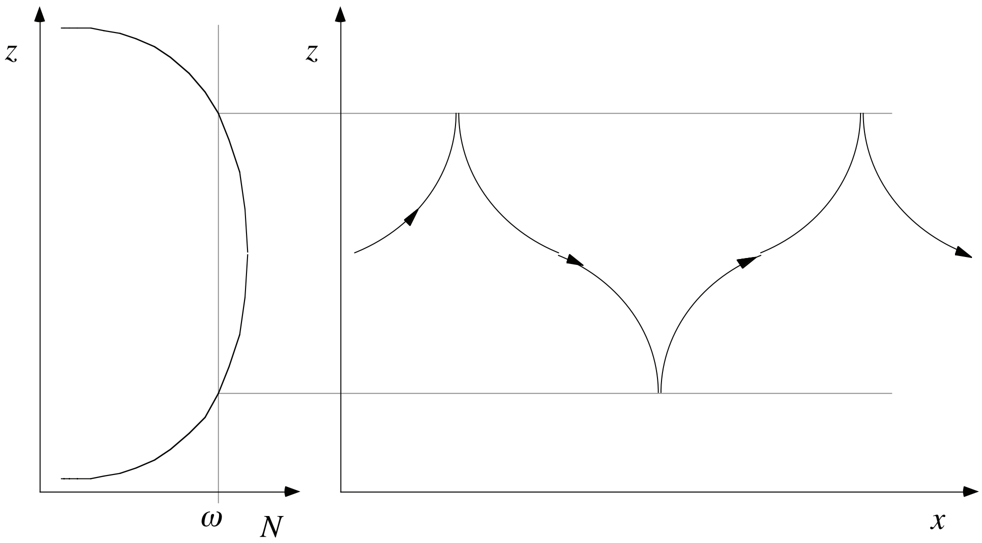


Fig. 2.2.1: Wave refraction and trapping due to a variation in N.

The trajectory of a ray moving into a region of decreasing N becomes more steeply inclined to the horizontal, and becomes vertical if it reaches a depth at which $N=\omega$. At this point the ray's group velocity vanishes and it cannot proceed into the region in which $N<\omega$ (a progressive wave solution is not possible in this region). The ray path forms a cusp as it doubles back on itself and propagates back through the region in which $N>\omega$. If the ray is bounded above and below by depths at which $N<\omega$ (as in the figure), it is confined to propagate in the intervening layer.

Such a situation exists in the oceans, where there is a layer close to the surface called the *thermocline* through which the density increases rapidly with depth due to a steep temperature gradient. Below the thermocline, the density varies much more gradually (resulting in a lower N), whilst the layer between the thermocline and the surface is well mixed by surface wave action and other instabilities, and so has essentially no gradient and zero N. Thus the higher-frequency oceanic internal waves are confined to propagate essentially two-dimensionally in this thermocline region (Lighthill, 1978).

Ray theory greatly simplifies calculations of complicated wave motions. With a knowledge of $\chi(\boldsymbol{r},t)$ the dispersion relation (2.2.6) can be derived, allowing the local group velocity to be found at any $(\boldsymbol{r},t)$ for a given $\boldsymbol{k}$. Starting from a point at which the values of ω and $\boldsymbol{k}$ are known, the ray propagation direction can be found from (2.2.7). Taking a small step in this direction, new values of ω and $\boldsymbol{k}$ can be found from (2.2.8) and (2.2.9), allowing the new value of $\boldsymbol{c}_g$ to be found. Ray paths can be found in this way by integrating $\boldsymbol{c}_g$ through the fluid.

Ray theory becomes invalid when the conservation of crests equation (2.2.5) breaks down, which occurs at places known as *caustics* where neighbouring rays cross. Ray theory also begins to lose its applicability when variations in the medium become large on time and length scales comparable to the period and wavelength of the rays, respectively.

The above results can be generalised for ray propagation in a moving medium. Let $\boldsymbol{U}(\boldsymbol{r},t)$ represent an ambient current which varies appreciably on length and time scales much greater than the period and wavelength, respectively, of the rays. Then a stationary observer will measure a *Doppler-shifted* frequency ω' given by

$$\omega' = \omega + \boldsymbol{k} \cdot \boldsymbol{U} \qquad (2.2.11),$$

and rays will propagate with velocity $\boldsymbol{c}_g+\boldsymbol{U}$, as they are advected by the current. The frequency ω defined earlier is that measured in a frame moving with velocity $\boldsymbol{U}$, and is often called the *intrinsic frequency*. The wavevector $\boldsymbol{k}$ is unaffected by the transformation between frames, and ω still satisfies the dispersion relation (2.2.6).

In a moving medium, (2.2.8) and (2.2.9) are replaced by (Lighthill, 1978)

$$\frac{d\boldsymbol{k}}{dt} = -\frac{\partial \omega}{\partial \chi} \nabla \chi - (\boldsymbol{k} \cdot \nabla) \boldsymbol{U} \qquad (2.2.12)$$

and

$$\frac{d\omega'}{dt} = \frac{\partial \omega}{\partial \chi} \frac{\partial \chi}{\partial t} + \boldsymbol{k} \cdot \left(\frac{\partial \boldsymbol{U}}{\partial t} \right) \qquad (2.2.13),$$

where (as before) the derivatives are taken in a frame moving with the advected rays. From (2.2.12) we see that the rays are refracted by the *shear* due to the background current. Equation (2.2.13) shows that the frequency observed by a stationary observer can also be changed by fluctuations in the current, as we would expect from (2.2.11).

2.3 *The effects of viscosity*

All the previous discussion has neglected the viscosity of the fluid medium. Thomas & Stevenson (1972) considered the case of two-dimensional linear waves in a *viscous* density-stratified incompressible fluid with constant N. They derived a similarity solution for the internal waves generated by a two-dimensional simple-harmonic localised disturbance of negligible dimensions.

To describe the wavefield of a downgoing internal wave beam produced by a two-dimensional simple-harmonic disturbance at the origin, it is helpful to use rotated coordinate axes x' and z' parallel to the group and phase velocity of the waves, respectively. Thus x' is directed along the beam and z' across it as shown in Fig. 2.3.1. Then in terms of η (a dimensionless measure of distance across the beam), Thomas & Stevenson (1972) derive displacement profiles of the form shown in Fig. 2.3.2 for several values of t.

The locus of points of constant η is given by

$$z' = \eta \left(\frac{x' \mu^*}{2\rho^* N \sin\theta} \right)^{\frac{1}{3}} \qquad (2.3.1),$$

where ρ^* and μ^* are the density and viscosity, respectively, at the depth of the wave source. This is curve is shown in Fig. 2.3.1.

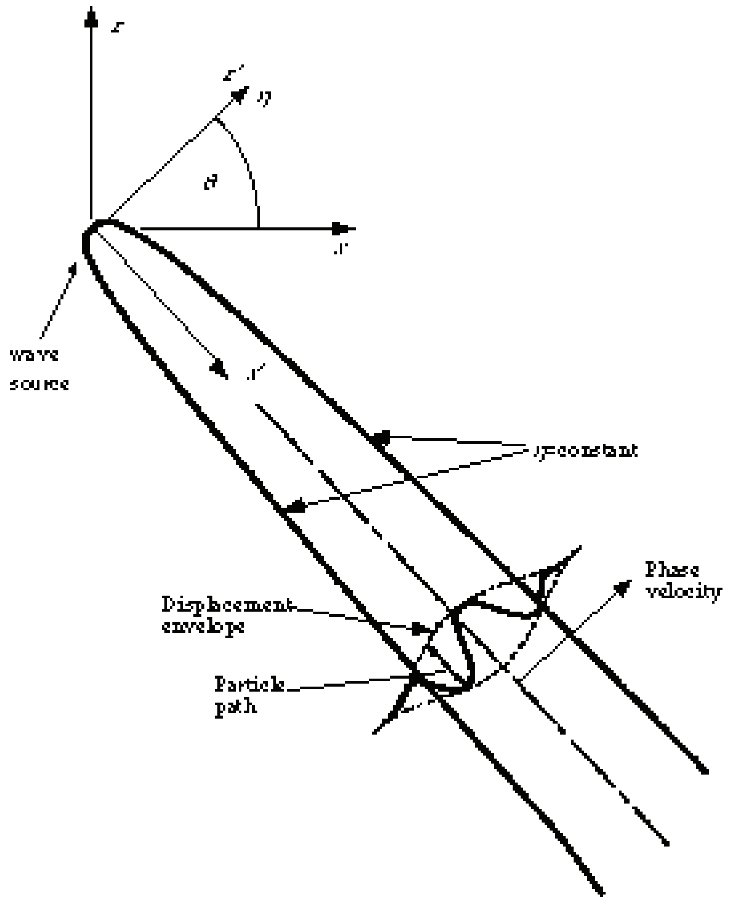


Fig. 2.3.1: Definition of coordinates aligned with the downgoing wave beam (adapted from Thomas & Stevenson, 1972).

The displacement profiles shown in Fig. 2.3.2 as a function of η will thus become broader as x' increases, since the value of z' for a given value of η increases with x', as given by equation (2.3.1) and shown in Fig. 2.3.1. This similarity solution imposes a minimum wavelength for the waves produced by a small source at a given distance x' from the source. A plot of z' against x' for various values of η is shown in Fig. 2.3.3 for typical parameters used in my experiments. By comparison with Fig. 2.3.2 we see that the longest wavelength expected from such a small source in a tank 60 cm long is about 5 cm. Thomas & Stevenson (1972) conducted experiments with a stratified salt solution which verified these predictions of their theory.

Kistovich *et al.* (1990) and Ivanov (1989) experimentally investigated wave generation by a source whose dimensions were not negligible. The sources used were vertically oscillated cylinders of various sizes. They found that the form of the beam generated depended on the ratio of the cylinder diameter to the viscous wave scale L_v given by

$$L_v = \frac{1}{N}\left(\frac{g\mu^*}{\rho^*}\right)^{\frac{1}{3}} \qquad (2.3.2).$$

If the diameter was less than L_v, the wave produced was a unimodal beam with a maximum at the centre, as predicted by Thomas & Stevenson. Cylinders whose diameters exceeded this scale produced bimodal beams, with two maxima corresponding to the top and bottom of the cylinder. For the stratifications used in my experiments, L_v is about 1.4 cm.

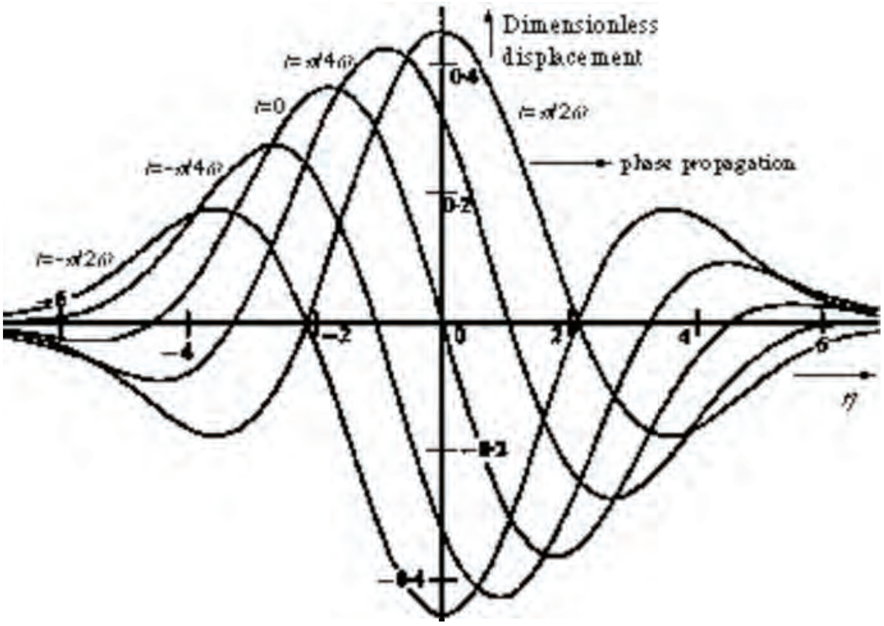


Fig. 2.3.2: Similarity displacement profiles at different times (adapted from Thomas & Stevenson, 1972).

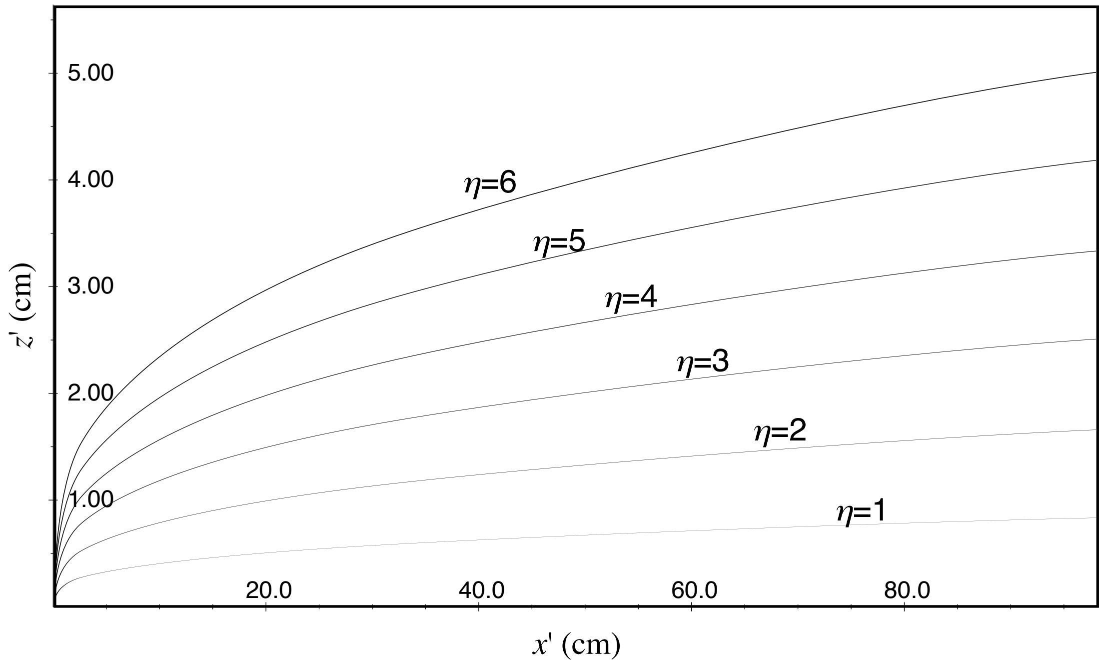


Fig. 2.3.3: Curves of constant η for my experiments.

2.4 Weak nonlinear interactions

Up to this point, we have considered only waves of infinitesimal amplitude, which are described by linear theory. Such waves will pass through one another without any change in their properties, and a linear combination of infinitesimal waves remains a solution of the linearised equations of motion. Under such conditions, there is no transfer of energy between different frequencies.

However, it is observed that the oceanic internal wavefield has a spectrum (see Fig. 1.2.1, for example) whose shape is essentially the same throughout the oceans (Garrett & Munk, 1972, 1975). Furthermore, distorted spectra are seen to return rapidly to this "universal" form. These intriguing empirical results have been the subject of intense theoretical investigation for several decades. An explanation consists of three parts: an understanding of energy sources, energy sinks and mechanisms by which energy may be redistributed in the spectrum. This redistribution must rely on some type of *nonlinear* interaction between waves in different parts of the spectrum, as linear theory prohibits such energy exchange.

In the process of linearisation, the nonlinear governing equations (A.1.12) and (A.1.13) become (A.1.14) and (A.1.15) by neglecting the terms

$$\boldsymbol{u}\cdot\nabla\rho' \,,\; \rho(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} \qquad (2.4.1\ a,b)$$

and

$$\rho'\frac{\partial \boldsymbol{u}}{\partial t} \qquad (2.4.2)$$

which involve products of the perturbation variables p', ρ' and $\boldsymbol{u}$ (see Appendix 1.3). As is clear from Fig. 2.1.2, the first two neglected terms are identically zero for a short plane wave, as $\boldsymbol{u}$ is perpendicular to gradients in ρ' and $\boldsymbol{u}$. Thus a short finite-amplitude plane wave is a valid solution for the *nonlinear* equations of a fluid with constant N, provided (2.4.2) remains negligibly small.

In general, a linear combination of such plane waves will *not* be a solution of the nonlinear equations, as the terms (2.4.1) may then be nonzero. However as Phillips (1960) pointed out (in the context of surface waves but equally valid here), if we have two plane waves of frequencies ω_1, ω_2 and wavevectors $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ whose amplitudes are not too large, the nonlinearities are essentially quadratic. These quadratic terms result in secondary temporal and spatial variations characterised by frequencies

$$\omega_3 = \omega_1 \pm \omega_2 \qquad (2.4.3)$$

and wavevectors

$$\boldsymbol{k}_3 = \boldsymbol{k}_1 \pm \boldsymbol{k}_2 \qquad (2.4.4),$$

which can be considered as small forcing terms to the linear equations.

In general ω_3, $\boldsymbol{k}_3$ will not satisfy the dispersion relation (so the phase velocity of these secondary components will not match that of a free wave with the same wavevector) and the amplitude of the secondary component will remain small.

However, if ω_3, $\boldsymbol{k}_3$ *do* satisfy the dispersion relation, the forced secondary component can then propagate as a free wave in its own right. In this case, the two initial waves together with the secondary wave form a *resonant triad*, and energy can be exchanged between the three waves. From (2.4.3) and (2.4.4) it is clear that in this case the three waves are on an equal footing (they are all free to propagate), so the distinction between primary and secondary waves loses its significance (Turner, 1973). The existence of resonant triads has been confirmed experimentally by McEwan (1971) and McEwan *et al.* (1972), amongst others.

These weakly nonlinear interactions have been incorporated into many *statistical* theories which attempt to explain the structure and evolution of the oceanic internal wavefield as a result of the formation of resonant triads amongst waves in different parts of the frequency and wavenumber spectra. Statistical theories are used because the wavefield is treated as being comprised of waves of all possible frequencies and wavenumbers, each with a spectral density which evolves with time. Such theories are reminiscent of those of quantum field theory and have actually been formulated in terms of Feynman diagrams of possible interactions, with propagating waves treated as particles and wave motions which fail to satisfy the dispersion relation described as virtual particles (Hasselmann, 1966). Such schemes can also incorporate the interaction of the internal wavefield with surface and seismic waves and with bottom topography.

2.5 Strong nonlinear interactions

Although a great deal of work has been done on describing the oceanic internal wavefield in terms of weak interactions (see Müller *et al.,* 1986), oceanic internal waves are often about 100 times too strong for weak theory to be applicable (Holloway, 1980, 1982). Thus it would seem that theories which incorporate *strong* nonlinear interactions will be needed to fully account for the observed oceanic wave spectra. The development of fully nonlinear *statistical* wave theories is a daunting task, as strong interactions allow much more indiscriminate energy transport between waves than interactions based on resonant triads. Thus several researchers have “started small” by investigating the strong nonlinear interactions of only two waves, in order to understand some of the basic processes which can occur when weakly nonlinear interaction theory breaks down.

Some of the most important interactions of the oceanic internal wavefield are between waves which are widely separated in terms of wavenumber (Müller *et al.,* 1986). Consequently, if the long wave is much stronger than the short wave, and of a much lower frequency, the effect of the long wave on the propagation of the short wave can be calculated by using ray theory. The weak short wave can be considered to be a “test” ray which does not modify the background long-wave flow. Using ray theory, the time-varying background flow due to the long wave can be arbitrarily strong, and the refraction and frequency shift of the short wave ray as it travels through this field can be found from equations (2.2.12) and (2.2.13). Note that the short wave frequency along a ray can change, since the background flow is time-dependent.

Consider a purely two-dimensional interaction in which no properties of the wave field depend on y. Thus both the background and test waves have group and phase velocities parallel to the plane y=0. It is helpful to define rotated coordinates (x', z') parallel to the group and phase velocity, respectively, of the background wave, which we assume to be downgoing (so that the vertical component of its group velocity is negative). We will denote quantities relating to the background wave by capital letters and reserve lowercase characters for the weak test wave; dashed quantities are measured with respect to the rotated (x', z') axes. This is illustrated in Fig. 2.5.1 - note that this rotated coordinate system is identical to that defined in section 2.3.

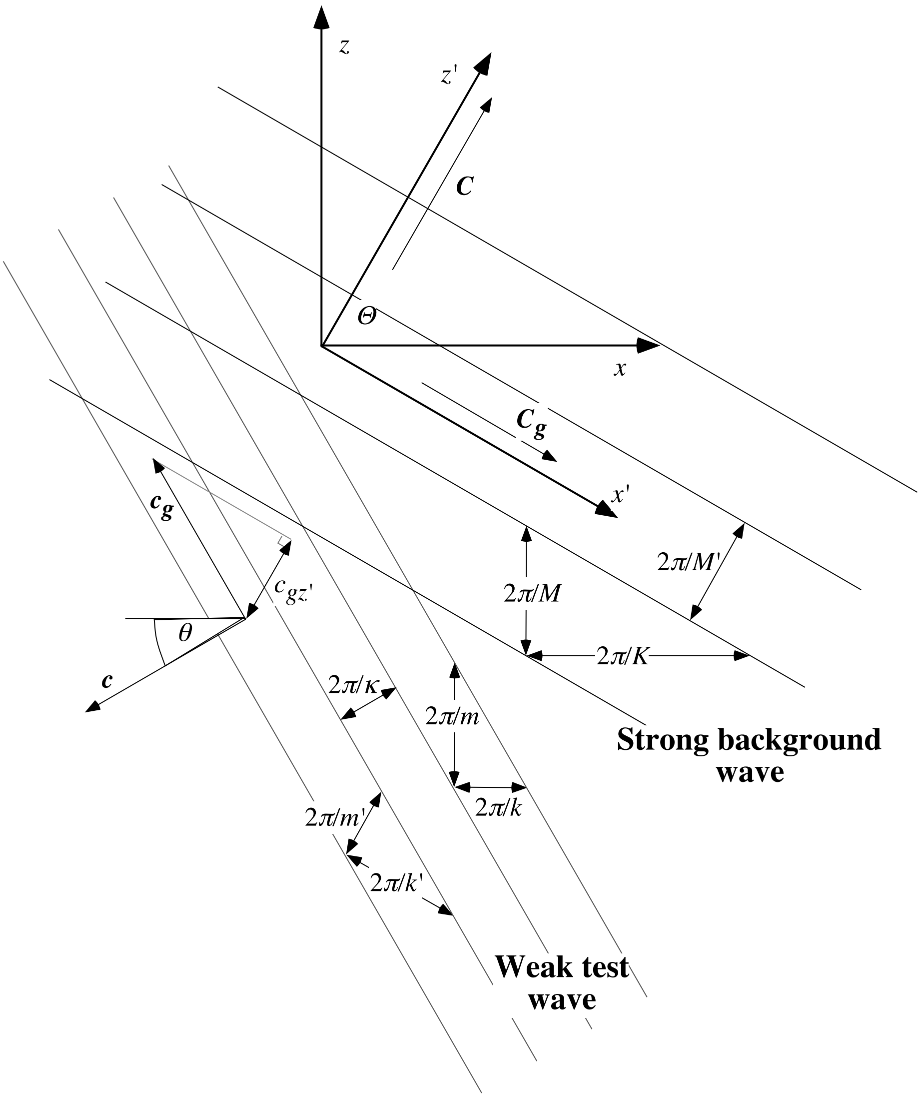


Fig. 2.5.1: Coordinates aligned with the background wavefield.

The background wave is considered to be infinite in extent in the x' direction, so the background field is a function of z' and t alone. It is clear from equation (2.2.12) that the component k' of the test wavevector in the x' direction is conserved. Thus the ray equations give the variation in m' of the test ray as a function of z' and t. Note also that only the component $c_{gz'} \equiv |\boldsymbol{C}|^{-1}(\boldsymbol{c}_g \cdot \boldsymbol{C})$ of the test wave group velocity in the z' direction is important, as the ray experiences no change in the background wavefield by moving in the x' direction.

The background field of the long wave affects the test ray through both the velocity and density gradient variations it produces. As is clear from equation (2.2.12), the ray is refracted through the action of the shear field produced by the z' dependence of the velocity field, and also refracted by the local modifications to N

arising from the density variations. The ray equations allow calculations of the test wave refraction to be made relatively easily (in some cases even analytic results can be found), yet places no restriction on the background wave amplitude. This approach was followed by Broutman and co-workers, and by Thorpe (1989), amongst many others.

Broutman and his collaborators looked at the effect of the shear field alone, by choosing a background which involved only horizontal fluid motions and hence no density variations (Broutman, 1984, 1986; Broutman & Grimshaw, 1988; Broutman & McIntyre, 1994; Broutman & Young, 1986; Macaskill & Broutman, 1988). As this work provided the theoretical framework for my experiments, I will discuss it in some detail. The background chosen was an *inertial wave*, a limiting case of a *gyroscopic wave*, which is an internal wave driven by the Coriolis forces arising in a rotating reference frame such as the Earth[1]. The phase velocity of such waves is vertical, so the (x', z') coordinates are identical to the unrotated (x, z) coordinates in this case. Apart from simplifying the analysis, an inertial wave was chosen because these waves are amongst the strongest in the oceanic internal wavefield (Broutman & Young, 1986).

The inertial background wave was chosen to have phase velocity pointing upwards. The wave had infinite horizontal extent but was confined vertically by a Gaussian envelope - the vanishing group velocity implies that this envelope is stationary. A test internal gravity wave was propagated upwards from below the inertial wave envelope, and its path traced by numerically integrating the ray equations.

It was found (Broutman & Young, 1986) that no matter how weak the background field, there would always exist a range of test wavenumbers m' in which test rays would meet caustic surfaces (surfaces on which neighbouring rays cross) at certain phases in the background field. Broutman (1986) showed that the unphysical singularities given by ray theory at these caustics can be removed by standard Airy integral methods.

The top figure in Fig. 2.5.2 shows the position of a test wave packet as a function of time as it propagates through the shear field of the background wave (the coordinate z' is scaled by the characteristic width L of the Gaussian envelope enclosing the background wave). It is evident from this figure that the test ray group velocity changes abruptly when it encounters a caustic surface. These surfaces enclose the local velocity maxima of the background wave and thus move upwards with the inertial wave phase velocity $\boldsymbol{C}$. The caustic surfaces close off at large $|z'|$ due to the Gaussian envelope.

---

[1] Gyroscopic waves have a similar anisotropic dispersion relation to internal gravity waves, but the maximum frequency is the *inertial frequency* f rather than the buoyancy frequency N. Inertial waves are gyroscopic waves at the inertial frequency and consist of horizontal, circularly polarised currents with a vertical wavevector and phase velocity (pointing either up or down) and zero group velocity. Further details on inertial waves can be found in many oceanographic texts, such as LeBlond & Mysak (1978) or Apel (1987).

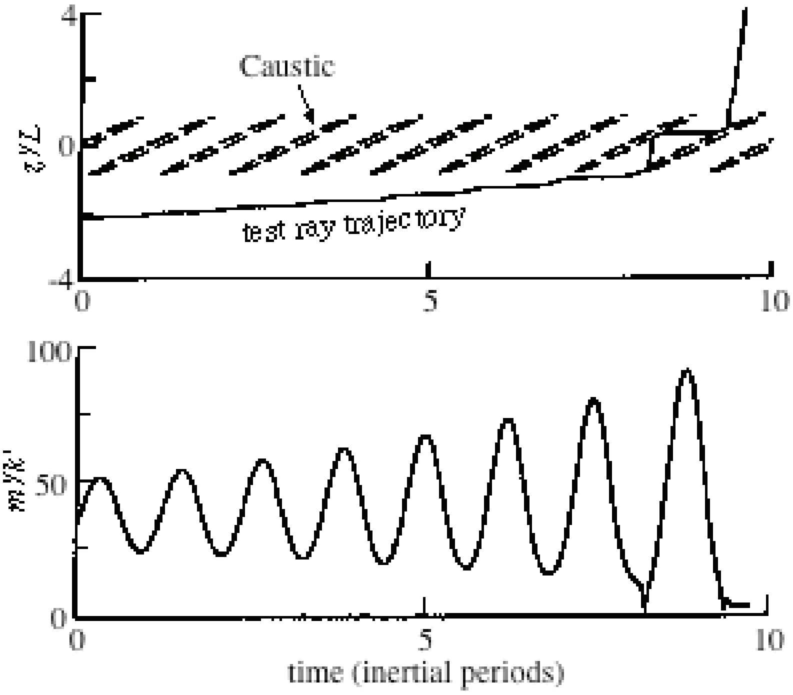


Fig. 2.5.2: Test ray refraction by caustics (from Broutman & McIntyre, 1994).

A test wave packet can encounter a caustic when the component $c_{gz'}$ of its group velocity parallel to ***C*** differs from C=|***C***|, so that the ray is in motion relative to the background wavefronts. When a ray arrives at a caustic (either by catching up with it or being overtaken), its value of $c_{gz'}$ is altered so that it cannot penetrate the caustic surface (it becomes less than C if it was initially greater, or greater if it was initially less). The caustics themselves are defined as the surfaces on which $c_{gz'}=C$, and the change in $c_{gz'}$ is greatest at these surfaces because the test wave packet experiences a sustained shear by remaining at the same phase of the background wave for an appreciable time. As shown in the lower figure in Fig. 2.5.2, there are accompanying abrupt changes in the wavenumber m' of the test wave at each encounter with a caustic, and also more gradual changes due to the shear field.

The propagation of test rays through the background wave can be more fully understood in terms of Fig. 2.5.3. Broutman & Young (1986) showed that the quantity ω'' defined by

$$\omega''=\omega'-Cm' \tag{2.5.1}$$

is conserved when the Gaussian envelope of the background wave is infinitely broad (ie. $L\rightarrow\infty$). Here ω' is the Doppler-shifted frequency defined in (2.2.11), where the velocity field is that due to the background wave. ω'' is the frequency of the test waves as measured by an observer moving with velocity ***C***. When L is finite, numerical experiments show that ω'' is slowly varying on the scale of the inertial wave envelope, and to a first approximation a wave packet may be considered to retain its initial value of ω'' as it passes through the background wave (Broutman & Young, 1986). Thus the wave packet is constrained to remain on a contour of constant ω'' such as those shown in Fig. 2.5.3. The vertical axis in these figures is the phase $\xi=M'(z'-Ct)/2\pi$ of the background wave. The small figures to the right show the variation in the background wave velocity field as a function of ξ.

Broutman & Young (1986) show that a given value of ω'' can be achieved by two distinct values of m' in the fluid at rest (a large distance away from the background wave), but a test ray cannot change from one of these roots to the other, since the intervening values of m' are forbidden (they correspond to different values of ω''). Near the centre of the background wave envelope, the ω'' contours become strongly distorted, resulting in a region of closed contours where these two roots are connected. This region of closed contours enables the test ray to have a large permanent change in m' by switching over to the other root, while still conserving ω''.

The process is illustrated in the succession of diagrams in Fig. 2.5.3. In Fig. 2.5.3a, the test wave approaches the inertial wavefield with $\omega''=\omega_t''$ and $c_{gz'}<<C$, and thus moves backwards relative to the wavefronts, as shown by the arrow. Note the second contour of $\omega''=\omega_t''$ is separate from the incident contour in this figure. As the test wave packet travels closer to the centre of the inertial wave envelope the contours become more distorted, until the two contours for $\omega''=\omega_t''$ join into a closed curve as shown in Fig. 2.5.3b. The test wave packet orbits this contour as shown by the arrows, alternately having $c_{gz'}<C$ and $c_{gz'}>C$. The test wave is stationary with respect to the background wave phase when the tangent to this contour is horizontal - this corresponds to an encounter with a caustic surface as shown in Fig. 2.5.2. As the background wave strength increases still further the closed contour of $\omega''=\omega_t''$ becomes confined so that the test wave cannot pass through all phases of the background wave (see Fig. 2.5.3c) - this corresponds to the finite width of the caustic surfaces in Fig. 2.5.2, which exclude the test wave from the vicinity of the velocity maxima of the background wave. The background wave amplitude experienced by the test wave decreases once the packet passes the maximum of the Gaussian envelope, and eventually the contour $\omega''=\omega_t''$ breaks into two separate curves once more.

The test wave packet can finish up on either of the disconnected curves after its encounter with the background wave, depending on the number of times it encountered a caustic. This is also clear from Fig. 2.5.2: $c_{gz'}$ "toggles" between two distinct values each time the test wave reaches a caustic, so when the test wave escapes it essentially retains the value of $c_{gz'}$ it had from its last encounter with a caustic. As is evident from Fig. 2.5.2, the test wave is most likely to escape from the background wave when $c_{gz'}>>C$, which from Fig. 2.5.3 corresponds to a value of $|m'|$ much less than the value $|m_c'|$ at the caustic (Broutman & Young, 1986). Thus the probable change in the test wave is most drastic when the incident test wave has $c_{gz'}<<C$ and $|m'|>>|m_c'|$ - this is a so-called "third-kind" encounter (Broutman & McIntyre, 1994). The decrease in $|m'|$ in such an encounter corresponds (by the dispersion relation (2.1.7)) to an increase in frequency, since k' is conserved. The decrease in $|m'|$ also corresponds to an *increase* in the wavelength of the short wave, contrary to the transport of energy from large to small scales predicted by statistical weak interaction theories and expected from studies of the oceanic internal wavefield. This remarkable result shows the difference it makes to consider fully nonlinear interactions, and it was the aim of my project to observe this effect experimentally.

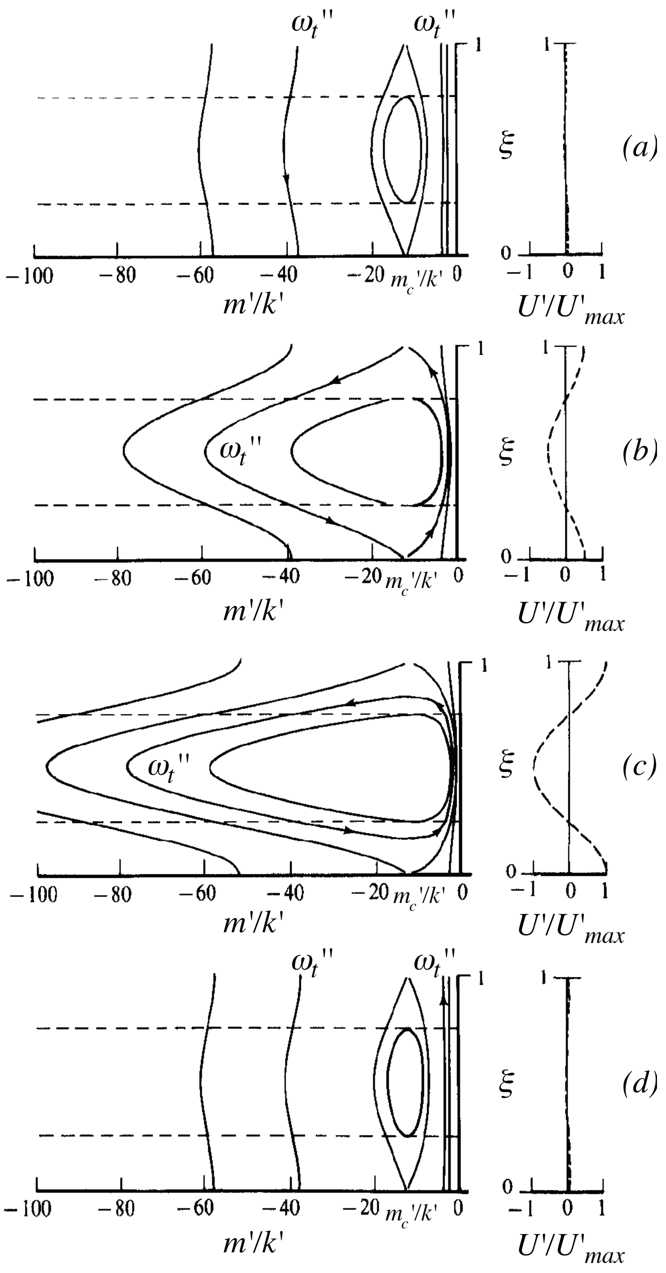


Fig. 2.5.3: Contours of constant ω'' vs. background phase ξ (adapted from Broutman & Young, 1986).

As is clear from Fig. 2.5.3, only test waves which have an initial value of m' sufficiently close to m_c' will find themselves on a closed ω'' contour at the maximum of the background wave envelope and thus be able to have a change in m'. The extent of the region of closed contours at the background wave maximum gives the range of incident wavenumbers m' which can be refracted - as shown in Fig. 2.5.3, this region is larger for greater peak background wave velocities. Furthermore, the larger regions of closed contours can connect more widely separated values of m' and so produce more drastic refraction. Thus refraction becomes both larger and more likely as the background wave is made stronger.

2.6 Adaptation of Broutman & Young's theory

Section 2.5 gave an overview of the main points of the theory developed by Broutman and collaborators to describe the effect of a strong inertial wave on the trajectory of a "test" internal gravity wave. This section describes an adaptation of this theory which replaces the inertial wave by a strong internal gravity wave, as was done in my experiments.

As was shown in section 2.1 (equations (2.1.11) and (2.1.12)), the magnitudes of the phase velocity $\boldsymbol{C}$ of the background wave and group velocity $\boldsymbol{c}_g$ of the test ray depend on their propagation angle and hence on their frequency. Thus the ratio $G \equiv c_{gz'}/C$ depends on the waves' *frequencies* as well as their wavelengths and the angle between their wavefronts. In section 2.5, the inertial wave frequency was fixed, and only the dependence on the test wave frequency needed to be considered. When both waves are internal gravity waves (as in my case), both frequency dependencies need to be considered. This is the subject of this section - the theory is generalised to cases in which the angle Θ is not fixed at 90° (see Fig. 2.5.1).

From (2.1.11), the phase speed of the background wave is

$$C = \Omega / M' \tag{2.6.1}$$

where Ω is the frequency of the background wave, and M' the magnitude of its wavevector. From (2.1.12) and the dispersion relation (2.1.7), the magnitude of the group velocity of the incident test wave is

$$c_g = \frac{N}{\kappa}\sqrt{1-\left(\frac{\omega}{N}\right)^2} \tag{2.6.2}$$

where κ is the magnitude of its wavevector and ω its frequency. Referring back to Fig. 2.5.1, it is clear that $c_{gz'}$ is given by

$$\begin{aligned} c_{gz'} &= c_g \sin(\Theta - \theta) \\ &= \frac{N}{\kappa}\sqrt{1-\left(\frac{\omega}{N}\right)^2}\,\sin\left[\cos^{-1}\left(\frac{\Omega}{N}\right) - \cos^{-1}\left(\frac{\omega}{N}\right)\right] \end{aligned} \tag{2.6.3}$$

Thus we have

$$G \equiv \frac{c_{gz'}}{C} = \frac{M'}{\kappa}\frac{N}{\Omega}\sqrt{1-\left(\frac{\omega}{N}\right)^2}\,\sin\left[\cos^{-1}\left(\frac{\Omega}{N}\right) - \cos^{-1}\left(\frac{\omega}{N}\right)\right] \tag{2.6.4}$$

which for a given M'/κ is a function of only Ω/N and ω/N. A plot of G is shown in Fig. 2.6.1, and will be discussed shortly.

The adaptation of predicted wavenumber variations to the case of internal gravity waves is somewhat more complicated, since we no longer have the simplification that $\boldsymbol{g}$ is parallel to the z' axis. Thus the dispersion relations are most simply expressed in terms of the unrotated coordinates, but the rotated coordinates are best for describing the geometry of the wave interactions.

Referring to Fig. 2.5.1, we can derive the following relations between the components of the test wavevector in these two coordinate systems:

$$k = -\kappa\cos\theta = k'\sin\Theta + m'\cos\Theta \qquad (2.6.5)$$

$$m = -\kappa\sin\theta = -k'\cos\Theta + m'\sin\Theta \qquad (2.6.6)$$

$$k' = -\kappa\sin(\Theta-\theta) = k\sin\Theta - m\cos\Theta \qquad (2.6.7)$$

$$m' = -\kappa\cos(\Theta-\theta) = k\cos\Theta + m\sin\Theta \qquad (2.6.8).$$

I use the sign convention that frequencies are positive, but the components of wavevectors may have either sign. Note from Fig. 2.5.1 that k, m, k' and m' are negative, whilst κ and M' are positive. As in section 2.5, uppercase symbols belong to the background wave, whilst lowercase is reserved for the test wave.

As in section 2.5, we define the frequency ω'' by

$$\omega''=\omega-U'k'-Cm' \qquad (2.6.9).$$

This is the test wave frequency as measured by an observer moving with velocity $\boldsymbol{C}$ - it suffers Doppler shifts due to the background wave velocity field, as well as the motion of the observer. The conservation of ω'' in a plane background wave follows from the proof in Broutman & Young (1986).

From the dispersion relation (2.1.7), the test wave frequency is given by

$$\omega = \frac{-kN}{\sqrt{k'^2+m'^2}} \approx \frac{kN}{m'} \qquad (2.6.10),$$

where the latter approximation applies when $|m'|>>|k'|$, corresponding to a small angle between the test wave and the background wave. This approximation is applicable to the incident test waves in my experiments, but may not be valid for the refracted ray if the refraction is strong (this prevented me from deriving values of m' and ω for a refracted ray). In this approximation, we have by (2.6.5) and the dispersion relation (2.1.7) that

$$\omega = \frac{Nk'}{m'}\sin\Theta + \Omega \qquad (2.6.11).$$

The conservation of ω'' implies, from (2.6.9) and (2.6.11) that

$$m' = \frac{1}{2C}\left[-(\omega''-\Omega+U'k') \pm \sqrt{(\omega''-\Omega+U'k')^2 + 4CNk'\sin\Theta}\right] \qquad (2.6.12).$$

This gives two different m' values for a given ω'' when U'=0. These m' values coalesce at a caustic, where from (2.6.12), (2.6.9) and (2.6.7) the value m_c' of m' at a caustic is given by

$$m_c' = -\sqrt{\frac{N}{\Omega}M'\kappa\sin(\Theta-\theta)\sin\theta} \qquad (2.6.13).$$

For a third-kind encounter, we need $|m_i'|>>|m_c'|$, where m_i' is the incident wavenumber of the test wave. Thus we need to minimise

$$\frac{m_c'}{m_i'} = \frac{1}{\cos(\Theta-\theta)}\sqrt{\frac{NM'}{\Omega\kappa}\sin(\Theta-\theta)\sin\theta} \qquad (2.6.14).$$

Figs. 2.6.1 and 2.6.2 show non-dimensional plots of G (from (2.6.4)) and of m_c'/m_i' (from (2.6.14)), respectively. Using the dispersion relation, these are shown as functions of the incident ω/N, with each curve corresponding to a different value of Ω/N. In both plots, curves intersect the horizontal axis when $\omega/N=\Omega/N$, allowing Ω/N to be determined for each curve. We need small values of both G and m_c'/m_i' for a third-kind encounter, and these plots show that this requires either $\omega\approx\Omega$ or $\kappa>>M'$ for the incident wave, as it is impractical to use $\omega/N\approx1$ since the refracted ray will have an even higher frequency. Thus to have a third-kind encounter in which the test wave frequency is significantly higher than that of the background wave, we also need the test wave to be much shorter than the background wave.

Equation (2.6.12) also allows us to find the value U_c' of U' at a caustic. The square root in (2.6.12) is zero in this case, giving

$$U_c' = \frac{-N\sin\Theta}{m_i'} + \frac{Cm_i'}{k'} + \sqrt{\frac{-4CN\sin\Theta}{k'}} \qquad (2.6.15),$$

where from (2.6.9) I have used the expression

$$\omega''=\omega-Cm_i' \qquad (2.6.16),$$

where $m'=m_i'$ when $U'=0$, far from the background wave. Equation (2.6.15) gives the minimum peak flow speed in the background wave required to create a caustic (and hence strong refraction) for a given m_i'. From (2.6.7), (2.6.8) and (2.6.1), (2.6.15) may be written in the non-dimensional form

$$\frac{\kappa}{N}U_c' = \frac{\sin\Theta}{\cos(\Theta-\theta)} + \frac{\Omega\kappa}{NM'\tan(\Theta-\theta)} + \sqrt{\frac{4\Omega\kappa\sin\Theta}{NM'\sin(\Theta-\theta)}} \qquad (2.6.17),$$

which (by the dispersion relation (2.1.7)) depends only on the dimensionless parameters ω/N, Ω/N and κ/M'. The implications of this equation will be discussed in Chapter 6, using the experimental value of κ/M'.

The analysis of this section is only a partial adaptation of the theory, whose implications need to be treated with considerable caution. It was intended as a guide for devising experiments rather than a rigorous theory to be tested. The changes introduced by replacing the inertial wave by an internal gravity wave are more far-reaching than was considered here, since an internal gravity wave produces a variation in the density gradient as well as a shear field, and this results in a time- and space-dependent local buoyancy frequency which also refracts the rays. This was entirely ignored in this section, but Thorpe (1989) asserts that refraction due to the shear field is *negligible* compared to that due to the density variations under the conditions encountered in my experiments.[2]

Another *caveat* is that the ray-theory approaches of both Thorpe (1989) and Broutman and collaborators may not be applicable to my experimental situation, as

---

[2] Thorpe (1989) produced a formulation of refraction by the background wave's density variations which used ray theory for the test wave and ignored the shear field of the strong wave. He calculated some numerical results, but produced no analytic results of use to my project. The numerical results were of little use as they did not apply to the parameters used in my experiments.

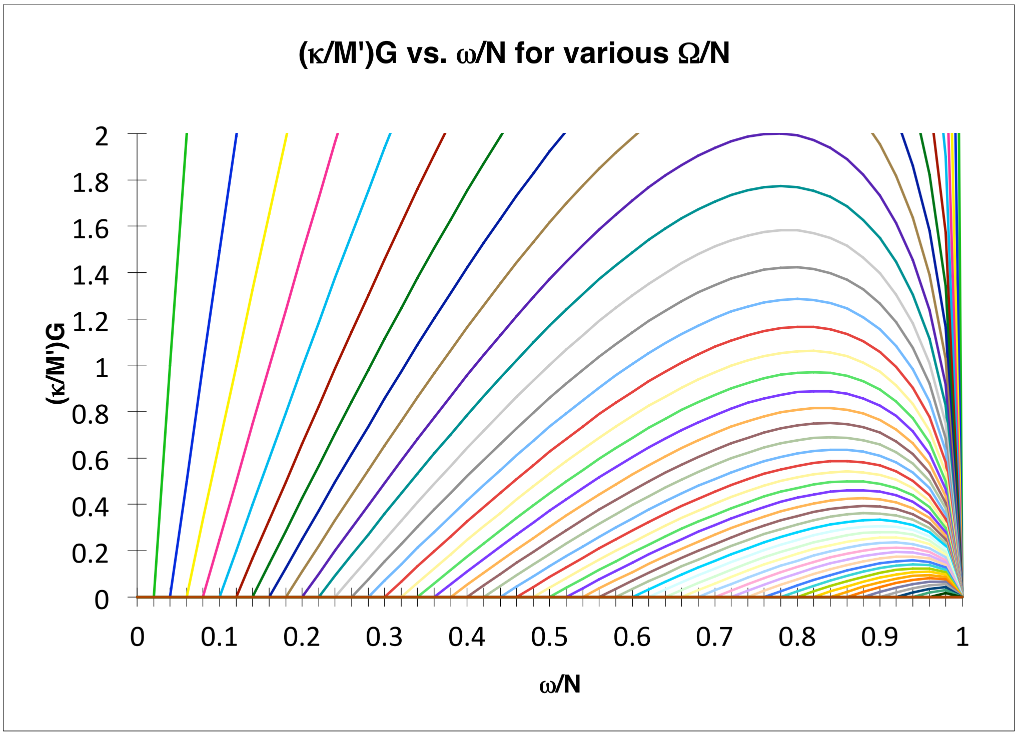


Fig. 2.6.1: Normalised G vs. ω for various Ω.

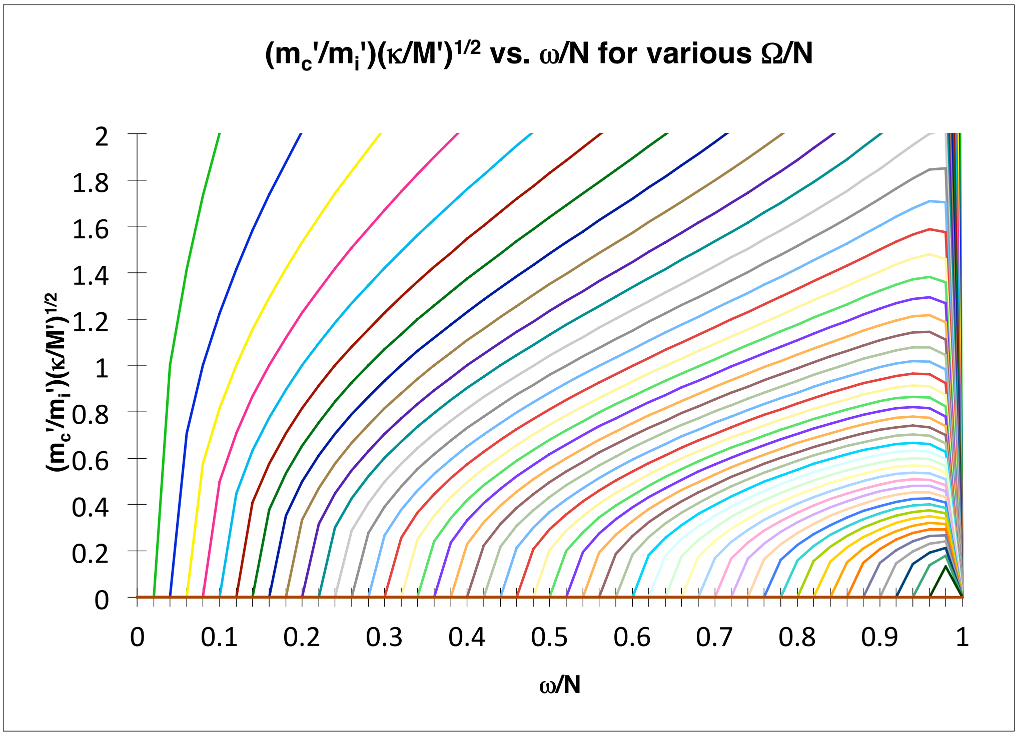


Fig. 2.6.2: Normalised m_c'/m_i' vs. ω for various Ω.

my apparatus could not produce a large separation in wavelength and frequency between the test and background waves.

Chapter 3: Experimental methods

3.1 Overview of the apparatus

This project aimed to use internal gravity waves to observe the strongly nonlinear interactions predicted by the numerical work of Broutman and others. To achieve this I needed a density-stratified medium and sources of both long, large amplitude waves and short, weak waves. The apparatus and methods used to observe the waves and their interactions are discussed in Chapter 4.

The experiments were conducted in a rectangular glass tank with internal dimensions 50 × 60 × 20 cm, giving a volume of 60 L (see Fig. 3.1.1). There was a hole in the centre of the tank's base through which it was filled with either a salt (NaCl) or sugar (sucrose) solution of varying concentration, creating a density-stratified solution. The tank filling process is presented in section 3.2, and the effects of diffusion on a gradient so produced are discussed in Appendix 2.

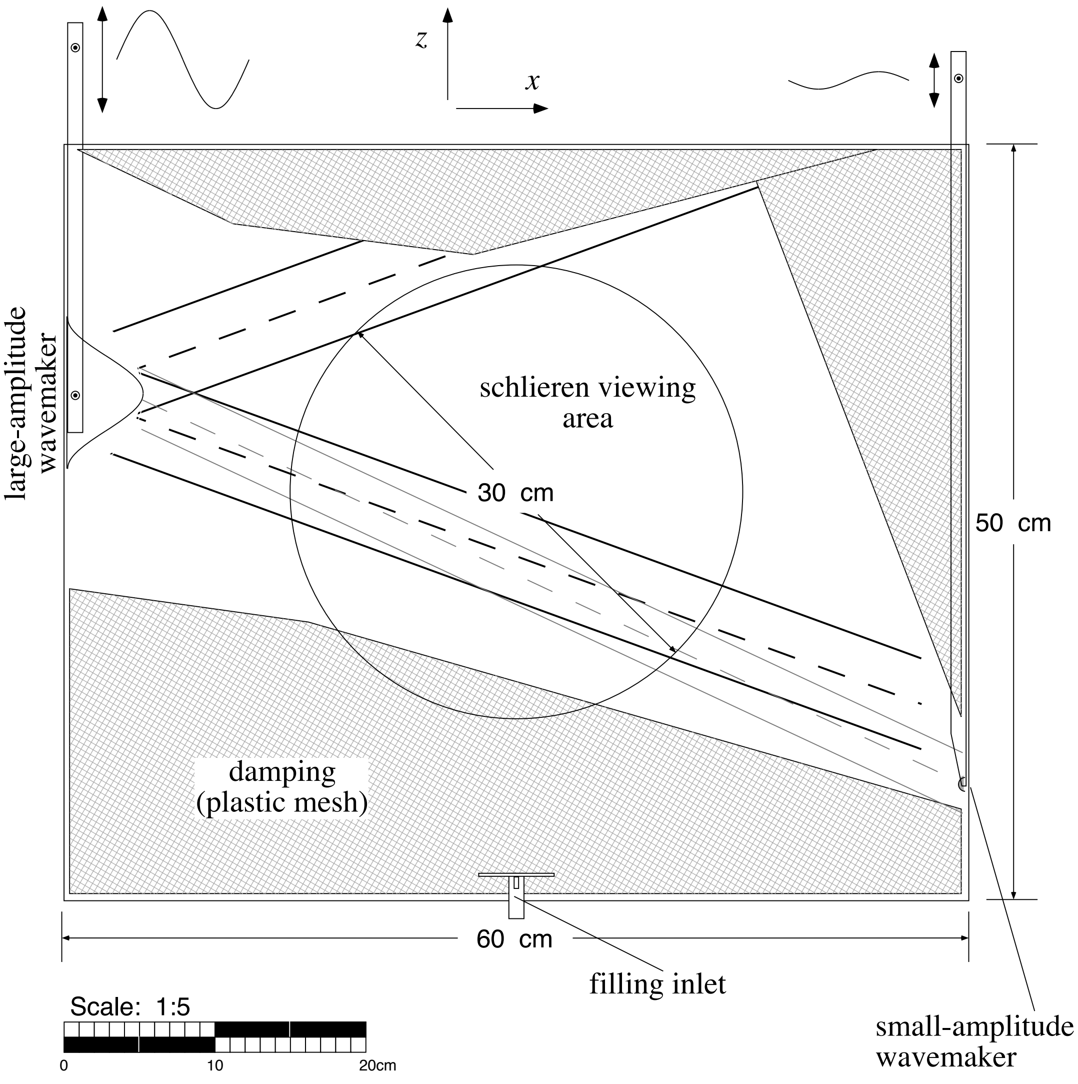


Fig. 3.1.1: The experimental tank.

Since both the theory and the optical visualisation methods are simpler for two-dimensional flows, the apparatus inside the tank was designed to have very little variation in the *y* direction throughout the 20 cm internal width of the tank. As salt water was used to produce most of the stratifications, the metal parts of the apparatus which were in contact with the solution were made from either aluminium or stainless steel to reduce corrosion.

At either side of the tank were mechanisms (built by Mr Tony Beasley) for supporting and oscillating the two wavemakers -see Fig. 3.1.2. Each consisted of a vertically sliding section driven by a DC electric motor via a crank arm. Each crank arm was 68 cm long and strokes did not exceed ±3 cm, so the wavemaker motion was very close to sinusoidal. The stroke of each crank was continuously variable between 0 and ±11 cm, and the speed of each motor was also continuously and independently variable, from 0 to 0.9 rad/s for the background wave and from 0 to 3.0 rad/s for the test wave. The wavemakers themselves were two-dimensional forms cut from Perspex and supported at an adjustable height by two metal arms. Thus the two wave sources were variable in frequency, stroke, mean vertical position and in profile. Several different wavemaker profiles were made and tested - some examples are shown in Fig. 3.1.3. The large wavemakers produced greater particle displacements and thus stronger waves.

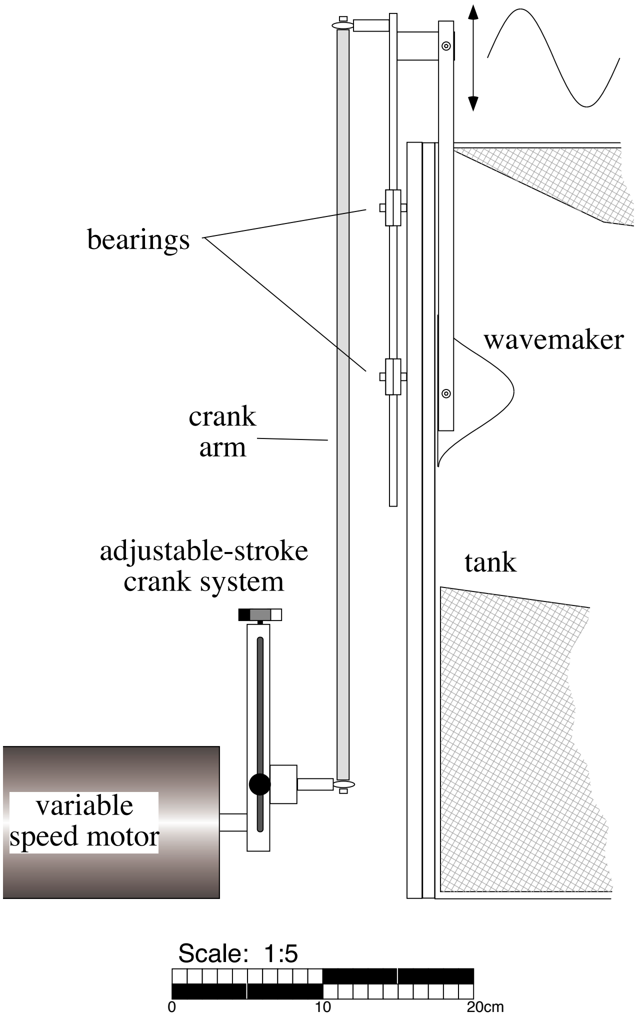


Fig. 3.1.2: The crank system for oscillating a wavemaker.

The wavemakers were held flush against the end walls to avoid mixing. As illustrated in Fig. 3.1.1, they produced waves in the manner shown in Fig. 2.1.3, but only one side of the cross pattern was produced due to the presence of the end wall. The wall coincides with a plane of symmetry of the flow pattern in Fig. 2.1.3 (considering it to be two-dimensional), so its presence makes no difference to the flow, apart from small boundary-layer effects due to viscosity. The wave generators were placed at the ends of the tank in order to increase the volume of the tank which could be used for experiments, and also to reduce the number of wave beams present.

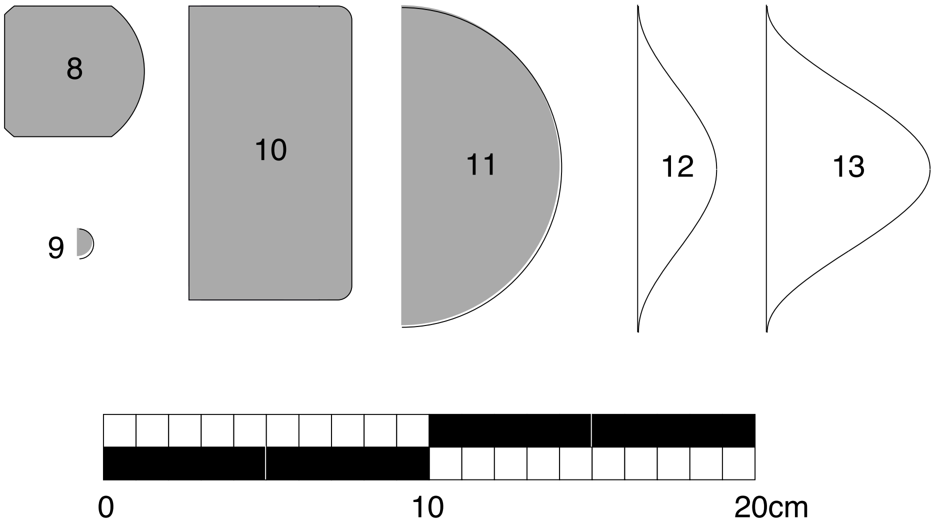


Fig. 3.1.3: Cross-sections of some of the wavemakers used.

The waves generated by these sources were strongly reflected by the tank walls and the water surface, so a considerable amount of damping was made in order to reduce this effect (see Fig. 3.1.1). Effective damping makes the tank response more like that of an unbounded system, as it reduces the Q of the resonant cavity formed by the tank so that resonant modes are not so important. The damping was made from plastic flyscreen mesh supported on wire frames, and reduced the wave amplitude by increasing the viscous dissipation of its energy. Damping was found to be most effective when its outer layers were very diffuse. This reduced reflection from the damping itself, and provided extra damping on the way out for any waves reflected by deeper layers of mesh or the walls of the tank. After some refinements the damping was sufficiently effective to make any reflected waves undetectable.

A colour schlieren system provided a means of visualising the wavefield. The system viewed a circular area of the tank 30 cm in diameter, as shown in Fig. 3.1.1. Further details on the schlieren system and other detection methods used can be found in Chapter 4.

3.2 Setting up a density gradient

A reasonably constant density gradient was created by the apparatus shown in Fig. 3.2.1, described in Oster (1965). To start with, the two identical buckets contained solutions at the maximum and minimum densities required for the tank. The buckets were connected by a pipe with a tap and the bucket containing the less dense solution had a second pipe connecting it to the bottom of the tank via a tap and flow meter. The volume of solution in each bucket was carefully adjusted to ensure that there was no pressure difference between the two ends of the pipe connecting the buckets, and that the total volume was that required to fill the tank.

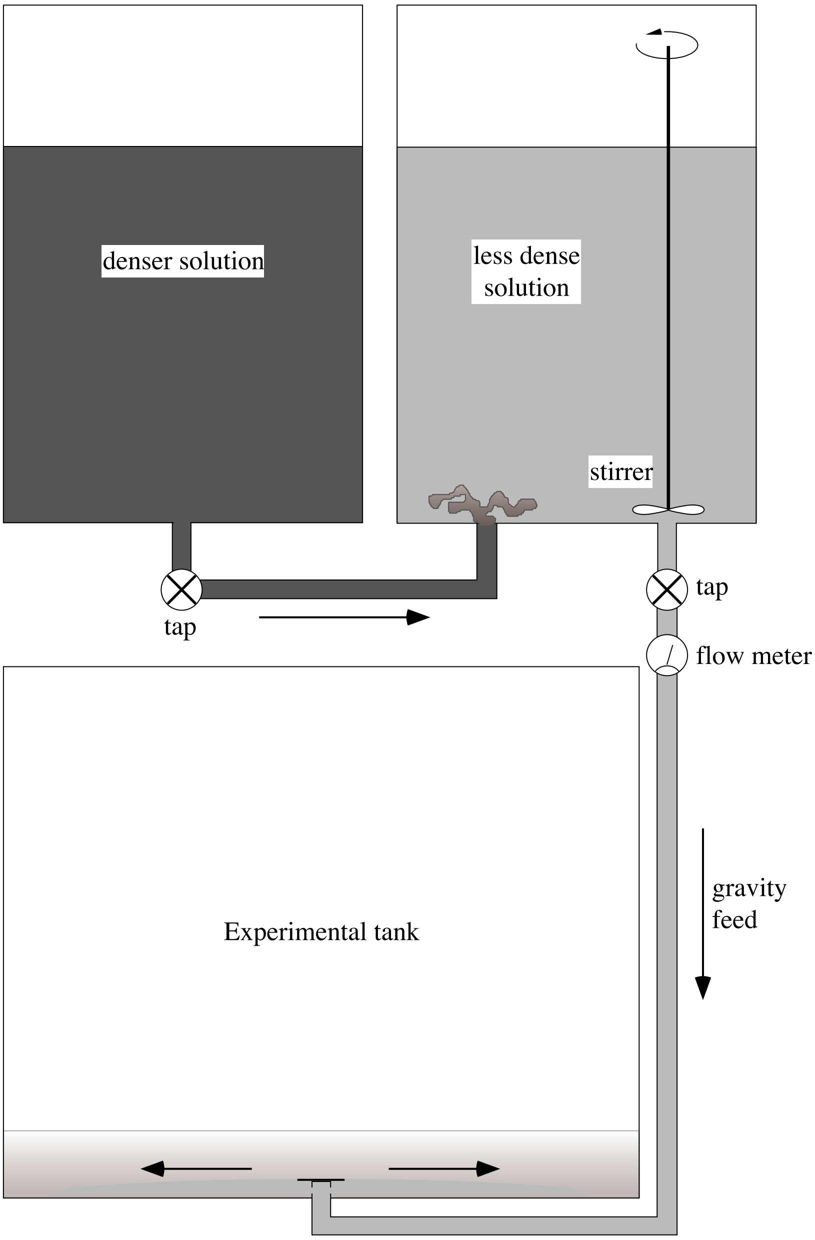


Fig. 3.2.1: Tank filling apparatus.

The bucket containing the less dense solution had a stirrer which ensured that its contents was kept thoroughly homogenised. Water in the mixed bucket was fed *slowly* by gravity into the bottom of the tank, where a baffle ensured that it spread horizontally along the bottom. Dense water flowed into the mixed bucket to equalise the bucket pressures, continually increasing the fluid density in the mixed bucket which fed into the tank. New fill water spread under the water already in the tank, gradually lifting the less dense older layers towards the top.

The result was a tank filled with water whose density increased monotonically with depth. In order to obtain a reasonably constant density gradient the volume flux into the tank needed to be kept constant, as incoming fluid mixed with that already in the tank to a degree which depended on the flow rate. To keep this mixing constant, the flow was continuously monitored by the flow meter and adjusted manually using the tap throughout the filling process.

If the flow rate from the dense bucket was half the flow from the mixed bucket into the tank, and the horizontal area of the tank was constant throughout its height, the tank fill had a linear variation of density with depth. To achieve this equipartition of flow from the buckets, the flow into the tank needed to be slow enough so that the two buckets remained close to equilibrium. This slow filling also reduced the amount of mixing in the tank and made the solution in the mixed bucket more homogeneous. When the horizontal area of the tank was reduced at some height (for example, by having a wavemaker of large horizontal size in the tank), the gradient was reduced in this region. This turned out to be more of a problem for the schlieren system than for the waves themselves - this will be discussed further in Chapter 4. The damping shown in Fig. 3.1.1 was sufficiently diffuse not to adversely affect the gradient in this way.

Once the tank had been filled, it was left to diffuse for a few hours, as diffusion acts to make the density gradient more constant in the body of the tank. Very smooth gradients were obtained by filling the tank over about 2.5 hours and leaving it to diffuse overnight. The effects of diffusion on salt or sugar gradients are discussed in detail in Appendix 2 where it is shown that in the short term (hours to days), diffusion tends to linearise the gradient at intermediate depths. However over a few weeks, diffusion reduces the depth range over which the density profile is linear, until eventually the entire tank has a uniform density. This large-scale diffusion is sufficiently slow to allow experimentation with an essentially constant stratification for a week or two in salt water, or three times as long with a sugar gradient (as sugar diffuses three times more slowly than salt). Typical density gradients used in my experiments had a density increase of 15% over the 50 cm depth of the tank, giving a buoyancy frequency of about 1.65 rad/s.

In practice, diffusion was not the only mechanism which disturbed the gradient. The flow around the wavemaker could separate in turbulent eddies which mixed the gradient, the mixed fluid intruding across the tank at its equilibrium depth. If the wavemakers were smoothly shaped, kept flush against the end walls and not moved too rapidly (ie. the product of frequency and stroke kept sufficiently small), this mixing could be kept at a negligible level. Convection could be caused by differences between the tank and ambient temperatures, and by the cooling and increased solute concentration of the uppermost layers by evaporation. The first cause was effectively countered by keeping the room temperature constant and avoiding direct sunshine on the tank, as well as filling the tank with water at the ambient air temperature. Evaporation was minimised by floating polystyrene foam on the top of the water, and enclosing the top of the tank in several layers of polyethylene plastic wrap.

A more serious danger for the stratification was gross mixing caused by moving the apparatus inside the tank. It was extremely difficult to change such things as the position of the damping or the wavemaker shape without so damaging the stratification that it would need to diffuse for several days to recover, if it could recover at all. Thus the tank contents needed to be planned and positioned before the tank was filled if good results were to be obtained. The process of mixing solutions, filling the tank and allowing diffusion to smooth the gradient took at least two full days, severely limiting the number of experimental geometries which could be tried.

Chapter 4: Wave observation methods

4.1 *Overview*

As shown in section 2.1 (equations (2.1.3) - (2.1.5)), internal gravity waves involve periodic variations in the velocity, density and pressure of the fluid, and any of these could in principle be observed to measure the properties of the waves.

The velocity at a point can be measured down to 1 mm/s by using a thermistor probe connected to a constant-temperature bridge circuit (McEwan, 1971; McEwan, Mander & Smith, 1972). Alternatively, by adding tracers to the fluid (such as dye streaks or neutrally buoyant particles), the velocity field can be inferred by measuring the displacement field (Merzkirch, 1974). Adding rheoscopic particles ("fish flakes") to the fluid allows the shear structure of the waves to be seen, again giving a (qualitative) indication of the velocity field.

The density field can be measured by a number of different methods. The refractive index of a salt (NaCl) or sugar (sucrose) solution is related to its density (see Fig. 4.2.1), as is the conductivity of a salt solution (Fig. 4.5.1) and the optical rotation of a sugar solution (Fig. 4.6.1). The variation of refractive index can be observed using shadowgraph, Moiré, schlieren or interferometric techniques (Merzkirch, 1974; Oster, 1965; Holder & North, 1963), whilst conductivity probes and polarimeters can be used to measure conductivity and optical rotation, respectively (Head, 1983; Lambert & Davey, 1974). In the course of my experiments, all of these methods except interferometry were used to observe the density field, and in addition dye streaks were used to observe the fluid displacements. In practice, the pressure fluctuations were too small to be easily measured.

4.2 *Shadowgraph*

The refractive index of a salt or sugar solution is nearly linearly related to the density of the solution, as Fig. 4.2.1 shows. The refractive index *gradients* resulting from spatial variations in the density of the solution will refract light rays, thus allowing the density variations to be detected.

The shadowgraph method is the simplest technique used for observing refractive index variations. It involves merely illuminating the test region (ie. the tank) with parallel light from a small source, and placing a screen on the other side of the test section (see Fig. 4.2.2 below). Light rays which pass through regions in which the second derivative

$$\nabla_v^2 n \equiv \frac{\partial^2 n}{\partial x^2} + \frac{\partial^2 n}{\partial z^2} \tag{4.2.1}$$

of the refractive index n is not constant will be brought closer together or spread further apart, depending on the sign of this second derivative (see Fig. 4.2.2 for the orientation of the coordinates). This results in variations in brightness of the screen's illumination, giving an image of the variations of $\nabla_v^2 n$ across the test region (Holder

& North, 1963; Merzkirch, 1974). This lensing process is identical to that which produces a pattern of ripples on the bottom of a pool under sunlight.

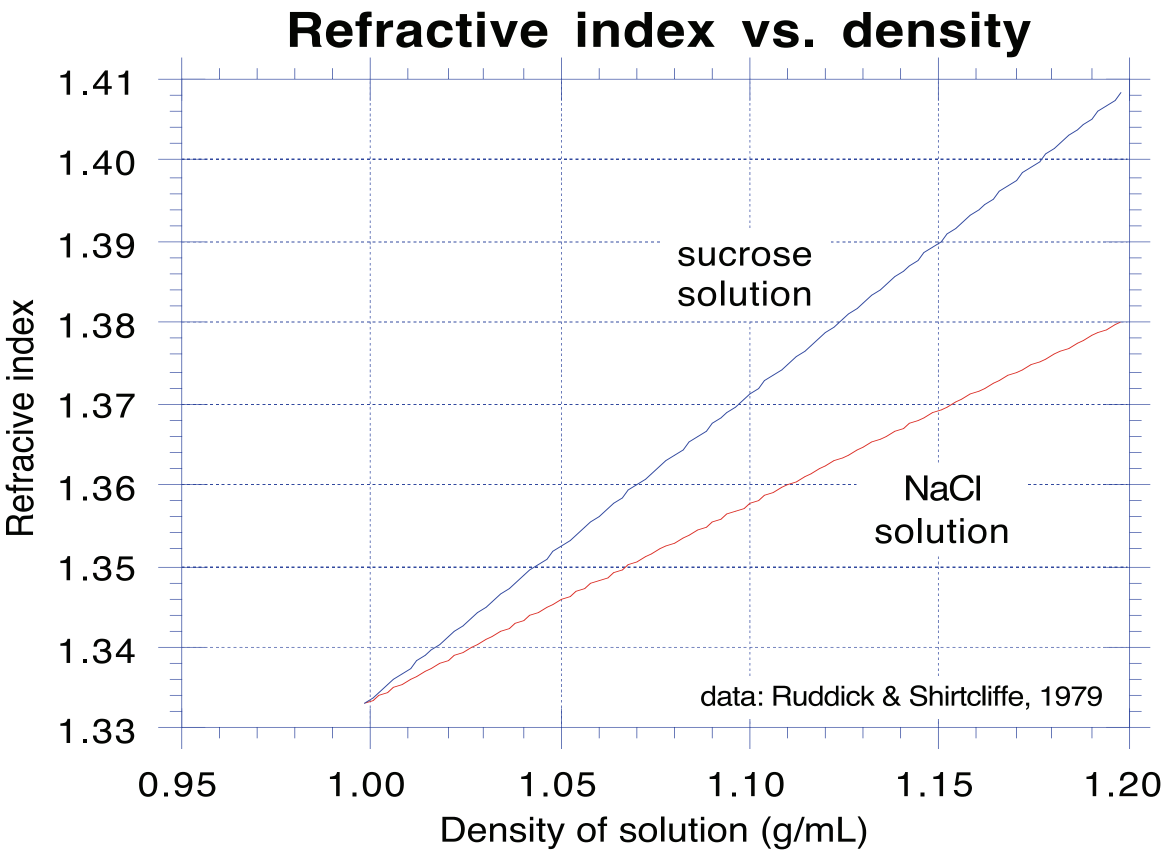


Fig. 4.2.1: Refractive index of salt and sugar solutions vs. density.

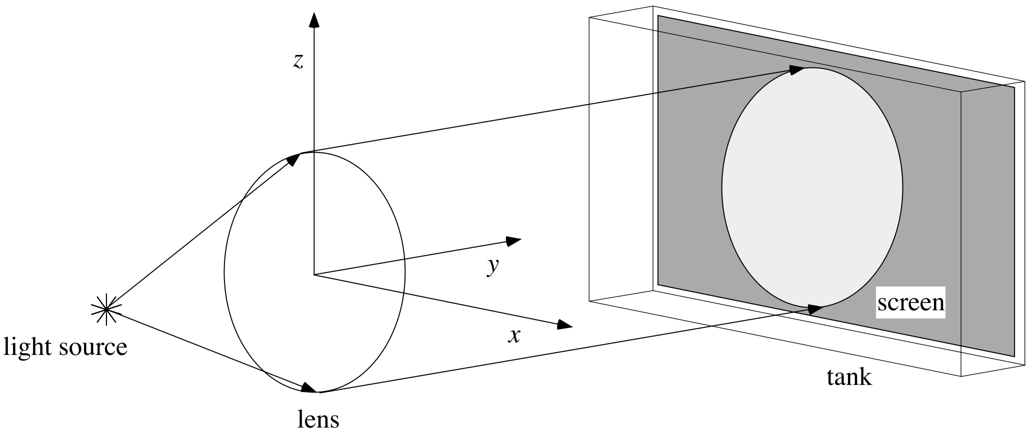


Fig. 4.2.2: A shadowgraph system.

Because the shadowgraph technique visualises *second* derivatives of the refractive index, it is not very sensitive and is difficult to use to observe internal gravity waves (although it has been done - see Thorpe, 1989). It has the advantage of being extremely simple to set up - all that is needed is a slide projector as a light source and a sheet of tracing paper to form the screen. The projector is placed at some distance from the test region so the light is reasonably parallel. I used a shadowgraph to monitor turbulent flow separation around the large wavemaker so that the attendant mixing could be minimised.

4.3 Moiré technique

A simple modification to the shadowgraph technique allows it to detect variations in the *first* derivative of the refractive index, greatly increasing its sensitivity (Ivanov, 1989; Oster, 1965). Transparent sheets printed with a large number of parallel black lines are placed on either side of the test section with their lines at a small angle. The screen shows overlapping shadows of the two sets of lines, which form a Moiré pattern. To a very good approximation, a refractive index gradient will refract light through an angle proportional to that gradient (Mowbray, 1967), thus shifting the relative positions of the two sets of lines on the screen. The Moiré pattern is very sensitive to such changes in the relative positions of the lines, and so enables the detection of slight changes in the refractive index gradient. The system is sensitive only to refractive index gradients normal to the lines, but this directional sensitivity can often be an advantage if gradients parallel to the lines are unimportant.

A Moiré system was set up briefly and found to be sufficiently sensitive to detect internal gravity waves, but was not used extensively as it was less effective than the colour schlieren system already set up. The Moiré technique may be worth considering for further work, as it is far easier to get working than schlieren but has similar sensitivity. This technique can also be used to view a larger tank area than is possible with a schlieren system.

4.4 Schlieren techniques

Due to their sensitivity, schlieren methods are often used for visualising the refractive index changes produced by internal gravity waves. The basic idea is shown in Fig. 4.4.1. Light diverging from a narrow slit at the focus of a parabolic mirror (M1) forms a collimated beam of light. This passes through the test section and is incident on a second mirror (M2) which directs the light to a lens which forms an image of the test section. A knife edge is placed in the focal plane of M2, where an image of the source slit is formed, and positioned to block half the source image, producing a uniform loss of intensity in the image of the test section.

A refractive index gradient in a region of the test section will refract the light from that region, forming an image of the slit at the focus of M2 which is slightly shifted in relation to the image formed by undisturbed rays. If that shift has a component normal to the knife edge (so the amount of the image cut out by the knife edge is different), the image of the test section will show the shift as a different brightness in the disturbed region. This is illustrated in Fig. 4.4.2.

Geometrical constraints require the light reflected by the mirrors to be off-axis (as in Fig. 4.4.1), which results in two main aberrations: coma and astigmatism. The coma of M1 is cancelled out by that of M2 if the angles of the Z-shaped setup in Fig. 4.4.1 are equal. The astigmatism is not removed in this way, and becomes more severe as the angles in the Z are increased (Holder & North, 1963; Merzkirch, 1974).

The astigmatism results in the source image having a focal distance from M2 which is slightly longer for horizontal than for vertical directions (considering the Z in Fig. 4.4.1 to be in a vertical plane). If the slit and knife edge are not vertical or horizontal, the image plane of the source is not well defined and the resulting schlieren image of a homogeneous test section is unevenly illuminated. This problem is eliminated if the slit and knife edge are vertical or horizontal, as this gives a well defined image plane for the source (the focus in only one direction is relevant).

A standard schlieren system as described above produces images of the test section which display differences in the refractive index gradient normal to the knife edge as different shades of grey. Such a system can be made very sensitive, and is a standard visualisation method used in shock wave research and many other fields (Kleine & Grönig, 1991). Since the first rather than the second derivative of *n* is visualised, schlieren techniques are usually more sensitive than shadowgraphs. However, as it is only sensitive to refractive index gradients in one direction (unlike shadowgraphs), it can sometimes be difficult to interpret the images of complex flow patterns.

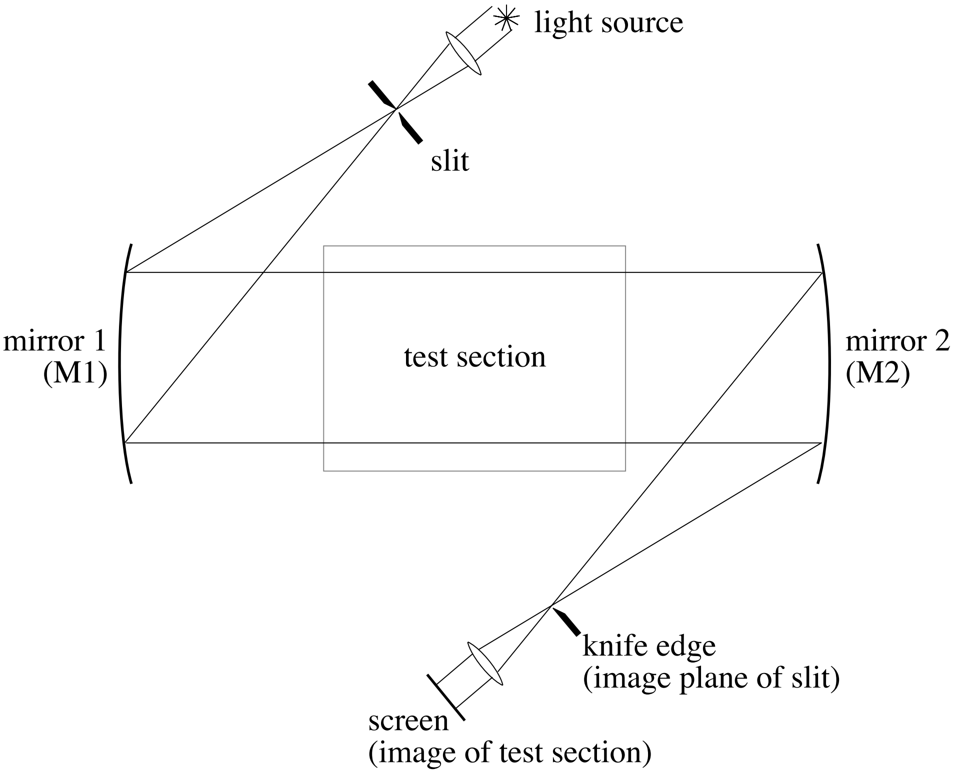


Fig. 4.4.1: A typical schlieren system.

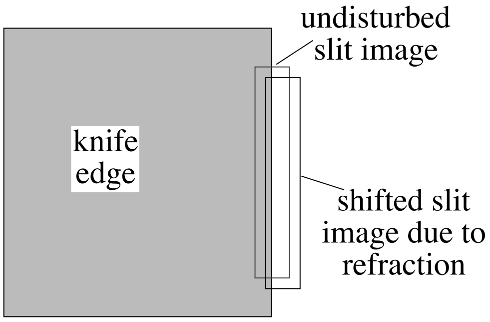


Fig. 4.4.2: Shift in the slit image due to refraction.

In order to display gradients in different directions in a single image, a colour-coding system can be used. In such a colour schlieren system, the source slit is replaced by a pinhole (with a white light source) and the knife edge by a transparent slide with a multicoloured pattern (see Fig. 4.4.3). The image of the pinhole passes through different colours on the slide depending on the refractive index gradients it has encountered, and by a suitable choice of slide pattern this allows the visualisation of two-dimensional gradients. Also, the position of the pinhole and slide may be interchanged, which can provide some advantages such as allowing the use of schemes to eliminate the effect of astigmatism and not limiting the imaging of the test section by the transparency of the slide.

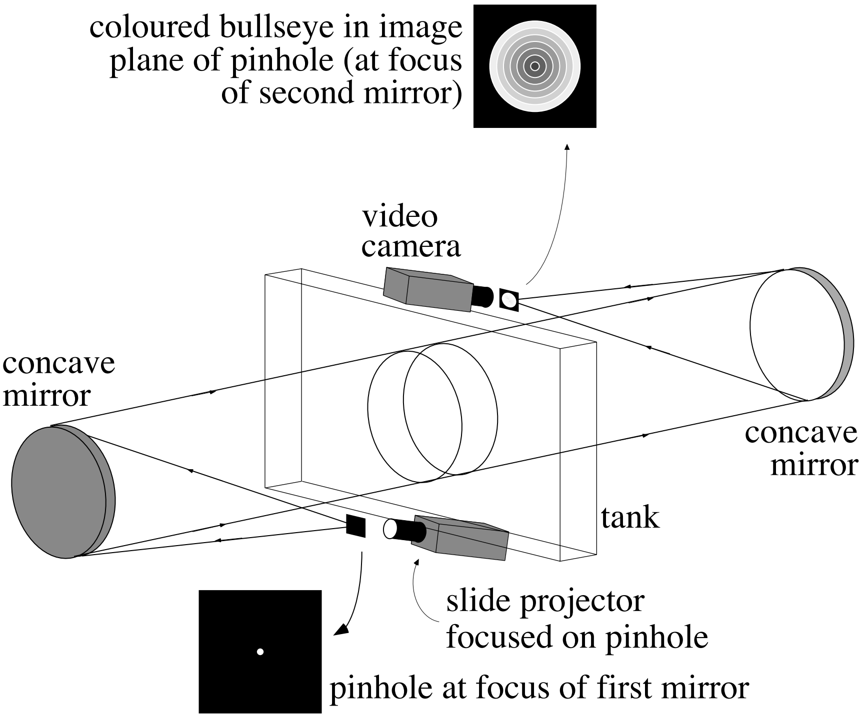


Fig. 4.4.3: The colour schlieren system.

The refractive index of sugar or salt solutions is very nearly a linear function of their density, as Fig. 4.2.1 illustrates. Thus an undisturbed stratified fluid will have some refractive index variation with depth, which will refract the parallel beam used in a schlieren system. Mowbray (1967) studied the Euler-Lagrange equations (which describe the refraction of light rays through a medium with nonuniform refractive index) and showed that if the refractive index (and hence density) increases linearly with depth, the rays will all be refracted downwards by the same angle. Thus the schlieren technique can be used in this case by making the simple modification of lowering M2 and tilting it slightly so that the refracted parallel beam is collected by M2 and reflected to the knife edge or filter. This will result in a uniform image in the absence of waves, and a system which visualises *perturbations* from this linear density stratification, that is, it shows $\nabla_v \rho'$ rather than $\nabla_v \rho$.

It is convenient that diffusion initially acts to produce such a constant density gradient through the middle section of the stratification, as shown in Appendix 2. However as diffusion proceeds, this linear portion of the stratification becomes narrower, gradually decreasing the region of the tank in which the schlieren technique is useable. As shown in Fig. A2.3, this degradation also reduces the depth range through which N is nearly constant and internal waves are relatively unrefracted.

The system used in my experiments is shown in Fig. 4.4.3. The mirrors used were 30 cm in diameter with a focal length of 2.44 m, made by OWL. The white light source was a standard Kodak slide projector (with a 250 W halogen lamp) placed underneath the bench which supported the tank. Many different pinholes were tried, but the most effective was a hole in black cardboard about 0.5 mm in diameter. About 35 different slides were produced by drawing the required pattern on a computer and outputting this to a colour slide recorder. The selection of slide patterns produced allowed a large range of sensitivities to both refraction direction and magnitude. The CCD colour video camera was fairly standard: a National F10, with a zoom lens set at 55 mm. The camera could display a stopwatch in a corner of the screen, giving a timebase for the film. A scale reference was provided by a 5 cm grid drawn on the tank, and the schlieren images were recorded in time-lapse or real time on VHS tape.

A considerable amount of effort was required to develop a colour schlieren system of sufficient sensitivity to detect internal gravity waves. The first problem encountered was bulging of the Perspex tank, which in effect formed a very poor lens, strongly refracting the light and making the schlieren image unusable. The glass tank used for all subsequent experiments was much more rigid, but still bulged sufficiently to alter the colours around the edges of the schlieren image. The glass used in the tank was 8 mm thick, and to rebuild it using thicker, more rigid glass would have been prohibitively expensive. In addition, the image was disturbed by the slightly uneven thickness of the glass.

In order to partly overcome these difficulties, a very strong gradient was used. This increased the refraction due to the waves compared to that caused by the tank, and also aided the detection of the waves. The gradients used typically had a 15% change in density over the 50 cm depth of the tank - the gradient could not have been much steeper over this depth without saturating the salt solution at the bottom of the tank.

To obtain sufficient sensitivity to detect the weak test waves, the colour bands in the slides needed to be about 0.5 mm wide or less. With this sensitivity, it was impossible to avoid also seeing the distortions of the tank. The image of the pinhole (and hence the pinhole itself) needed to be at least as small as the stripe width in order to give good colour definition in the image. This reduced the available light, but thankfully the CCD camera was just sensitive enough to produce useable images (although images became excessively noisy if the pinhole size was reduced much below 0.5 mm).

Another difficulty with the schlieren system was that there was a depth range with a reduced gradient due to the effect of the large-amplitude wavemaker on the filling process. This had only a minor effect on the waves (waves were refracted by less than 5° for most frequencies), but produced strong vertical refraction of the light passing through this region. Any slide whose pattern varied vertically thus produced an image in which the colours were distorted in this region. Slide patterns consisting of vertical stripes were used instead of bullseyes to eliminate this effect (see Fig. 5.2.4 for an example).

The angle between the parallel beam through the test section and that from the pinhole (or to the camera) was about 12°. This angle could not be reduced due to space restrictions and resulted in considerable astigmatism, the vertical and horizontal foci of M2 being about 4.5 cm apart when the tank was empty. This made it very difficult to obtain an even background colour when using a slide with a two-dimensional pattern. One solution I developed was to put the coloured slide in front of the light source, then use two slits in front of the camera. A vertical slit was placed at the horizontal focus, and a horizontal slit at the vertical focus, the net result being the same as if a square aperture had been used in a non-astigmatic system. Although this system worked very well when the tank was empty, any identifiable vertical or horizontal foci were destroyed by the lensing which occurred when the tank was full (due to the tank's distortions). Thus in practice the two-slit approach was no better than just putting a pinhole approximately in the focal “plane” of the second mirror.

4.5 The conductivity probe

The conductivity of a salt (NaCl) solution depends strongly on the electrolyte concentration, which in turn determines the density of the solution at a given temperature. This is clear from Fig. 4.5.1, which shows a curve fit from Ruddick & Shirtcliffe (1979).

A commercial miniature four-electrode conductivity probe (model 5021 from Precision Measurement Engineering) and associated electronics were used to measure the density fluctuations in a very small region around the probe tip, which is shown in Fig. 4.5.2. The probe tip was attached to a long stainless steel shaft and lowered vertically into the tank to the desired location (the upper damping had holes through it to allow this).

The probe electronics applied a 10 kHz alternating current between the outer two electrodes and monitored the potential difference between the voltage probes. Virtually no current was drawn by the voltage probes, so the reading they give was unaffected by their surface impedances (Head, 1983). An alternating current was used to avoid electrolytic effects.

The control electronics included a feedback loop which adjusted the current to maintain a constant RMS potential difference between the voltage sensing electrodes. The DC output voltage signal from the electronics was proportional to the RMS current needed to maintain the constant RMS voltage between the probes, and was thus proportional to the conductivity of the solution. The output was recorded by a chart recorder at first, and later by a data acquisition and analysis program I developed for the purpose (this is discussed in section 4.7 and Appendix 3).

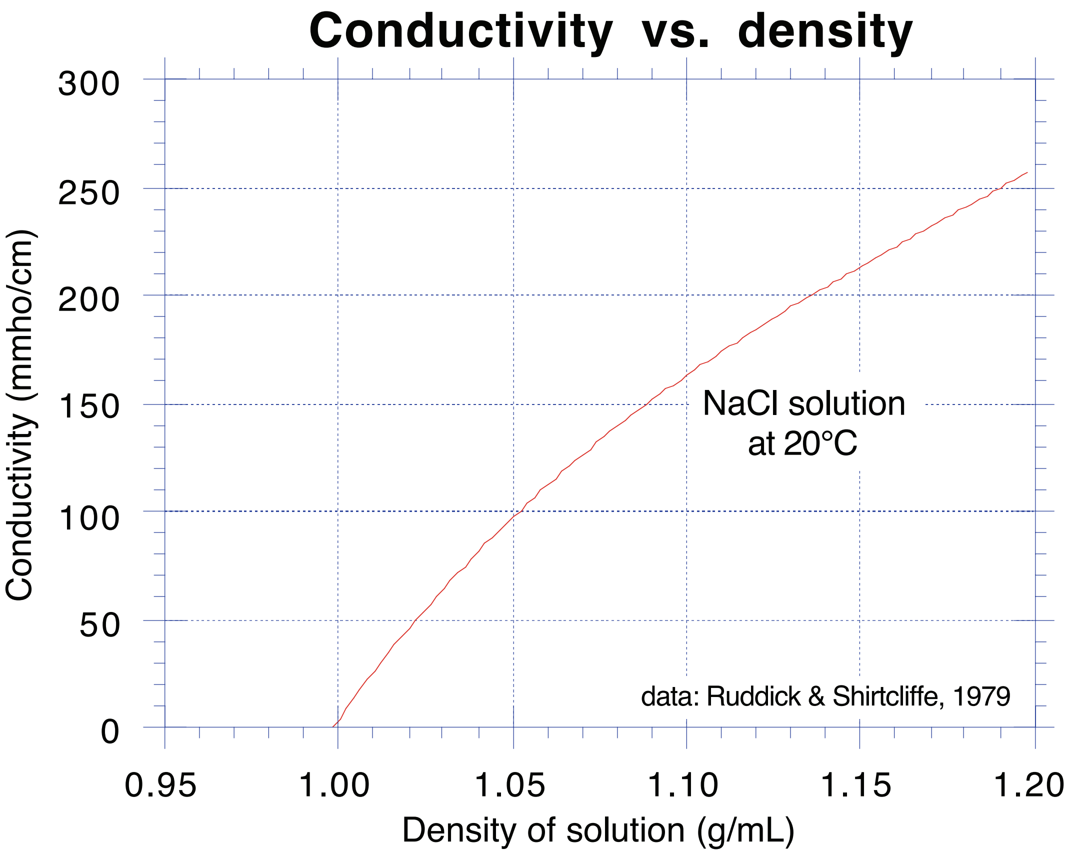


Fig. 4.5.1: The relationship between density and conductivity of a salt solution.

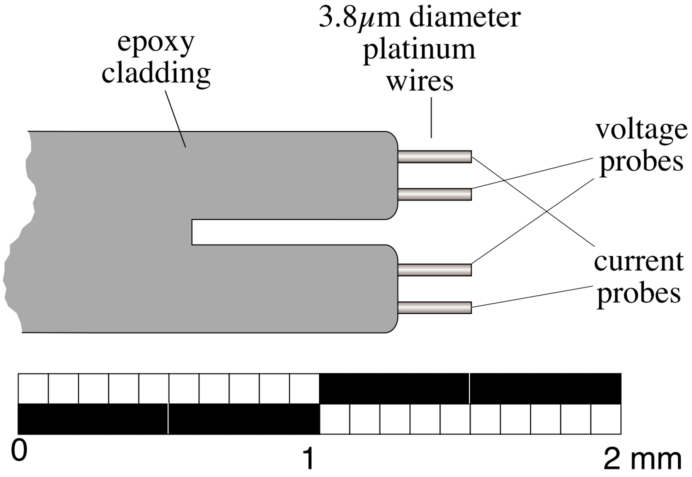


Fig. 4.5.2: The four-electrode conductivity probe (scale diagram).

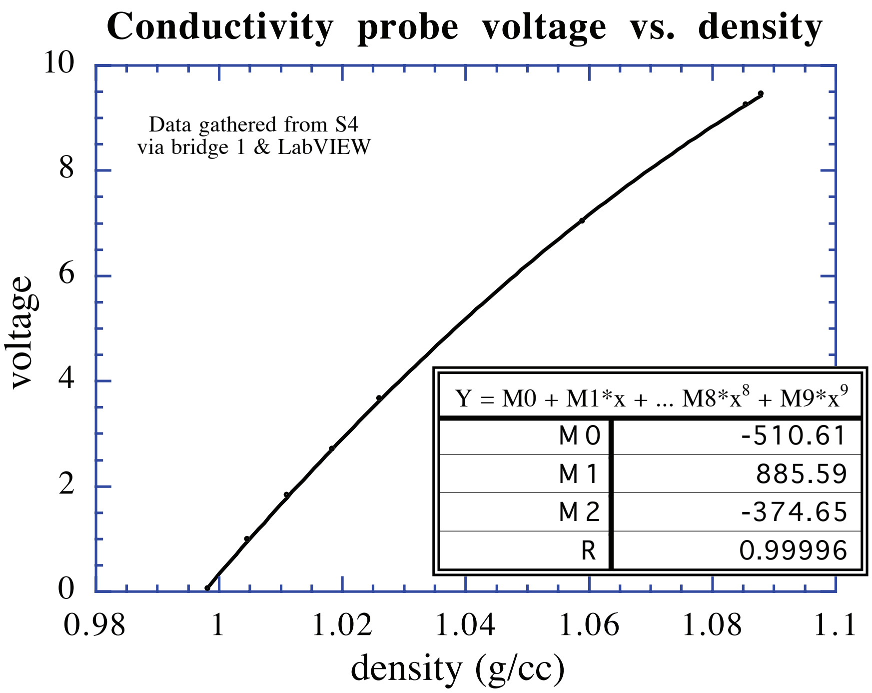


Fig. 4.5.3: Calibration of the conductivity probe.

A calibration curve for the conductivity probe is shown in Fig. 4.5.3. The nonlinear relationship between conductivity and density evident from this figure was insignificant for the very small density fluctuations of the waves I investigated, and so conductivity changes were taken to be directly proportional to density changes. As I was interested only in the spectral content of density variations, the output signal was not converted into an absolute density. The probe reading averaged over a region of radius 0.2 mm in the immediate vicinity of the electrode wires and had a response time of 2 ms (Head, 1983) - this was more than adequate for observing the internal gravity waves I produced.

4.6 Polarimetry

Many media exhibit the property of *optical activity*, which means that they rotate the plane of polarisation of linearly polarised light traversing the medium. This effect is due to the medium exhibiting a different refractive index for left- and right-circularly polarised light. These are orthogonal polarisations, so *any* polarisation may be expressed as a linear combination of them. Thus a plane polarised light wave can be decomposed into left- and right-circularly polarised waves with a phase lag between them determined by the linear polarisation angle. In an optically active medium, the left- and right-circular polarisations travel with different speeds, so the phase difference between them (and thus the direction of plane polarisation) changes by an amount proportional to the distance traversed in the medium - that is, the plane polarisation direction rotates as the wave propagates through the medium (Guenther, 1990).

Only media with some "handedness" can treat the two circular polarisations differently and thus be optically active. Many molecules are *chiral*, meaning they are different from their mirror images and thus come in distinct right- and left-handed forms. Solutions of such molecules are optically active. The degree of rotation is proportional to the wavenumber of the light, the distance traversed and the refractive index difference for the two circular polarisations. In a solution of a chiral solute, this last quantity is closely proportional to the solute concentration, and can thus be used as method for measuring the concentration. The concentration in turn gives the density of the solution, which fluctuates in the presence of internal gravity waves and can thus be used for their detection.

Sodium chloride ionises in solution to the obviously achiral ions Na^+ and Cl^-, so salt water is not optically active. However sucrose (table sugar) *does* produce optically active solutions, and this was the solute used when polarimetry was utilised to detect the waves.

Fig. 4.6.1 (based on the cubic curve fits of Ruddick & Shirtcliffe, 1979) shows the angular rotation per centimetre of a sucrose solution as a function of its density.

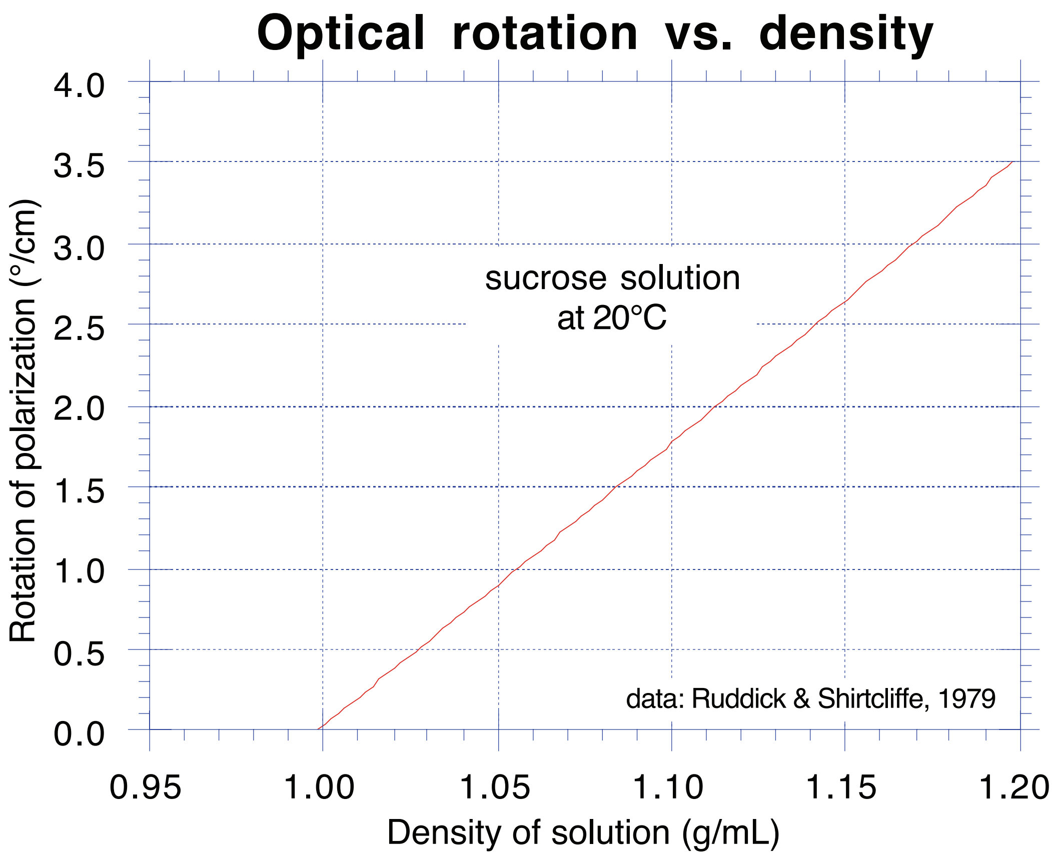


Fig. 4.6.1: Optical rotation at 589 nm vs. density of a sucrose solution.

An internal gravity wave involves small oscillations in the density, which will result in small changes in the direction of polarisation of polarised light passing through the medium. As Fig. 4.6.1 makes clear, the change in polarisation direction is very nearly linearly related to the change in density (especially when the density fluctuations are small). Detection of these small changes in the polarisation direction (less than 1°) requires a sensitive instrument.

The polarimeter used had been designed and built at RSES some years previously by Mr. Derek Corrigan and Dr. Barry Ruddick, based on a paper by Lambert & Davey (1974). It allowed a continuous, direct readout of polarisation, and was sensitive to changes as small as 0.03° when correctly aligned. Fig. 4.6.2 is a simplified diagram showing its principle of operation. The polarised light source was a Spectra-Physics 4 mW helium-neon laser (632.8 nm), whose output thus had an extremely well-defined wavenumber and was also highly polarised, with an extinction ratio of approximately 1000:1 due to Brewster windows in the laser cavity. Thus it formed an excellent source as it gave an unambiguous rotation angle and also sampled the medium along a narrow beam, effectively making a point measurement of the two-dimensional flow.

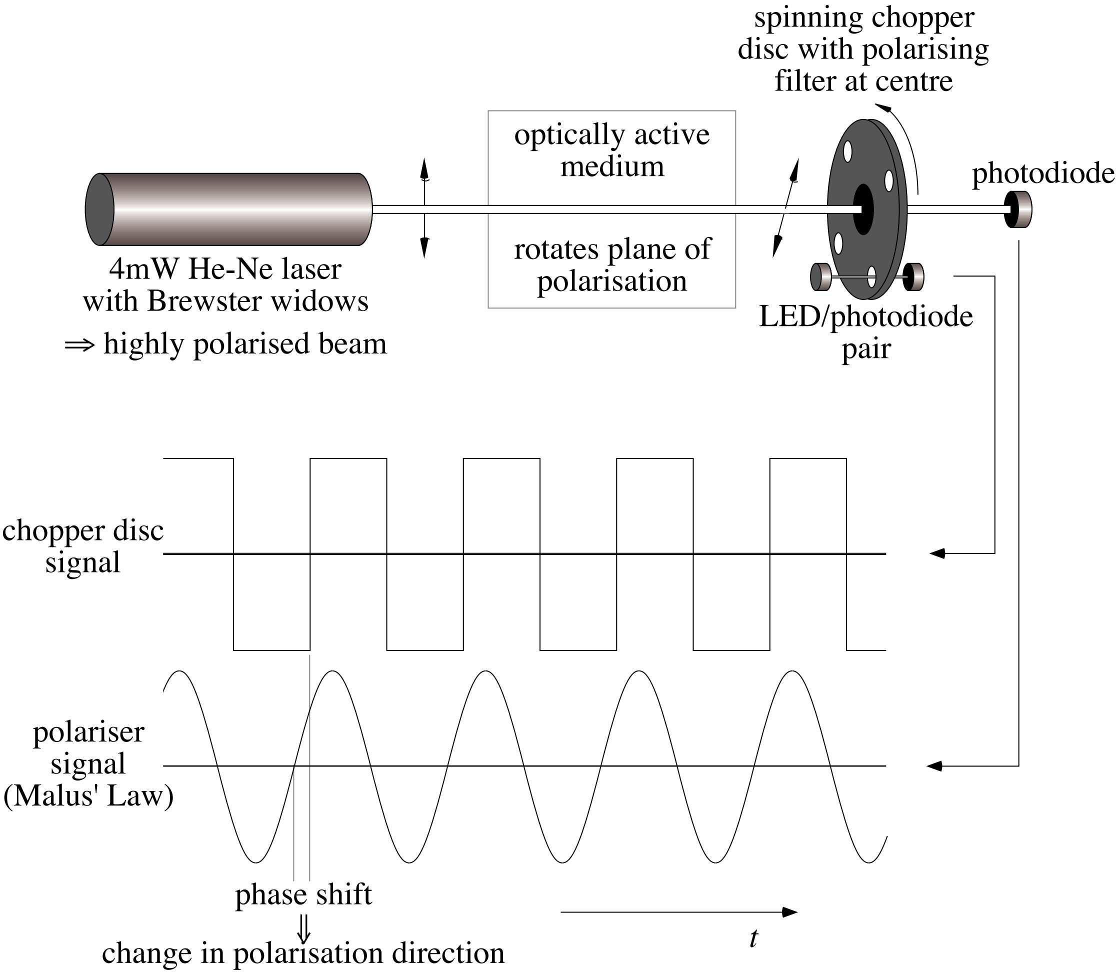


Fig. 4.6.2: Schematic of the polarimeter.

The beam was aligned along the *y* axis and passed through the tank to the detector apparatus. This consisted of a spinning chopper disc (driven at 136 rad/s (1300 rpm) by a servo-stabilised motor) with a polarising filter at its centre. The laser beam passed through this polariser and onto a photodiode, where, by Malus' Law, it produced a sinusoidal signal at twice the frequency of the rotation. An LED/photodiode pair had its light path interrupted by the chopper disc, and the signal produced by this was used to generate a square wave which provided a reference phase. The phase shift between this signal and the sine wave produced by the polariser was proportional to the angle by which the polarisation had been rotated and hence to the integrated density along the light path through the tank.

The electronics detected the time difference between the zero-crossings of the two signals and converted this to an analog output proportional to the angle of rotation, and also provided a digital display of the rotation angle in degrees. A data acquisition program (see section 4.7) was used to record and analyse the analog output signal. The zero point for the rotation angle was arbitrary and could be adjusted by simply rotating the laser about its axis - as I was interested only in *fluctuations* in the angle rather than an absolute measure of the angle itself, the zero point was set to any convenient value.

Alignment of the polarimeter was a difficult task and involved a considerable amount of adjustment to minimise the noise level. This was done mainly by using an oscilloscope to monitor the output of the photodiode behind the polariser and carefully adjusting the laser alignment to obtain a stable sinusoidal signal. A lens could be placed at a variable position in front of the polariser in order to correct for slight misalignments. The lens also served to increase the acceptance angle of the detector, thus allowing for refraction of the beam by internal gravity waves in the tank. In order to avoid reflections from the tank feeding back into the laser, the beam was put on a slight angle to the *y* axis.

4.7 Data acquisition

I wrote a program for data acquisition and analysis using the LabVIEW package on a PC. The digitising board in the computer could acquire data from any analog source, such as the conductivity bridge, polarimeter, or the photodiode used in the attempts at making quantitative schlieren measurements. The program I wrote allowed the user to choose the sampling frequency and the number of samples to acquire and to perform a Butterworth lowpass filtering of this data, with an adjustable cutoff frequency and rolloff slope (this filtering capability was not actually needed in my experiments). The AC one-sided power spectra of both the raw and filtered data were automatically calculated, with the option of normalising the spectra. Plots of the raw and filtered waveforms and their power spectra were displayed. A timer was also provided, giving an absolute timebase to the samples which could be synchronised with an external timer (such as the on-screen stopwatch in the video recordings), allowing a direct comparison between the schlieren image and the detected signal.

In order to reduce high-frequency random noise the program included an oversampling capability. The user could choose some oversampling factor (say 50), increasing the actual sampling frequency by this factor. The data returned from the sampling would consist of averages taken over every 50 samples, thus giving the same number of data points and the same effective sampling frequency as if there had been no oversampling. If the signal was essentially constant during those 50 samples but was obscured by some symmetrically-distributed high-frequency noise, this averaging procedure could dramatically reduce the noise level, especially at high frequencies.

All the acquired data and analysis could be saved to a file readable with Microsoft Excel, so that further analysis, calculations, plots or printouts could easily be made. A data file was tabulated with the raw and filtered samples listed against the synchronised timebase provided by the timer (in both hours:minutes:seconds and total seconds). The raw and filtered power spectra were tabulated against frequency in Hz. A data file included the settings used to acquire it, providing a complete record for later reference.

A printout of the program and further technical details can be found in Appendix 3.

Chapter 5: Results

5.1 *Observed wave properties*

A basic requirement for a laboratory system to display significant shear-induced refraction of the type predicted by Broutman and collaborators was to have sufficiently small initial values of G and m_c'/m_i'. As shown in section 2.6, these conditions depend strongly on the wavelengths and frequencies of the waves - as the frequency was easily varied, the challenge was to produce an appropriate wavelength ratio. Furthermore, I needed a background velocity field of sufficient strength to produce caustics in the background wave.

When wave frequencies were sufficiently high to separate the upgoing and downgoing waves from a source, schlieren images could be used to determine wavelengths. This was not altogether straightforward, as the viscous wave envelopes contained only about one wavelength (see section 2.3), and the effect of a number of co-propagating wavelengths at the same frequency would be difficult to differentiate from that of an envelope. Furthermore, since the schlieren system detects refractive index *gradients*, it is more sensitive for short waves than long waves of the same amplitude.

Further wavelength information was found by using tracers. Small crystals of potassium permanganate ($KMnO_4$) dropped into the tank left dyed trails which were advected by the velocity field. Tiny bubbles occasionally rising from the bottom damping also served as tracers, as they dragged up denser fluid in a trail behind them, which was visible in the schlieren image for a few seconds before diffusing away[3].

Fourteen different wavemakers in a wide variety of shapes and sizes were used in an attempt to produce waves of significantly different wavelength (see Fig. 3.1.3 for some examples). It had been hoped that the wavelength would scale with the wavemaker size, allowing a large wavelength difference between the background and test waves, and thus a small value of G. However, as shown in section 2.4 the situation is complicated by the effects of viscosity. As predicted in that section, at a distance of 30 cm from the source the shortest wavelength which could be produced was about 4 cm. Wavemakers smaller than the viscous wave scale L_v (about 1.4 cm in my experiments) all produced the same wavelengths - reducing the wavemaker diameter from 1 cm to 0.5 cm only reduced the wave amplitude. This result was

---

[3] Although the bubble trails involved only small density changes, these changes were over very small distances, and so produced refractive index gradients sufficient to be detected by the schlieren system. I attempted to make a bubble injector to utilise this effect, but it proved very difficult to produce sufficiently small *single* bubbles. A fine "spray" of hundreds of bubbles of the right size (and hence sufficiently small rising velocity) could easily be produced, but this was of little use as a tracer and was more damaging to the gradient than a single bubble.

disappointing, as it meant that the scale of the waves would not be so small compared to that of the schlieren viewing field and the tank.[4]

Having a fairly large minimum wavelength would have been tolerable if much larger background waves could be produced. Unfortunately, as mentioned in section 2.4, increasing the wavemaker diameter beyond L_v results in bimodal wave beams rather than a longer wavelength. Such a large diameter source acts in a similar way to a pair of sources smaller than L_v located at the top and bottom of the large wavemaker (wave generation is most effective at the top and bottom of the wavemaker because that is where the greatest fluid displacements occur). Thus a large wavemaker produced two pairs of strong beams (one pair propagating upwards, the other downwards). The wavelengths of these beams was no more than about 7 cm, giving less than a factor of two difference between the wavelengths of the test and background waves.

As illustrated by Fig. 2.6.1, to achieve a small value of G under these circumstances requires the test wave to have a frequency only slightly greater than that of the background wave. Although the validity of theory based on the ray equations is highly questionable under such circumstances (little difference between the ray frequency and wavelength and the temporal and spatial variation of its surroundings), there is no firm line at which ray theory suddenly breaks down. The experiment went ahead to see whether the predictions of Broutman's theory could be extrapolated to this laboratory situation.

Throughout this thesis the “short waves” approximation (2.1.7) to the dispersion relation (A1.30) has been used, which applies when $\kappa >> N^2/2g$. The smallest wavenumber was that of the background waves, about $2\pi/7$ cm$^{-1} \approx 90$ m^{-1}. N was typically 1.6 rad/s, giving $N^2/2g \approx 0.13$ m^{-1}, clearly showing the validity of the “short waves” approximation.

Wave amplitudes were found by measuring the displacements of potassium permanganate streaks in the tank. It was found that the displacements were about ±1 cm for the background wave, and around ±0.05 cm for the test ray, giving a large contrast between the strengths of the waves. This is important, since the theory assumes that the test ray is too weak to affect the background wave. At a typical frequency of 0.55 rad/s, this gives peak velocities of 0.55 and 0.025 cm/s for the two waves. The peak background wave velocity was about a factor of 13 less than the minimum required to produce caustics predicted by equation (2.6.17). The required amplitudes could not be generated with my apparatus (see section 6.1) - unfortunately, this was not realised until it was too late to build anything new. However, the background waves were still sufficiently strong to exhibit nonlinear

---

[4] From (2.3.1), the limiting wavelength is proportional to the cube root of the kinematic viscosity μ/ρ. Unfortunately, there are few fluids which are significantly less kinematically viscous than water (apart from exotic things like liquid hydrogen or superfluid helium). The most appropriate fluid I found in Lide (1991) was $CHCl_2F$, a CFC refrigerant with one-quarter the kinematic viscosity of water which would thus reduce the limiting wavelength to 63% of the value in my experiments. This is not a very significant reduction, and this liquid boils at 8.9°C, making it difficult to work with. Apart from the question of its effectiveness as a solvent (so it could be density-stratified), using this liquid would also have reduced the background wavelength as discussed in the next paragraph.

behaviour, as will be shown in section 5.4. Their phase velocity was about 0.6 rad/s - this value is only slightly greater than the fluid velocities, and is an indication that the waves cannot be treated as infinitesimal.

Measurements were made of the ray propagation angle as a function of the wave generator frequency. As shown in Fig. 5.1.1, the measured values of $\cos\Theta$ fell on a straight line intersecting the origin when plotted against Ω. The line in this figure is the theoretical prediction using the value of N calculated from the density change over the depth of the tank. Although this is not a new result (it was verified in 1966 by Mowbray & Rarity), it at least shows that the behaviour of these nonlinear waves remains constrained by these predictions of linear theory.

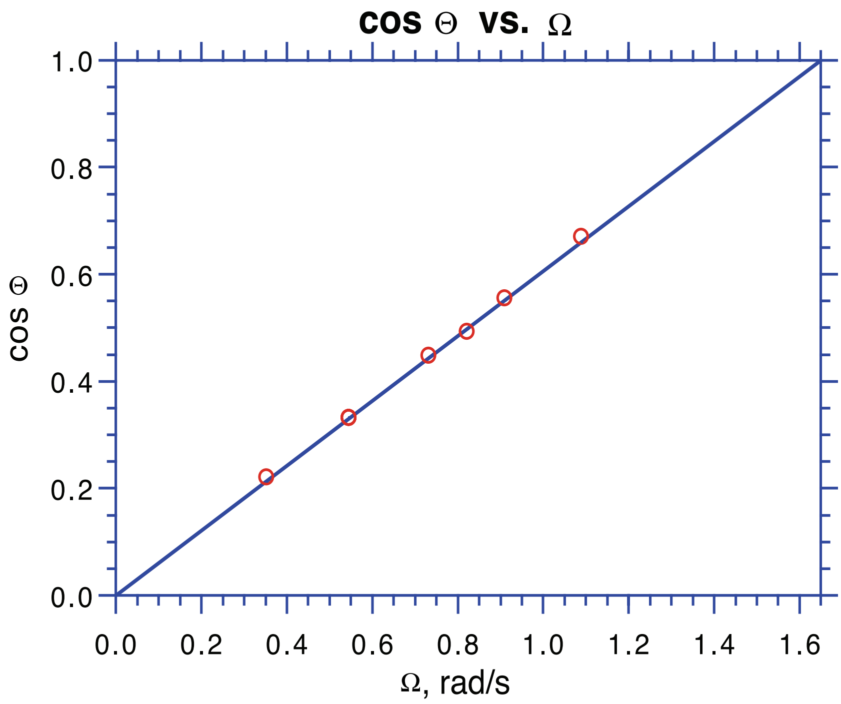


Fig. 5.1.1: cosΘ *vs.* Ω *for N=1.65 rad/s.*

5.2 *Schlieren results*

The initial idea for the experiments was to observe the effect of the strong wave on the test wave by using the schlieren system. Wavelengths could be seen directly, and frequencies could be deduced from the angles of the wavefronts (or speed of the motors for the incident waves). As I had no clear theoretical predictions of the frequency of a refracted ray (and hence its angle), it seemed ideal to have a technique such as schlieren which allowed the whole interaction region to be surveyed. The schlieren system was used extensively to observe the waves and fine-tune the design of the damping to intercept the waves in the desired way. It was invaluable as a means of seeing the wavefield, but proved to be less useful for observing wave refraction itself.

About 10 hours of video footage was recorded through the schlieren system, much of which was recorded at 3/14 time lapse to aid visual perception of these low-

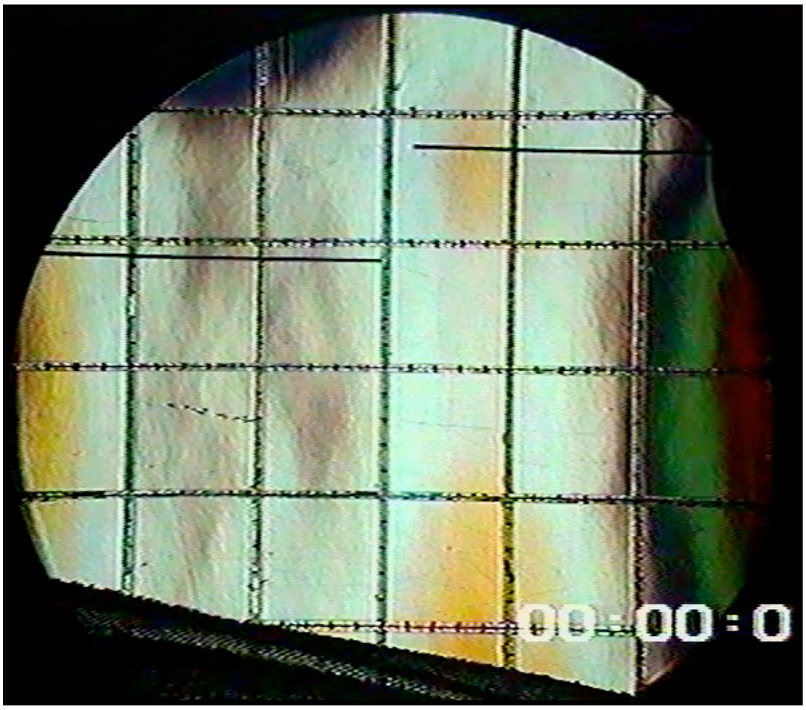


Fig. 5.2.1: Schlieren image of tank at rest.

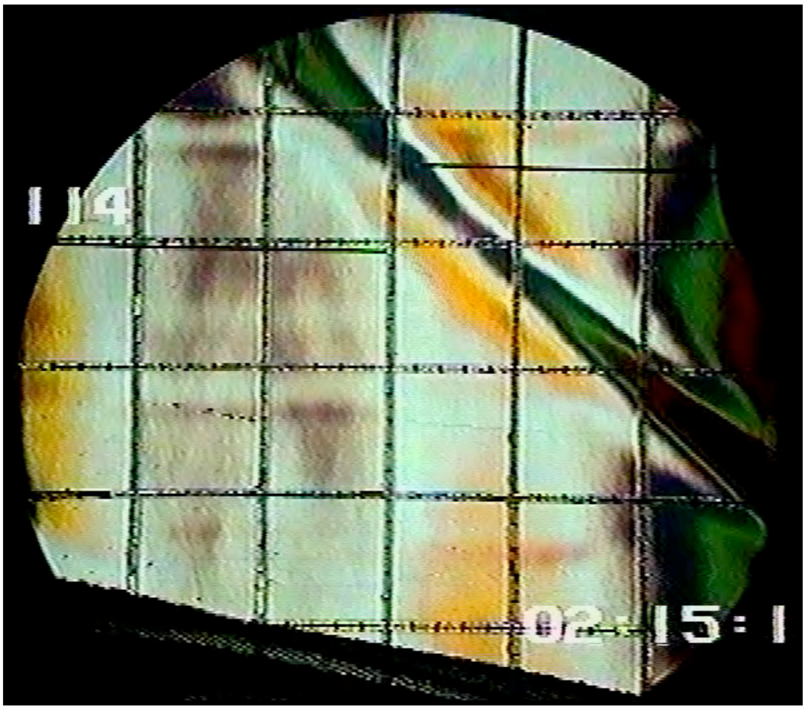


Fig. 5.2.2: Schlieren image of test wave.

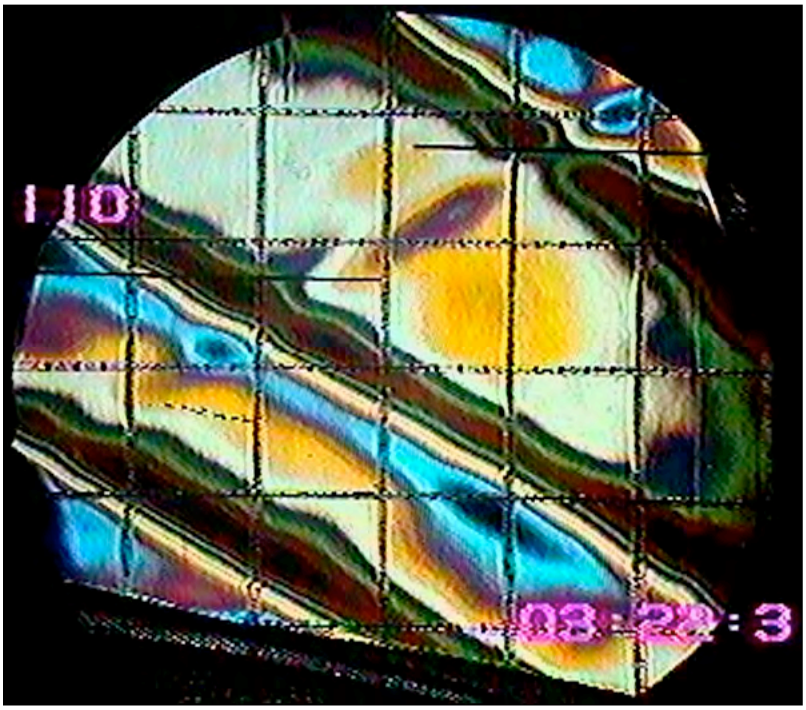


Fig. 5.2.3: Schlieren image of background wave.

frequency waves. The schlieren system was refined until it was sufficiently sensitive to detect the weak test waves - Figs. 5.2.1, 5.2.2 and 5.2.3 show typical images of the tank at rest, with test waves and with background waves respectively. The area visible in these figures is shown in Fig. 5.2.5. Both wave sources were outside the field of view: the test wavemaker was at the lower right, and the background wave source was to the left. The scale grid in these figures has lines with a 5 cm spacing, and damping is visible at the bottom of each image. A reference time is shown at the bottom right and the run number is displayed at the top left.

These images were obtained with the slide shown in Fig. 5.2.4, which was aligned vertically and was thus sensitive only to horizontal refractive index gradients. This made the system oblivious to the strong vertical refraction through the region of reduced density gradient due to the large amplitude wavemaker (as discussed in section 4.4). Unrefracted rays passed through the clear central stripe, giving an uncoloured background. Horizontally refracted rays passed either through the primary or secondary colours depending on the sign of the horizontal refractive index gradient, and were given different colours from each group depending on the degree of refraction. Rays refracted by a very large amount missed the stripes altogether and produced black regions in the schlieren image. From Fig. 2.1.2 it is clear that the maxima of the refractive index *gradient* coincide with the regions of maximum flow speed in a single wave beam - thus refraction maxima correspond to regions of maximum velocity either parallel or antiparallel to the group velocity, these two flow directions producing horizontal refractive index gradients of opposite sign so wave “peaks” and “troughs” appear with different colours when using the slide in Fig. 5.2.4.

The colours visible around the edges of Fig. 5.2.1 are due to a slight bulge in the tank (a few millimetres over 60 cm), and those in the central region result from slight inconsistencies in the thickness of the glass tank walls. This was confirmed by draining the tank - the colours at the edges disappeared, but those in the centre remained. The density gradient was made very steep to increase the refraction due to the waves relative to that caused by the tank, but when the schlieren system was sensitive enough to detect the test wave, it also saw the tank's distortions. Rays which pass through these distortions have a constant offset to their refraction in addition to the fluctuating refraction from the waves, so the colours in these regions vary about a colour other than white. This results in the apparent distortion of the wavefronts in Fig. 5.2.2. Although this spoilt the aesthetic impact of the images, it was of little importance for my experiments since the visibility of wavelengths and wavefront angles was virtually unaffected.

The background wave image (Fig. 5.2.3) displays a pair of downgoing wave beams from the top and bottom of the wavemaker (as shown in Fig. 5.2.5), and a reflected upgoing beam at the top right (the upper damping had been removed for this run). A weaker upgoing ray is also visible.

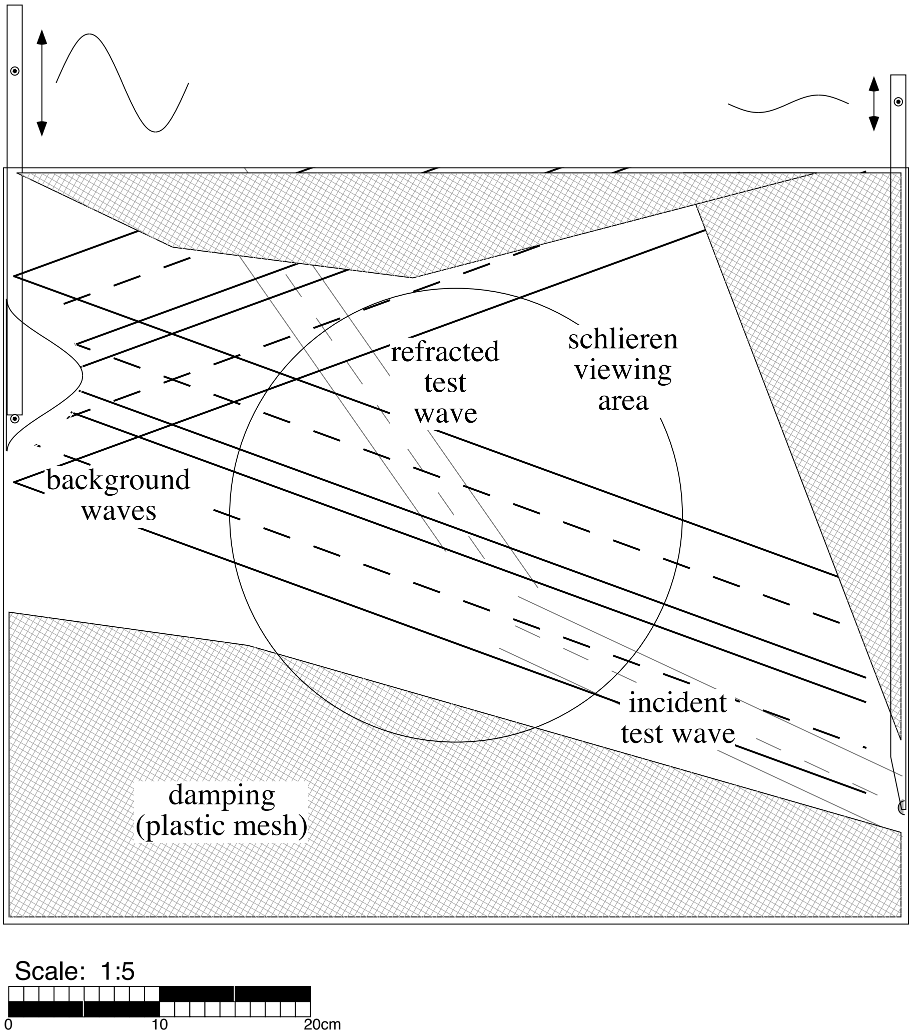


Fig. 5.2.5: Wave interaction geometry.

The experimental geometry is shown above in Fig. 5.2.5. The upgoing test wave intersected the downgoing background wave beam at a shallow angle, and any downshift in $|m'|$ as predicted by ray theory would result in refraction to a higher frequency and hence a steeper angle to the horizontal. Unfortunately, this sent any refracted wave into the upgoing beam from the background source. Although no strong refraction would be expected from this wave[5], it greatly obscured the view of any test wave refracted by the downgoing beam. This was a severe problem because the background wave was so much stronger than the test wave. Having vertical stripes for the schlieren slide helped slightly by making the system more sensitive to steeper wavefronts.

Several approaches were tried to reduce the obscuring effect of the upgoing beam. The angular separation between the beams could be enlarged by increasing their frequency, but this was of limited use since the downgoing beam needed to remain above the test wave source (which in turn needed to be kept at least 5 cm

---

[5] G is negative in this case because the wavefronts in the upgoing beam are propagating downwards, so caustics cannot form because an upgoing test wave packet will always be in motion relative to the background wavefronts.

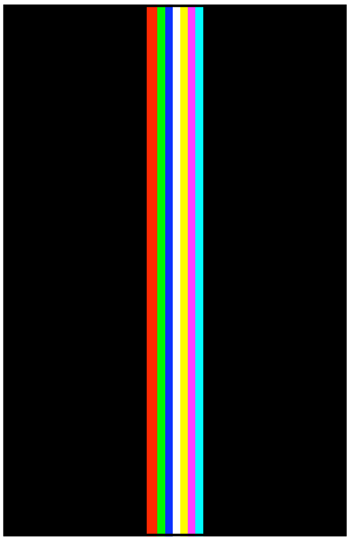


Fig. 5.2.4: A colour schlieren slide (twice actual size).

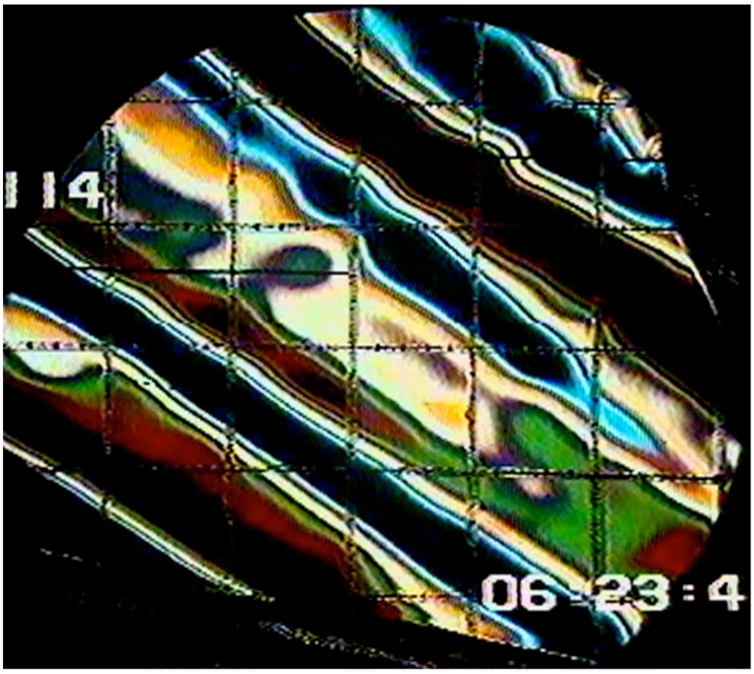


Fig. 5.2.6: Schlieren image of test waves (Fig. 5.2.2)
with strong background waves.

from the bottom of the tank to avoid the diffused region). Damping the upgoing beam was also found to be of little use, since there was insufficient space between the source region and the viewing area for the thickness of damping needed. This problem was exacerbated by the relatively large distance required for the two beams to separate out enough so that the upgoing beam could be damped without substantially affecting the downgoing beam.

Another problem with damping was that it restricted the range of wave frequencies which could be used, since beams of different frequencies (and hence angles) may not have been intercepted by the damping in an appropriate way. Damping could not be moved without badly damaging the gradient, so the tank had to be drained and refilled whenever a new damping geometry was tried. In the end, no satisfactory method was found for removing the unwanted waves.

The view was further cluttered by *second harmonic generation* from both wave sources. It was found that when the frequency of either wave source was less than $N/2$, a second, steeper set of wave beams was often produced in addition to the expected beams. The ray angles suggested that these extra waves had twice the frequency of those in the main beams, and measurements with a conductivity probe confirmed this and showed that the amplitude of the second harmonics was about 40% that of the primary wave beams. These waves were especially conspicuous because the schlieren system was more sensitive to their steeper wavefronts, so I spent about a week trying to understand and eliminate them (this is discussed in Appendix 4). Plausible generation mechanisms were hypothesised, but these did not provide a direct method for eliminating the effect.

Two remedies of limited usefulness were to position damping to remove the second harmonic waves (only suitable when the main wave beams were narrow and source frequency was close to $N/2$ so that the second harmonic was well separated), or to choose a wavemaker frequency above $N/2$ so that second harmonics could not form spatially separated beams. The first approach was effective in some cases, but the second was not of much use for the background waves, since they needed a frequency less than $N/2$ in order to remain above the test wavemaker. As the frequency of the test wave needed to be close to that of the background, this also had to be below $N/2$.

Thus the view of the test ray in the schlieren image was almost completely obscured by the two broad upgoing background waves, and often by the second harmonics as well. This is clear from Fig. 5.2.6, which is the same as Fig. 5.2.2 except that the background wave is also present (this is a stronger wave than that shown in Fig. 5.2.3). In the 32 experimental runs with both wavemakers operating (encompassing a wide range of different experimental arrangements and conditions), there was no detectable evidence for refraction of the test ray, mainly because it was almost invisible in the background wavefield.

If the gradient had not had a strong disturbance due to the large-displacement wavemaker, it would have been possible to align the stripes in the slide so that the schlieren system was insensitive to the gradients produced by the background wave. Although this would have also made it difficult to detect the incident test wave (due to its similar angle), a significantly refracted test ray would have been very conspicuous. As it was, the stripes needed to be vertical so that the region at the level of the background wavemaker could be seen.

5.3 Digital image subtraction

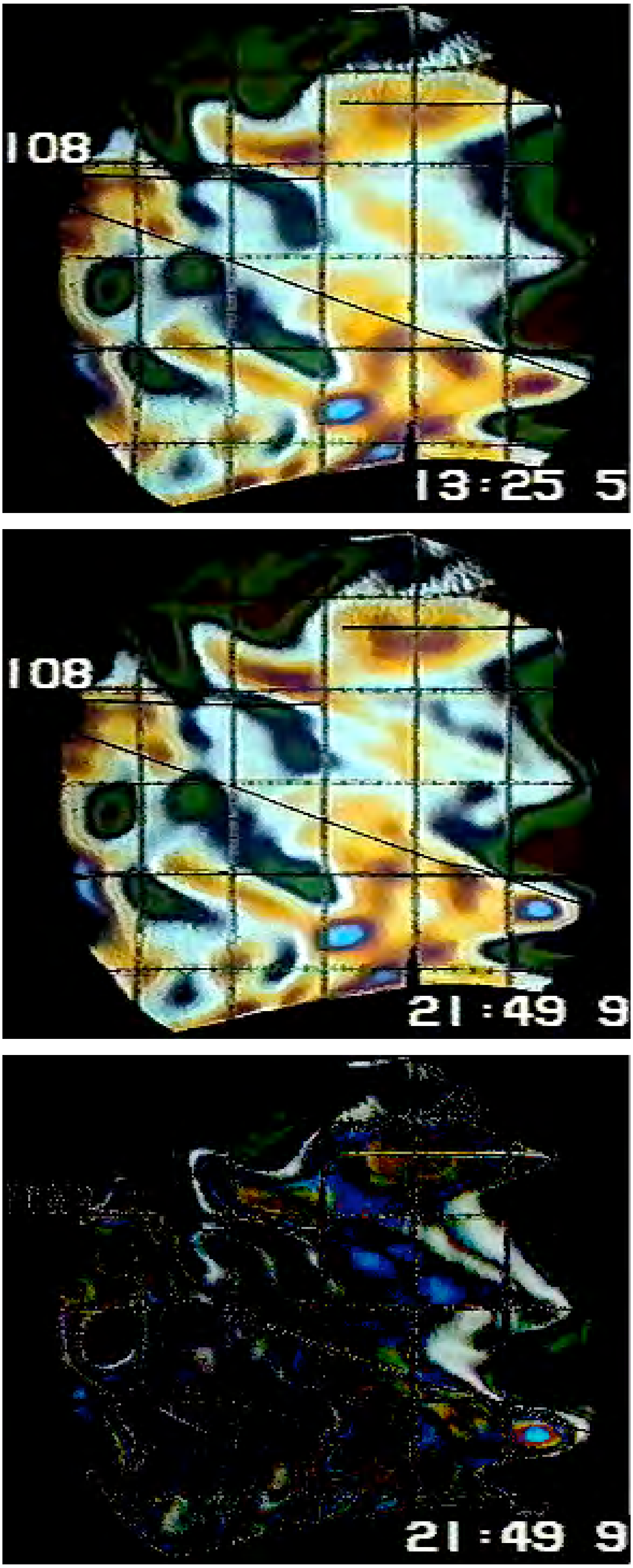
I attempted to use a digital image subtraction technique to remove the obscuring effect of the background waves. The principle was simple: obtain digitised images of the background field with and without the test waves, and “subtract” them, leaving only the effect of the test waves.

To create subtractions, I needed to select images which had the same phase of the background wave so that they would differ only by the presence of the test wave. Wire markers were made which attached to each wavemaker mechanism and extended outside the tank to the schlieren viewing area (these are the horizontal lines visible in Figs. 5.2.1 - 5.2.3). The markers moved up and down with the wavemakers, giving a visual indication of the phase of the test and background waves in the video recording. I digitised sections of suitable video footage using the FusionRecorder package on a Macintosh computer. Each recording consisted of a complete cycle of the background wave at one-quarter scale resolution, digitised in colour at 10 frames per second with JPEG compression.

I selected frames to subtract using the phase markers as a guide, coupled with the timer in the video recordings (accurate to 0.1 s) and precise measurements of the wavemaker's period. These frames were copied into Adobe Photoshop to subtract the images using the “difference” function. This returned an image in which the red, green and blue intensities of each pixel were the absolute values of the differences between these primary colours in the corresponding pixels of the original images. Thus the subtraction produced dark areas where there was little difference between the two images, and brightly coloured or white areas where the test wave produced some change in the schlieren image.

About a dozen difference images were produced, using a number of different wave frequencies. Fig. 5.3.1 is a typical example showing two original images and their difference. The effect of the test wave can be seen as the bright regions at the right and the top of the difference image. The rest of the image is not completely dark, due to slight misalignments between the images. These have many origins: air currents in the room (which produced visible fluctuations in the schlieren image), jitter and noise in the video recording, artefacts of the image compression process and of course slight temporal mismatching of images due to the finite sampling frequency of the frames. Although the difference image clearly shows the presence of the test wave and its rough location, it doesn't show any crucial details (such as wavefronts) which are needed to determine the test wave's frequency and wavelength. None of the difference images produced gave any such indications of the behaviour of the test wave in the background field.

Part of the problem was that discrete coloured stripes were used in the schlieren system. The test wave shifted the positions of the background wave's colour bands in the schlieren image, but this shift was not detectable unless it changed the colour of a pixel. Thus the difference image only shows the effect of the test wave at the *edges* of the colour bands produced by the background wave.



Background waves only

Background and test waves

Difference (enhanced)

Fig. 5.3.1: Digital image subtraction.

To avoid this problem I tried subtracting images produced using a linear greyscale slide in the schlieren system to produce a continuously graded image. The greyscale on the slide needed to be quite broad to so that the refraction due to the background wave would not go beyond it - as a result, the test wave produced only small changes in the image intensity. Although these could be seen by eye, they were undetectable with the 8-bit resolution of the greyscale digitising system, and the difference image was uniformly black. A colour “rainbow” slide could have been used instead of the greyscale, but this would have resulted in difference images whose sensitivity depended on the background wave phase (although not so strongly as for discrete colours).

It would have been nice to have a program which could subtract corresponding frames of two digitised sections of video footage to produce a movie of the difference, which may have revealed the temporal behaviour of the test wave and perhaps given an indication of whether refraction was occurring. As it was, I had to do subtractions on a frame-by-frame basis, which I quickly found to be an inordinately laborious way to construct an animation since each subtraction took around 10 minutes.

5.4 Point probe results

Once I had concluded that the schlieren system would not enable the detection of refracted test rays even with digital image-processing techniques, a new detection approach needed to be found. The schlieren method was defeated by the cluttered wavefield, and it was hoped that this problem would be overcome by looking at the frequency spectrum of the system, as only the two incident wave frequencies were expected in addition to that of a refracted ray. This spectral analysis was achieved by using the LabVIEW program I wrote to acquire and analyse data from point measurement probes, first a conductivity probe and later the polarimeter.

Unfortunately, point probes have several intrinsic drawbacks. By definition, a single point probe can only detect fluctuations at a point in the two-dimensional flow and so cannot be used to determine wavelengths[6] (as frequency and wavelength are not related). A more serious limitation of point probes is that they must be placed at the right location to detect a wave of interest, since wave energy is strongly confined to narrow beams. Thus the position of a refracted wave needs to be quite well known before it can be detected. Unfortunately I didn't have any theory which could be directly applied to predict refracted ray paths, so there was a certain amount of guesswork involved in positioning the probe.

The procedure used with both the conductivity probe and the polarimeter was to position the probe at a location where the refracted ray was likely to appear and where the background wavefield was relatively weak. Then four power spectra were collected: the background noise in the absence of waves, the test wave alone, both

---

[6] Wavelengths can be determined by using correlations between two or more spatially separated probes, but this was too complicated (because of the narrow envelopes relative to the wavelengths) for a project of this scope. It was also largely unnecessary since wavelengths of at least the incident waves could be determined from the schlieren images.

the test and background waves, and finally the background wave alone. A nonlinear refraction of the test wave was expected to result in a peak in the spectrum obtained with both waves present which did not appear in the spectra of either wave separately and could not be explained by the weak interaction theory of section 2.4.

Power spectra were usually obtained for the frequency range from DC to 0.5 Hz, as the buoyancy frequency was usually around 0.25 Hz so most wave energy should have fallen in this band. A typical spectrum of the background waves is shown in Fig. 5.4.1. The most remarkable result at first was that both the test and background waves appeared to have higher harmonics - up to the 6th harmonic in the case of the background wave. As the figure shows, most of these harmonics lay *above* the buoyancy frequency, a situation which is forbidden under linear theory. At first it appeared that this was simply an artefact of the detection method used - any nonlinearity in the signal processing from the tip of the conductivity probe to the data acquisition board of the PC would have resulted in higher harmonics by producing a non-sinusoidal periodic signal in response to the (presumed sinusoidal) conductivity variations at the tip of the probe.

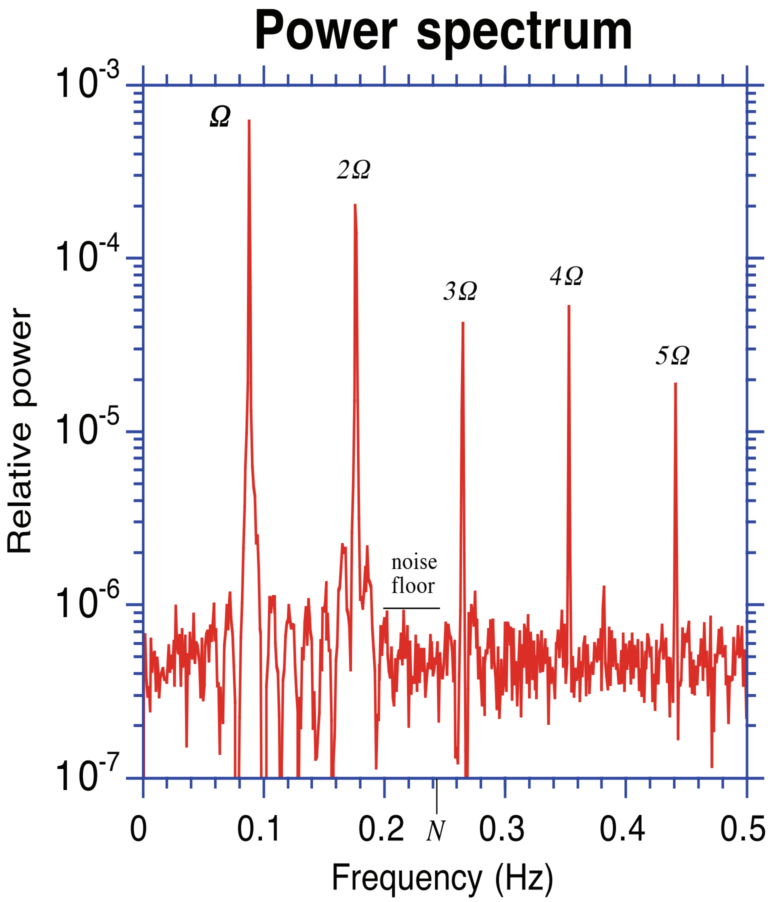


Fig. 5.4.1: A typical background wave power spectrum.

The conductivity probe was carefully calibrated using the data acquisition system, so nonlinearities in the entire measurement apparatus could be investigated. The nonlinearity of the calibration curve (see Fig. 4.5.3) was insufficient to account for the observed spectral content, due to the very small density fluctuations of the waves.[7]

The probe produced a wake visible in the schlieren system which washed back and forth past the probe tip in the oscillating velocity field of a wave, so the probe was measuring the conductivity variation of its disturbed wake rather than that of the wave itself. This may have explained some of the higher harmonic generation since the effect was stronger when a sheath was put around the probe tip to increase the disturbance of the wake.

Thus I decided to try some sort of non-invasive technique for detecting the waves which would not disturb the flow in this way. The first attempt was to build a quantitative schlieren system (with the help of Mr. Derek Corrigan), using a photodiode to detect intensity variations of a black and white schlieren image formed about 1.2 m behind the focus of the second mirror (using a linear greyscale slide). Although the circuit displayed sufficient sensitivity to detect the waves, the refraction in the wavefield was so strong that a shadowgraph image was formed in addition the schlieren image. This gave the photodiode a strongly nonlinear response to the waves, with sharp peaks corresponding to the passage of a bright focused region. As this nonlinearity could not be calibrated out, the quantitative schlieren approach was abandoned.

The next approach was to use polarimetry. The 9th tank fill was stratified with sugar[8] rather than salt so that the optical rotation would increase linearly with depth. After painstaking alignment, the polarimeter was sensitive to rotations as small as 0.03°, which was sufficient to detect the test waves. However when spectra were obtained, they displayed higher harmonics in much the same way as the waves detected by the conductivity probe. Although some slight nonlinearity in the polarimeter response was expected due to refraction of the laser beam by the waves (so it integrated the optical rotation over slightly different light paths at different wave phases), this was insufficient to explain the strength of the higher harmonics.

Thus I concluded that the higher harmonics were real features of the internal wavefield. The existence of such anharmonic waves in an *exponentially* stratified fluid was predicted and experimentally verified by Kistovich *et al.* (1990). Here is a simplified overview of their theory: When particle displacements in an exponentially stratified fluid are small, the nonlinearity of the density gradient is insignificant and a displaced particle experiences a Hookean restoring force, giving oscillations with a sinusoidal time dependence. When particle displacements are large enough to “see” that the density gradient is nonlinear, the restoring force is no longer Hookean and this results in non-sinusoidal temporal waveforms. In this case, the anharmonic waves form beams confined to propagate at an angle to the vertical determined by

---

[7] Also, the quadratic dependence on density could produce nothing higher than the second harmonic.

[8] As the viscosity of sugar solutions increases rapidly with concentration for relative densities above ~1.15 (Lide, 1991) and thus increases the minimum wavelength, the density increase was reduced to 10% over the depth of the tank. The schlieren sensitivity was not greatly affected, since refractive index increases more rapidly with density for a sugar solution than for a salt solution (see Fig. 4.2.1). The kinematic viscosity at the depth of the test wavemaker was approximately doubled, but the increase in wavelength was only 25% because of the cube-root dependence of (2.3.1).

the *fundamental* frequency ω they contain, according to the same dispersion relation as for linear waves. The higher harmonics (integer multiples of the fundamental frequency ω) are all packed into the same beam, rather than forming spatially separated modes as predicted by linear theory. Although the fundamental frequency is limited by N, the higher harmonics are not band-limited, since they simply express the fact that "natural" large-amplitude oscillations have a non-sinusoidal time dependence in a nonlinear density gradient.[9]

The results of Kistovich *et al.* (1990) applied in the case of an exponential stratification (giving a constant N). However, the stratifications I used were close to linear by virtue of the filling technique used and the effect of diffusion. A perfectly linear gradient gives a Hookean buoyant restoring force regardless of the displacement magnitude, resulting in a monochromatic frequency spectrum (Kistovich *et al.*, 1990). However, the density gradients in my experiments were not perfectly linear, even away from the diffused regions at the top and bottom of the tank (see Appendix 2). Apart from minor kinks in the gradient due to mixing from the wavemaker and other disturbances, there was a strong decrease in the gradient at the height of the background wavemaker, caused by the effect of its large displacement on the filling process. Due to its large vertical size (around 10 cm), this defect in the gradient was not removed effectively by diffusion. Thus this region had a significant and essentially permanent nonlinearity in the density profile, which could explain the anharmonic waves detected in this region, and the fact that the degree of anharmonicity varied with position in the tank. Furthermore, since I was observing the *density* fluctuations produced by the waves, even a simple-harmonic displacement field would produce non-sinusoidal variations in density in a region with a nonlinear gradient.

Unfortunately, the disturbed gradient at the depth of the large-amplitude wavemaker was the ideal location for detection of refracted test rays, as (by definition) it lay in the relatively calm region between the up- and down-going background beams. Thus it was unfeasible to move the detection to a more linear part of the gradient. If the density profile of the tank was known, it could have been possible to calibrate out some of the effect of the nonlinear gradient. Due to awkward alignment problems, this would have been impractical to attempt by traversing the polarimeter. The measurement would have been much easier with the conductivity probe, but the tank would need to be refilled with salt water, which would have meant recollecting all the data to make use of the calibration.[10]

When both wavemakers were running, the spectrum became more cluttered by weakly nonlinear interactions between the two waves. The nonlinear response of

---

[9] It is important to distinguish this phenomenon from the generation of spatially separated second-harmonic beams by the source (as discussed in Appendix 4). The anharmonicity is manifested only through the anharmonic particle motion, and not through the formation of spatially separated beams - the dispersion relation is modified and only the *fundamental* frequency determines the wave angle. Thus both the "fundamental" and "second harmonic" wave beams generated by a source can actually be anharmonic waves.

[10] Incidentally, filling the tank with a gradient produced with both sugar *and* salt (so conductivity and polarimetry could be used together) would have produced further complications, due to the double-diffusive convection which can arise as a result of the different diffusivities of these solutes (Turner, 1973).

the fluid resulted in frequencies which were the sum or difference of the test wave frequency (or frequencies) and a harmonic of the background wave. These were non-resonant triads, and represented forced oscillations rather than travelling waves (see section 2.4). As shown in Fig. 5.4.2, this effect contributed a large number of additional peaks to the spectra, which could all mask a peak due to a refracted ray[11]. Thus I couldn't escape from clutter by looking at the frequency spectrum as I had hoped.

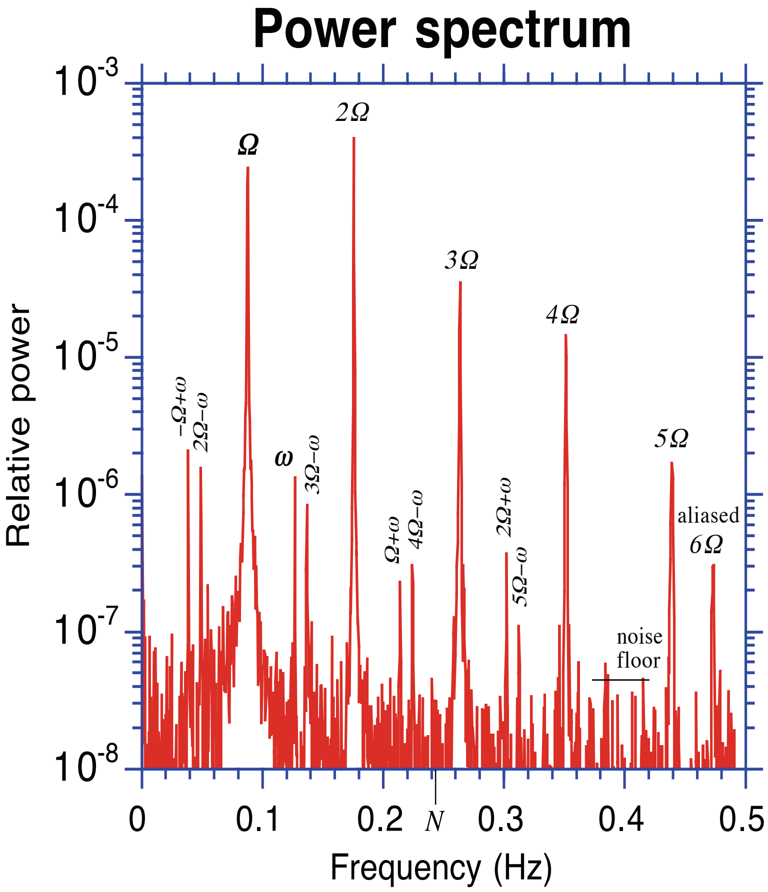


Fig. 5.4.2: Typical power spectrum with test and background waves.

Given that I had no reliable theoretical predictions of the frequency of a refracted ray, I could not be sure that its peak would be unobscured by the other peaks (or even that I had the detector in the right place). Thus I tried a number of different frequency combinations in order to shift the peaks slightly and alter the angle at which the rays intersected, which in turn effectively altered the position of the detector relative to the intersection region. The range of possible frequencies was limited by damping and other geometrical constraints, as well the requirement that G be kept small.

To obtain sufficient spectral resolution to discriminate the peaks, at least 1024 samples were required, giving 512 distinct frequency bands in the one-sided spectrum (the number of samples was a power of two so that a fast Fourier transform

---

[11] The second harmonic appears more powerful than the fundamental in this figure because the detector was closer to the second-harmonic beam than the primary beam.

could be used). To obtain a spectrum from 0 to 0.5 Hz required a sampling rate of 1 Hz, resulting in a collection time of at least 1024 s (about 17 min). Better resolution was obtained by taking 2048 samples, but this took 34 min. Given that four spectra needed to be obtained for each run, this placed a severe time constraint on the number of experiments which could be attempted.

Of the 16 different experimental runs, covering a range of parameters, there were no spectral peaks observed which could not be explained as either harmonics of the incident waves, or sum or difference frequencies due to weak non-resonant interactions.

Chapter 6: Discussion

6.1 Why it didn't work

As explained in the previous chapter, I found no evidence for refraction of the test rays in any of my experiments. Although the results were not conclusive due to the extreme difficulties encountered in observing the wavefield, there are several good reasons to tentatively conclude that such refraction did not occur.

Firstly, the test and background waves were very similar in both wavelength and frequency: the wavelengths differed by a factor of 1.75, and the frequencies typically differed by only 50%. In such circumstances the validity of theory based on the ray equations is very dubious - the spatial and temporal variations of the medium cannot be ignored on the scale of the test wavelength and period, especially when the variations are large as in my experiments.[12]

Fig. 6.1.1 shows how the frequency and wavelength problems are coupled. This figure is similar to Fig. 2.6.1, but in order to clarify the difficulty in separating the frequencies, $G\kappa/M'$ is plotted as a function of the ratio Ω/ω of the wave frequencies for different values of Ω/N. These curves intersect the horizontal axis when $\Omega/\omega=\Omega/N$ (implying $\omega=N$) and when $\Omega/\omega=1$. A small value of Ω/ω is required in order to use ray theory, and the figure shows that this implies a large value of $G\kappa/M'$ when ω is not approaching N. Thus to achieve a small G, κ/M' must be large, and this is exactly where the difficulty lay.

Although wavelengths shorter than the viscous minimum are nearly impossible to make, much longer wavelengths can be produced using a mechanically complex wavemaker such as that described by McEwan (1972). By the time the seriousness of the wavelength problem was realised it was too late to begin construction of such a mechanism, especially since a larger tank would have been needed to accommodate the necessary damping (damping thickness needs to be comparable to a wavelength to be effective).

A second reason to doubt the applicability of the theory is that it ignores the density variations produced by the background wavefield, which according to Thorpe (1989) are a much stronger source of refraction than the background wave shear under my experimental conditions. Thus the dynamics of the interaction between the two waves may have been completely different from that described by the theory I used. Viscosity was ignored by Broutman and collaborators and by Thorpe (1989), and this may also have had an important role to play in the interaction since the wavelengths are at the viscous limit.

---

[12] Some of the calculations of Broutman & Young (1986) and Broutman & McIntyre (1994) had frequencies or wavelengths differing by factors as small as 2.3 and 1.5, respectively. However when the frequency difference was small a very large wavelength difference was chosen, and *vice-versa* - no calculations had *both* the wavelength and frequency so similar as they were in my experiments. In any case, the ray-theoretical results under these conditions were not checked against calculations made without making the ray approximation, so the validity of the theory in these cases was not verified.

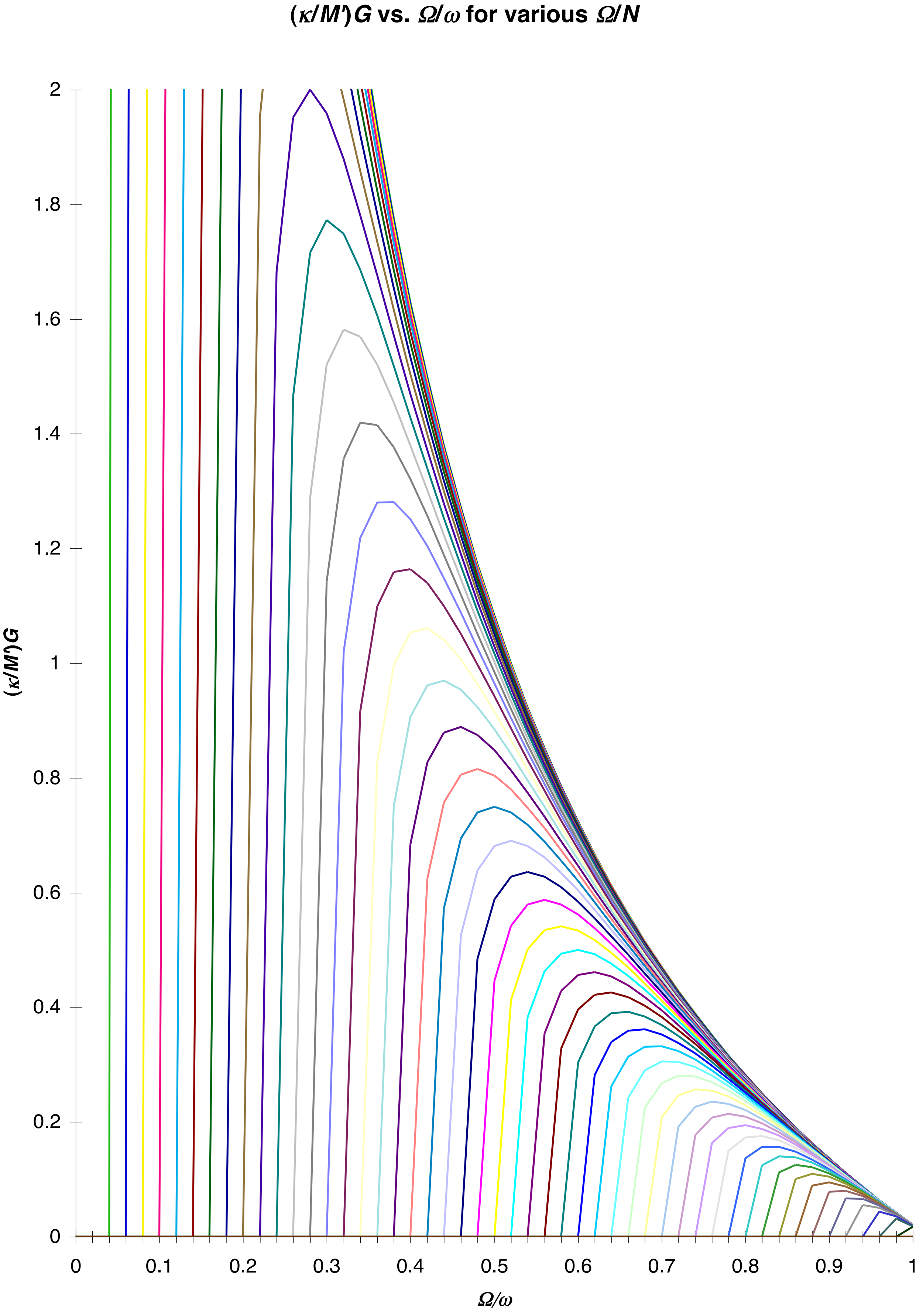


Fig. 6.1.1: Normalised G vs. Ω/ω for various Ω/N.

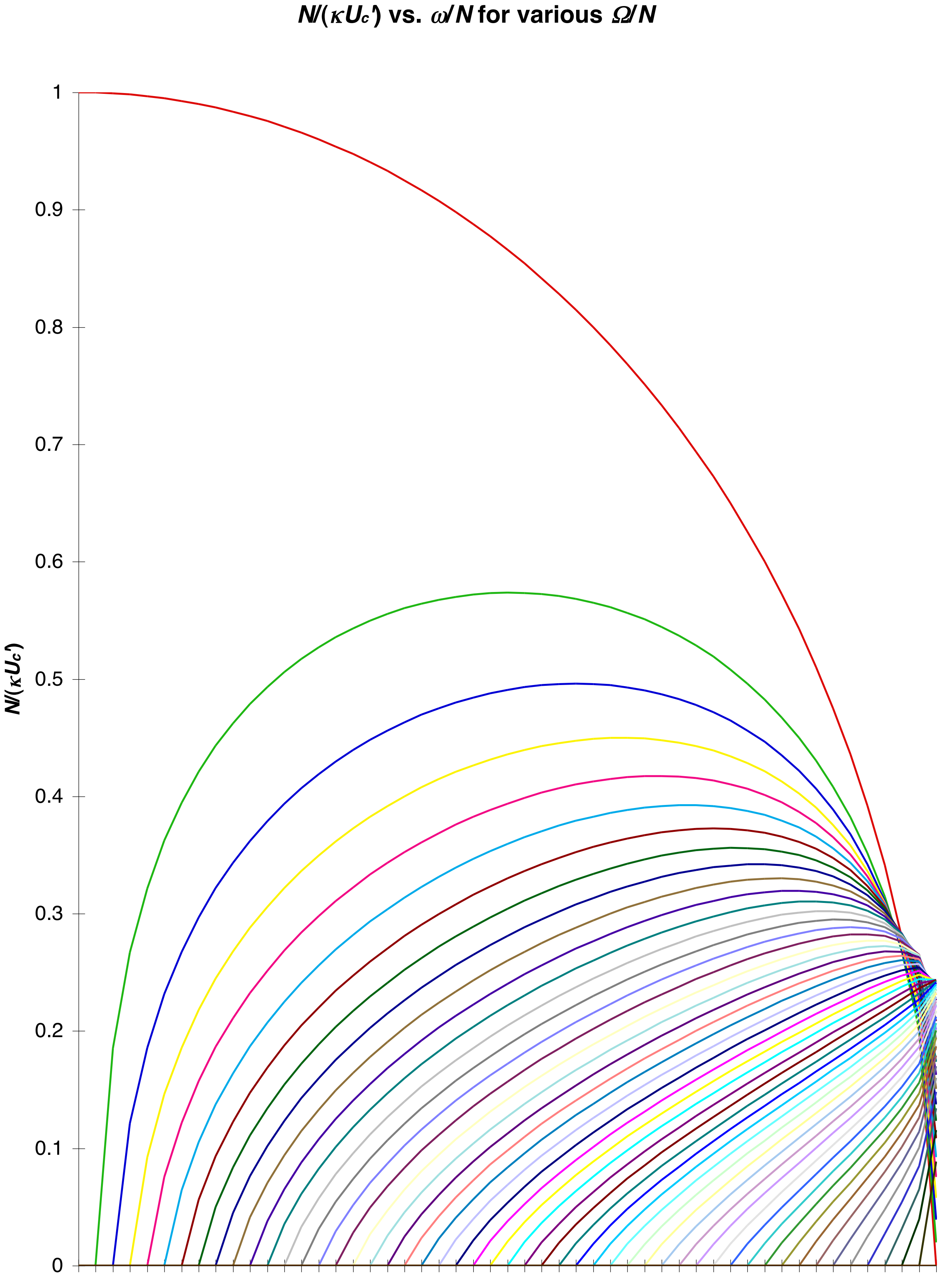


Fig. 6.1.2: Normalised $1/U_c'$ vs. ω/N for various Ω/N, with $\kappa/M'=7/4$.

Even if the theory had been applicable to my experimental circumstances, it predicted that a third-kind interaction could not occur. In my experiments I was able to achieve small values of G and m_c'/m_i' as required for a third-kind encounter, but the background wave amplitude fell well short of that required to produce caustics, which are a prerequisite for this kind of refraction. Equation (2.6.17) gives the minimum peak background wave current U_c' required to produce a caustic for given incident wave frequencies and wavelengths. In units of N/κ, U_c' varies between 1 and $+\infty$, so it is more convenient to plot its inverse as in Figs. 6.1.2 and 6.2.1. As for Figs. 2.6.1 and 2.6.2, each curve corresponds to a constant value of Ω/N given by the intersection of the curve with the horizontal axis (the widely separated uppermost curve is for $\Omega/N=0$). Fig. 6.1.2 was plotted with $\kappa/M'=7/4$ as in my experiments.[13] For the typical values $\omega/N=0.5$ and $\Omega/N=0.35$, this gives $N/(\kappa U_c')\approx 0.14$. With $N=1.6$ rad/s and $\kappa=2\pi/4$ cm^{-1}, this implies a minimum peak current of 7.3 cm/s, well in excess of the 0.55 cm/s produced by the background waves in my experiments.

The wave flow velocity could be increased by making the wavemaker longer in relation to its height and increasing the stroke. Velocity was also proportional to frequency, but this was limited by the requirement that the downgoing background wave beam remain above the test wavemaker. Unfortunately, none of these parameters could be optimised more than they already were without greatly increasing the mixing at the wavemaker, which produced intrusions that spread across the tank and interfered with the wave detection. At a frequency of $\Omega=0.56$ rad/s (ie. with $N=1.6$ rad/s and $\Omega/N=0.35$ as above), particle displacements of ±13 cm are needed to produce the required peak velocity. The required 26 cm wide wavemaker would be much too large for the 60 cm tank width, and would also have produced large distortions to the gradient in the filling process.

Furthermore, this flow velocity is extremely difficult to achieve at the background wavelength without producing substantial mixing, due to the large shears involved. Indeed, waves of such strength are likely to have been intrinsically unstable. The *Richardson number*, *Ri*, is a dimensionless measure of the potential for instability in a sheared, stratified fluid. It is defined by

$$Ri = \frac{N^2}{\left(\partial u/\partial z\right)^2} \qquad (6.1.1)$$

(where u is the flow speed) and it can be shown that a stratified fluid is stable when $Ri>1/4$ at all points in the flow (Apel, 1987). Thus a necessary (but not sufficient) condition for instability is that $Ri<1/4$ at some point. The minimum Richardson number for low-frequency background waves is approximately

$$Ri_{\min} = \left(\frac{N}{M'U'_{\max}}\right)^2 \qquad (6.1.2),$$

which gives $Ri_{\min}\approx 10$ for the background waves in my experiment, indicating their stability. If these waves had the peak flow speed required to produce caustics, (6.1.2) gives $Ri_{\min}=0.06$, showing that these strong waves are very likely to be unstable.

---

[13] the plotted curves are inaccurate when ω differs greatly from Ω, since the approximation (2.6.10) begins to break down.

The required value of U_c' could be made comparable to the experimental values by reducing N by a factor of 13 or more. Unfortunately, this would reduce the gradient by a factor of about 180 and make the waves virtually undetectable. Also, the minimum Richardson number would not be changed, so the waves would still probably be unstable.

6.2 How it could be done (maybe)

The previous section listed several important parameter values which could not be achieved with my apparatus. This section outlines an experimental apparatus which could overcome these difficulties and satisfy the requirements of the (dubious) theory I used.

Let us start from scratch and use the required parameters to determine the experimental arrangement (rather than the other way around). Due to the lack of trustworthy theoretical predictions (as density variations are ignored), a method such as schlieren is needed to survey a large part of the wavefield and search for refracted rays. This in turn would require a region which gave a clear view of the test waves after they had interacted with the background wave, so a solution would need to be found to the cluttered wavefield which plagued my schlieren observations.

A major part of the problem was the upgoing background wave, and this could be eliminated using the wavemaker design of McEwan (1972). This consists of a series of flat plates which pivot about parallel horizontal axes as sketched in Fig. 6.2.2. They are joined by rubber sheets to created an undulating surface with zero total displacement which mimics the standing wave pattern of an intersecting pair of up- and down-going waves. The top of the wavemaker is against a horizontal reflecting surface, at which point the wavemaker motion has an *antinode*. The downgoing wave propagates freely, but the upgoing is reflected to form part of the downgoing beam. The net effect is the same as a wavemaker of twice the height (made by putting a mirror image of the wavemaker on top of it) which produces only a downgoing wave beam.

We can now choose all the wave parameters and see what requirements this places on the apparatus. I will start by choosing the test wave frequency to be $\omega=N/2$, which will eliminate second-harmonic generation from this source while still allowing ample scope for frequency increases as a result of the interaction. This frequency results in test waves which propagate at 30° to the horizontal.

I choose a background frequency of $\Omega=N/6$, which gives a reasonable frequency contrast with the background wave, and results in wavefronts inclined at about 10° to the horizontal. Also, having the test frequency an integer multiple of the background frequency results in sum and difference peaks in the frequency spectrum which overlap those of the background wave harmonics instead of cluttering the spectrum.

Choosing a moderately small value of G=0.3 should produce a weak third-kind refraction, and from Fig. 2.6.1 this implies $\kappa/M'=6$, giving a significant wavelength contrast. Using these values, Fig. 2.6.2 gives $m_c'/m_i'=0.65$, which indicates that only moderate changes in m' are possible.

Accepting the viscous limitation in the minimum wavelength implies $\kappa=2\pi/4$ cm^{-1} or less in an interaction region, so the background wavenumber is at most $M'=2\pi/24$ cm^{-1}. The waves produced by McEwan's design have a wavelength determined by the spacing of the pivoting plates, and have no upper wavelength limit imposed by viscosity as they do when generated by the simple oscillating forms used in my experiments (McEwan, 1972). Thus this mechanism is capable of producing the large wavelength required.

The background wavelength and angle determine the scale and geometry of the apparatus. The long wavelength implies a very large tank, as shown in the scale diagram Fig. 6.2.2. - note the size of my experimental tank for comparison. The test wave source is positioned so that refracted rays will not be intercepted by the background wave reflector. The upgoing test waves pass through the downgoing background wave and into the viewing area, with the left beam removed by damping (having both the test and background waves coming from the same side increases the size of the viewing area). This damping could also be positioned to remove any second-harmonic wave beam produced by the background wave source.

The wave reflector needs to be a thin sheet so that it will have a negligible effect on the gradient in the viewing area, but must still be rigid enough to effectively reflect the strong background waves without allowing their energy to pass into the area above. The viewing area is located at a depth which would be largely unaffected by the distorted gradient produced by the displacement of the background wavemaker.

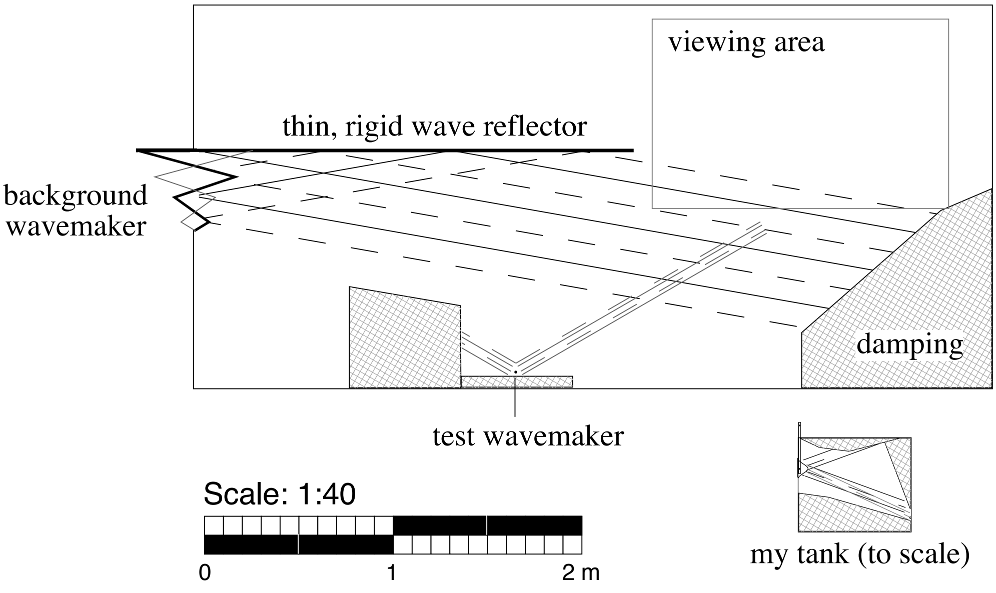


Fig. 6.2.2: Proposed wave-interaction apparatus.

The apparatus in Fig. 6.2.2 produces interacting waves which have significant differences in both wavelength and frequency, and sufficiently small values of G and m_c'/m_i'. We now need to address the question of the wave amplitude required to produce caustics in this case. Fig. 6.2.1 is the same as Fig. 6.1.2, but it was plotted for $\kappa/M'=6$. From this figure, we see that the peak background flow velocity must exceed $U_c'=N/(0.15\kappa)$.[14]

The tank in the figure has a depth of 2 m, and a linear change of density by 15% over this depth gives a buoyancy frequency N=0.84 rad/s, about half the value used in my experiments.[15] Using this value of N and the viscous-limited test wavenumber, we have U_c'=3.6 cm/s - this is less than the value required in my experiments, due mostly to the reduced value of N. For the frequency Ω=0.14 rad/s (implied by $\Omega/N=1/6$ and N=0.84 rad/s), this requires particle displacements of ±25 cm, which needs a wavemaker which moves through at least 50 cm as shown in the figure. Although these displacements are large, the shear is reduced by the long wavelength and the background wave has a minimum Richardson number of 0.8, indicating stability. It would be a technical challenge to build a wavemaker which could combine the required displacements and wavelength, however.

Thus the apparatus shown in Fig. 6.2.2 should be just capable of producing the appropriate conditions for these interactions. Due to the sheer size of the viewing area, it would be prohibitively expensive to use a schlieren system to observe the waves (mirrors of about 1 m diameter would be required), but the Moiré technique could be used instead (see section 4.3). However there would probably be some difficulty in achieving the required sensitivity, since the gradient is one-quarter of that used for my experiments.

It hardly needs to be stated that a 2 × 4 m tank would be very expensive to build and use. If it had a thickness of 0.2 m as in my experiments, it would have a volume of 1600 L; the solution alone would weigh over 1700 kg and about 160 kg of salt would be required for each fill. Despite the large dimensions, the wavelength and frequency separation between the waves may still be insufficient for the predicted interactions to occur, and an even larger tank could be needed (since reducing Ω increases the length of the wave reflector required). Before embarking on such a large-scale project it would be prudent to construct a theory which explicitly included the effects of density fluctuations (and even those of viscosity) and addressed the validity of ray theory, so that there would be good reasons for expecting the experiment to work.

Alternatively, one could test the theory of Broutman and collaborators more directly by using a rotating system so that inertial waves could be used, as was done in their analysis. Such a system would still need to be large to obtain sufficient scale separation between the viscous test ray and the inertial wave, and a rotating system

---

[14] This value is somewhat tentative, as the rays meet at a fairly steep angle (20°) so (2.6.10) is not such a good approximation.

[15] This is close to the steepest gradient possible with salt (and hence the largest value of N), as the solution saturates when the density increase reaches 20% (Lide, 1991). Sugar solutions are less useful for steep gradients due to the rapid increase in viscosity for densities greater than about 15% above water. Other solutes may enable the formation of steeper gradients than are possible with NaCl.

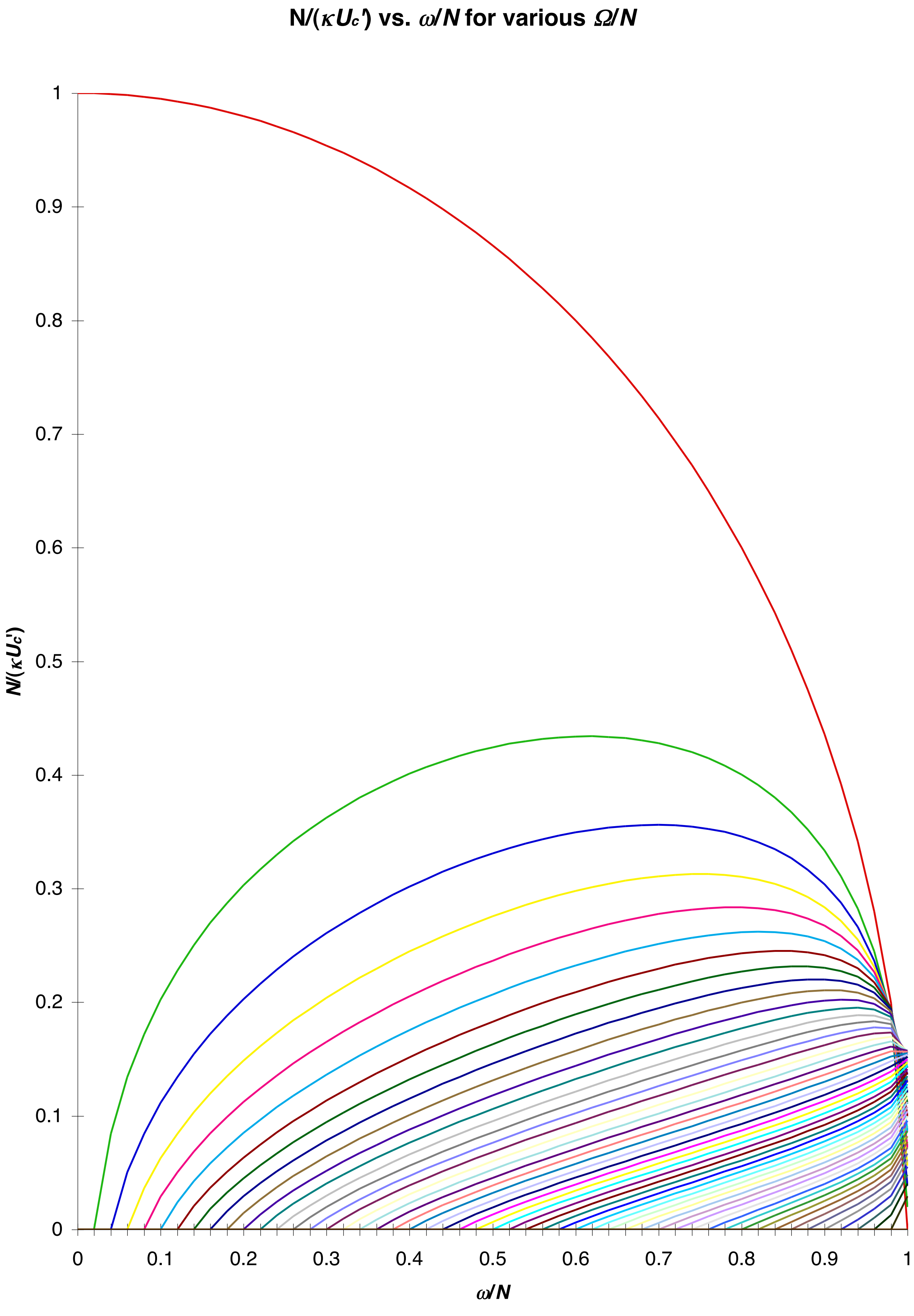


Fig. 6.2.1: Normalised $1/U_c'$ *vs.* ω/N *for various* Ω/N*, with* $\kappa/M'=6$*.*

could make observations very difficult. Apart from needing to place the detection apparatus in the rotating reference frame, a rotating tank would need to be cylindrical and this would make refractive-index techniques very difficult to implement. However, an advantage of using inertial waves is that they involve no density variations and so would not interfere with measurements of the density field.

In any case, it is clear from Fig. 6.2.2 that the tank and other apparatus used for my experiments were quite inappropriate to observe the shear-induced refraction predicted by Broutman and others.

Conclusion

Broutman and others have developed a theory based on the ray approximation to describe the strongly nonlinear interaction of a weak, short test ray with the shear field of a strong inertial wave of large wavelength. This theory was modified to replace the inertial wave by a long, large-amplitude internal gravity wave. The density perturbations due to the long wave were ignored in this adaptation.

An apparatus was developed for the generation of strong internal gravity waves in a stratified salt or sugar solution, with the objective of observing the predicted nonlinear interaction. Schlieren and polarimetric techniques were refined to improve their suitability for detection of internal gravity waves, and a data acquisition and analysis system was developed to provide Fourier analysis of the wavefield.

Waves were detected and found to obey the linearised dispersion relation despite being in the nonlinear regime. Second-harmonic generation was observed from the wave sources and plausible explanations were developed for this phenomenon. Also observed were nonlinear effects such as anharmonic waves and sum and difference frequencies due to the large wave amplitude.

However, the particular nonlinear interaction I set out to observe was not detected. Although its absence was not conclusively demonstrated due to the serious observational difficulties encountered, there were sound theoretical reasons for suspecting that the interaction did not occur. Due to inescapable limitations of the apparatus, the background wave was more than an order of magnitude too weak to generate the required shear field. Also, the wavelength and frequency separation between the waves was limited by viscosity and could not be increased to a sufficient level. There were doubts as to the validity of the theory used, since it ignored the density fluctuations of the strong wave.

An apparatus was proposed which could generate waves with appropriate frequencies, wavelengths and amplitudes for the predicted interactions, but due to its large size it would be prudent to develop a more directly applicable theory before attempting construction.

Appendix 1: Derivation of the basic properties of internal gravity waves

The derivation of the basic properties of internal gravity waves is quite standard. The following derivation is based on those given by Gill (1982), Lighthill (1978) and LeBlond & Mysak (1978).

A1.1 The equations of motion

We wish to find the modes of free oscillation under the influence of gravity of a fluid whose density is a function of position when at rest. First of all, we make the *continuum approximation* that the fluid has the same basic properties at all length scales - that is, we ignore the fact that the fluid is comprised of discrete molecules. The behaviour of this fluid will be described from the Eulerian perspective, as a function of time t and of position $\boldsymbol{r}$ relative to coordinate axes which are fixed in space relative to an inertial reference frame.

The fluid is described by the density $\rho(\boldsymbol{r},t)$, pressure $p(\boldsymbol{r},t)$ and velocity $\boldsymbol{u}(\boldsymbol{r},t)$, and is assumed to be inviscid.

The *substantive derivative* of any scalar or vector function $\phi(\boldsymbol{r},t)$ is defined as

$$\frac{D\phi}{Dt} = \frac{\partial\phi}{\partial t} + (\boldsymbol{u}\cdot\nabla)\phi \qquad \text{(A1.1).}$$

This is the time derivative of ϕ taken by following the motion of the fluid element under consideration.

Using the substantive derivative, the basic equations relating ρ, p, and $\boldsymbol{u}$ are the *continuity equation* (conservation of mass)

$$\frac{D\rho}{Dt} + \rho\nabla\cdot\boldsymbol{u} = 0 \qquad \text{(A1.2)}$$

and the inviscid *momentum equation* (conservation of momentum)

$$\rho\frac{D\boldsymbol{u}}{Dt} = -\nabla p + \rho\boldsymbol{g} \qquad \text{(A1.3),}$$

where (A1.3) assumes that the only external force is that due to the gravitational acceleration $\boldsymbol{g}$.

In order to close this set of equations relating ρ, p, and $\boldsymbol{u}$ we need a third equation, this being the *equation of state* for the fluid. Salt or sugar solutions were the fluids used in my project, which have equations of state of the form

$$\rho=\rho(p,S,T) \qquad \text{(A1.4),}$$

where S and T are the concentration of solute and the temperature, respectively.

Molecular diffusion of the solute and conduction of heat are insignificant on the time and distance scales defined by the frequency and wavelength of the waves considered in this project, so a fluid element may be assumed to preserve its values of S and T:

$$\frac{DS}{Dt} = 0 \qquad \text{(A1.5)}$$

$$\frac{DT}{Dt} = 0 \qquad \text{(A1.6).}$$

Furthermore, the density of the solution is only a weak function of pressure, and for the small pressure variations in this project this dependence can also be ignored - that is, the liquid is assumed to be *incompressible*:

$$\frac{\partial \rho}{\partial p} = 0 \qquad \text{(A1.7)}$$

In effect, this eliminates acoustic waves from consideration, as (A1.7) implies an infinite sound speed. Given (A1.5), (A1.6) and (A1.7), the density of a material fluid element remains constant:

$$\frac{D\rho}{Dt} = 0 \qquad \text{(A1.8)}$$

which combined with (A1.2) implies that the velocity field is solenoidal (divergence-free):

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{(A1.9).}$$

To summarise, the equations of motion governing this inviscid, incompressible, diffusionless fluid in an inertial frame in which gravity provides the only body force are (A1.3), (A1.8) and (A1.9).

A1.2 The hydrostatic equilibrium state

It will greatly clarify later working to introduce the *resting state* of the fluid, a state of hydrostatic equilibrium defined by $\boldsymbol{u}(\boldsymbol{r},t)=\boldsymbol{0}\ \forall \boldsymbol{r},t$. The pressure and density in the resting state are denoted p_o and ρ_o, respectively - note that these are functions of $\boldsymbol{r}$ but not of t.

In general the pressure and density are given by

$$p = p_o + p', \quad \rho = \rho_o + \rho' \qquad \text{(A1.10),}$$

which defines the perturbations p', ρ' to the resting state.

The resting state defined by $\boldsymbol{u}=\boldsymbol{0}$, $p'=\rho'=0$ satisfies (A1.8) and (A1.9) identically, whilst (A1.3) implies

$$\nabla p_o = \rho_o \boldsymbol{g} \qquad \text{(A1.11).}$$

In the non-equilibrium case, combining (A1.11) and (A1.10) with (A1.8) and (A1.3) gives

$$\frac{D\rho'}{Dt} + (\boldsymbol{u} \cdot \nabla)\rho_o = 0 \qquad \text{(A1.12)}$$

and

$$\rho \frac{D\boldsymbol{u}}{Dt} = -\nabla p' + \rho' \boldsymbol{g} \qquad \text{(A1.13),}$$

respectively. Note that (A1.11) removes the contributions of p_o and ρ_o from the right-hand side of (A1.13), showing that only the *perturbations* p' and ρ' to the equilibrium state provide any driving forces to alter the momentum.

A1.3 Linearisation

For the moment, let us consider *infinitesimal* perturbations $\boldsymbol{u}$, p' and ρ' to the static equilibrium state. In this case, any products of these small quantities can be neglected, a procedure known as *linearisation.*

Equation (A1.9) is linear in $\boldsymbol{u}$ and hence unchanged by linearisation, whilst (A1.12) becomes

$$\frac{\partial \rho'}{\partial t} + (\boldsymbol{u}\cdot\nabla)\rho_o = 0 \qquad \text{(A1.14)}$$

and the linearised form of (A1.13) is

$$\rho_o \frac{\partial \boldsymbol{u}}{\partial t} = -\nabla p' + \rho' \boldsymbol{g} \qquad \text{(A1.15).}$$

We now define right-handed Cartesian coordinate axes with unit vectors $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$ such that the z axis points vertically upwards so that $\boldsymbol{g} = -g\hat{\boldsymbol{z}}$, where $g=|\boldsymbol{g}|$. In these coordinates we write $\boldsymbol{r} = (x,y,z)$ and $\boldsymbol{u} = (u,v,w) = \boldsymbol{u_h} + w\,\hat{\boldsymbol{z}}$. Note that (A1.11) implies $p_o=p_o(z)$ and hence $\rho_o=\rho_o(z)$, since ∇p_o must also be a function of z alone.

Now (A1.9) can be written

$$\nabla_h \cdot \boldsymbol{u}_h + \frac{\partial w}{\partial z} = 0 \qquad \text{(A1.16),}$$

where $\nabla_h = (\partial/\partial x, \partial/\partial y, 0)$ is ∇ acting only on the horizontal coordinates. Similarly, (A1.14) can be written

$$\frac{\partial \rho'}{\partial t} + w\frac{\partial \rho_o}{\partial z} = 0 \qquad \text{(A1.17)}$$

and (A1.15) becomes

$$\rho_o \frac{\partial \boldsymbol{u_h}}{\partial t} = -\nabla_h p' \qquad \text{(A1.18)}$$

and

$$\rho_o \frac{\partial w}{\partial t} = -\frac{\partial p'}{\partial z} - \rho' g \qquad \text{(A1.19).}$$

Taking the time derivative of (A1.19) and substituting (A1.17) yields

$$\rho_o\left(\frac{\partial^2 w}{\partial t^2} + N^2 w\right) = -\frac{\partial^2 p'}{\partial t \partial z} \qquad \text{(A1.20),}$$

where the *buoyancy frequency* $N=N(z)$ is defined by

$$N^2 = \frac{1}{\rho_o}\boldsymbol{g}\cdot\nabla\rho_o = \frac{-g}{\rho_o}\frac{\partial \rho_o}{\partial z} \qquad \text{(A1.21).}$$

The buoyancy frequency is a parameter of fundamental importance in a density-stratified fluid, as will become clear shortly.

Taking the time derivative of (A1.16) allows $\boldsymbol{u}_h$ to be eliminated via (A1.18) to give

$$\rho_o \frac{\partial^2 w}{\partial t\, \partial z} = \nabla_h^2 p' \qquad \text{(A1.22).}$$

Taking the horizontal Laplacian of (A1.20) and differentiating (A1.22) with respect to both t and z allows us to eliminate p', yielding an expression for w alone:

$$\nabla^2 \frac{\partial^2 w}{\partial t^2} + N^2 \nabla_h^2 w - \frac{N^2}{g} \frac{\partial^3 w}{\partial z \partial t^2} = 0 \qquad \text{(A1.23).}$$

A1.4 The dispersion relation

If we assume N to be constant, (A1.23) is a fourth order partial differential equation for w with constant coefficients. We will look for plane wave solutions to this equation, ie. solutions of the form

$$w = W \exp\left[i(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t)\right] \qquad \text{(A1.24),}$$

where W is some constant, $\boldsymbol{k}=(k,l,m)$ is the wavevector and ω the frequency of the waves. We use the usual convention whereby the actual physical quantity is the real part of (A1.24).

Substitution of (A1.24) into (A1.23) yields the following dispersion relation linking ω to $\boldsymbol{k}$:

$$k^2 + l^2 + m^2 - \frac{N^2}{\omega^2}\left(k^2 + l^2\right) + im\frac{N^2}{g} = 0 \qquad \text{(A1.25).}$$

The factor i in the third term of (A1.25) indicates that at least one of the components of the wavevector must be complex. The exponential growth or decay associated with the imaginary component is in fact only an apparent problem. The condition that N is constant implies

$$\rho_o \propto \exp\left(\frac{-N^2 z}{g}\right) \qquad \text{(A1.26).}$$

If we assume equipartition between kinetic and potential energy, the total wave energy density is

$$E = \rho_o |\boldsymbol{u}|^2 \qquad \text{(A1.27),}$$

so for E to remain constant, we must have

$$|\boldsymbol{u}| \propto \exp\left(\frac{N^2 z}{2g}\right) \qquad \text{(A1.28).}$$

Thus we would expect m to be complex, the imaginary part giving the exponential behaviour of (A1.28). Substituting $m=m_R + im_I$ (where m_R and m_I are both real) into (A1.25) and equating real and imaginary parts gives

$$m_I = \frac{-N^2}{2g} \qquad \text{(A1.29),}$$

as expected, as well as the real dispersion relation

$$\omega^2 = \frac{N^2\left(k^2 + l^2\right)}{\kappa^2 + \dfrac{N^4}{4g^2}} \qquad \text{(A1.30)},$$

where κ is the magnitude of Re($\boldsymbol{k}$). To simplify notation, we will consider only the real part of m from now on, and write $m=m_R$ and $\boldsymbol{k}=(k,l,m_R)$. Note that (A1.30) implies that $\omega^2 \le N^2$, that is, the internal wave spectrum is *band limited* in frequency.

For short waves, we have $\kappa >> N^2/2g$ and (A1.30) simplifies to

$$\omega = N\cos\theta \qquad \text{(A1.31)},$$

where θ is the angle of $\boldsymbol{k}$ to the horizontal (see Fig. 2.1.1). Note that in this short wave case, (A1.31) implies that the wavelength $\lambda=2\pi/|\boldsymbol{k}|$ is *independent* of ω - a disturbance at a given frequency can produce waves of *any* wavelength, provided they are short enough to satisfy (A1.30).

A1.5 The nature of the wavefield

We now derive the form of the other components $\boldsymbol{u}_h$, p' and ρ' of the wavefield, given that w is given by (A1.24) and N is constant. Substituting (A1.24) into (A1.22) and (A1.17) gives

$$p' = \frac{-\omega m \rho_o}{k^2 + l^2} w \qquad \text{(A1.32)}$$

and

$$\rho' = \frac{i\rho_o N^2}{\omega g} w \qquad \text{(A1.33)},$$

respectively. Note that the factor of i in (A1.33) indicates a $+\pi/2$ phase shift in ρ' relative to p' and w. Now $\boldsymbol{u}_h$ can be found using (A1.32) and (A1.18):

$$\boldsymbol{u}_h = -(k,l)\frac{m}{k^2 + l^2} w \qquad \text{(A1.34)},$$

which when combined with (A1.24) gives the complete velocity field:

$$\boldsymbol{u} = \left(\frac{-km}{k^2 + l^2}, \frac{-lm}{k^2 + l^2}, 1\right) W \exp\left[i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)\right] \qquad \text{(A1.35)}.$$

It is clear from (A1.35) that

$$\boldsymbol{u}\cdot\boldsymbol{k} = 0 \qquad \text{(A1.36)},$$

that is, plane internal gravity waves are *transverse*.

The *phase velocity* is defined as

$$\boldsymbol{c} = \frac{\omega \boldsymbol{k}}{\kappa^2} \qquad \text{(A1.37)},$$

whilst the *group velocity* is defined by $\boldsymbol{c}_g = \nabla_k \omega$, where $\nabla_k = \left(\partial/\partial k, \partial/\partial l, \partial/\partial m\right)$ is ∇ acting in wavenumber space. For short waves, substituting (A1.30) into (A1.38) we find (after some tedious algebra) that the group velocity is

$$\boldsymbol{c}_g = \frac{Nm\sqrt{k^2 + l^2}}{\kappa^3}\left(\frac{km}{k^2 + l^2}, \frac{lm}{k^2 + l^2}, -1\right) \qquad \text{(A1.38).}$$

By comparison with (A1.35) it is clear that $\boldsymbol{c}_g$ is parallel to $\boldsymbol{u}$ and hence

$$\boldsymbol{c}_g \cdot \boldsymbol{c} = 0 \qquad \text{(A1.39),}$$

showing that the group and phase velocities are *perpendicular*, so that the wave energy propagates *along* the wavefronts. The implications of this peculiar behaviour will be elaborated in the main text.

In terms of the wavelength $\lambda=2\pi/|\boldsymbol{k}|$, the magnitudes of the phase and group velocities are

$$|\boldsymbol{c}| = \frac{\lambda\omega}{2\pi} = \frac{\lambda N \cos\theta}{2\pi} \qquad \text{(A1.40)}$$

and

$$\left|\boldsymbol{c}_g\right| = \frac{\lambda N \sin\theta}{2\pi} \qquad \text{(A1.41).}$$

Thus both velocities are proportional to wavelength, but $|\boldsymbol{c}|$ increases with increasing frequency, whilst $|\boldsymbol{c}_g|$ decreases.

Appendix 2: The effects of diffusion

The solute concentration in the tank becomes increasingly uniform as a result of diffusion. In the absence of waves, the density (and hence concentration) must depend spatially only on the vertical coordinate z, as horizontal density variations are destroyed by the unbalanced pressure gradients they induce. Thus the diffusion can be described by the one-dimensional diffusion equation

$$\frac{\partial S}{\partial t} = D\frac{\partial^2 S}{\partial z^2} \qquad \text{(A2.1)},$$

where we ignore the weak dependence of the diffusivity D on the solute concentration $S=S(z,t)$.

It is clear from this equation that a linear variation of S with z will be stable, as the time derivative is zero. However, in a finite tank we have boundary conditions of zero flux of solute through the top and bottom of the tank, forcing the density *gradient* to be zero there. Thus a nonzero gradient in the body of the tank cannot be stable, as $\partial^2 S/\partial z^2$ must be nonzero at at least two depths. The time derivative is largest where the gradient changes rapidly with depth, thus quickly smoothing out small-scale nonlinearities in the density profile. Nonlinearities over large scales (such as between the top of the tank and the linear gradient at mid-depth) take much longer to be affected. Thus in the short term diffusion acts to produce a constant gradient away from the top and bottom boundaries, which is utilised to smooth out the gradient after a fill. In the longer term the influence of the boundary conditions spreads out into the body of the tank, gradually reducing the extent of the linear density variation.

It was of great importance for the feasibility of this project that the gradient remained essentially constant through most of the tank for a reasonable number of days, to allow experiments to be conducted. The general solution to (A2.1) subject to the boundary conditions of zero flux through the planes $z=0$ and $z=d$ is easily shown to be

$$S(z,t) = \sum_{j=0}^{\infty} A_j \exp\left(-\left(\frac{j\pi}{d}\right)^2 Dt\right)\cos\left(\frac{j\pi}{d}z\right) \qquad \text{(A2.2)},$$

where the coefficients A_j are determined by Fourier analysis from the initial conditions. Note that small-scale nonlinearities in the gradient (ie. terms with large j) decay very much more rapidly than larger scale features. If at $t=0$ the concentration varies linearly between S_{max} at $z=0$ and S_{min} at $z=d$, the coefficients are

$$A_0 = \frac{S_{\max} + S_{\min}}{2},$$
$$A_{j>0} = \begin{cases} \dfrac{4(S_{\max} - S_{\min})}{j^2\pi^2} & \text{for odd } j, \\ 0 & \text{for even } j \end{cases} \qquad \text{(A2.3).}$$

This solution and the variation it implies for N (see equation (2.1.1)) were plotted for a number of different values of t in order to plan where the wavemakers needed to be placed to ensure that they could generate travelling waves at the frequency required. An example is shown in Figs. A2.1, A2.2, which were calculated for a typical gradient made with salt and a depth d=50 cm. These calculations were made assuming that density is proportional to salt concentration - for my purposes, the error introduced by ignoring the slight nonlinearities in this dependence is negligible (see Ruddick and Shirtcliffe, 1979). In these figures, note that the gradient relaxes very rapidly to begin with, then changes much more slowly once $\partial^2 S/\partial z^2$ has been reduced, so the gradient through most of the tank remains useable for over a week. Note also that N is not constant for a constant density gradient, as it also depends on ρ_o. The Boussinesq approximation ignores this dependence, and gives a constant N at t=0, as shown in Fig. A2.2.

The values of D are 1.484×10^{-9} m^2/s for salt and 0.521×10^{-9} m^2/s for sugar (Lide, 1991), so a gradient made with sugar will diffuse around three times more slowly than the salt gradient shown in the figure. It is fortunate that D is sufficiently small for the overall gradient to have a lifetime of a week or two, but sufficiently large for small-scale disturbances to diffuse out in a matter of days. It is clear that the diffusion is negligible on the timescale of a wave period (of the order of one minute), as was assumed in Chapter 2.

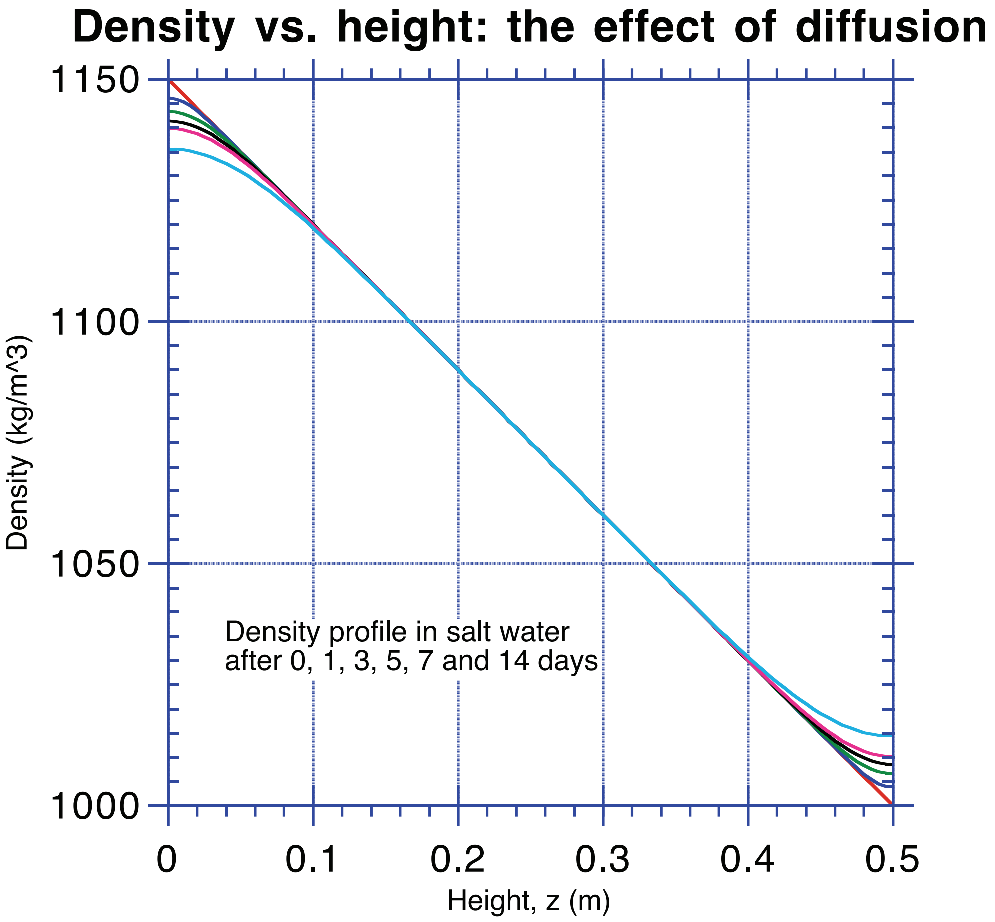


Fig. A2.1: The effect of diffusion on a density gradient in salt water.

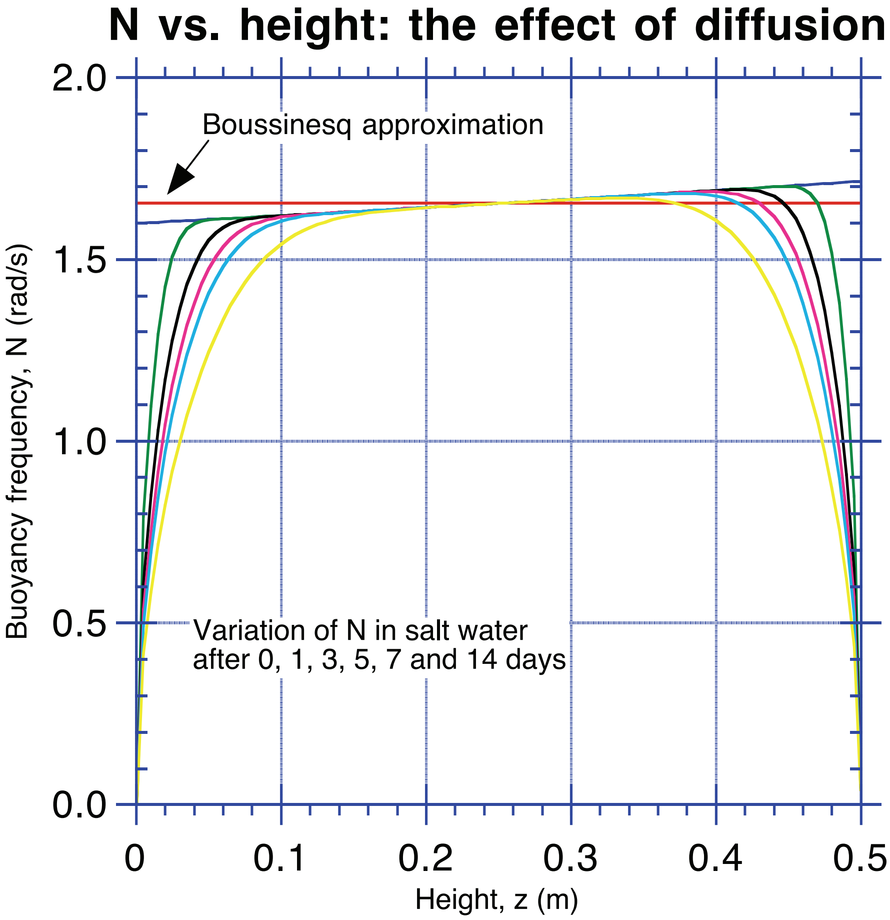
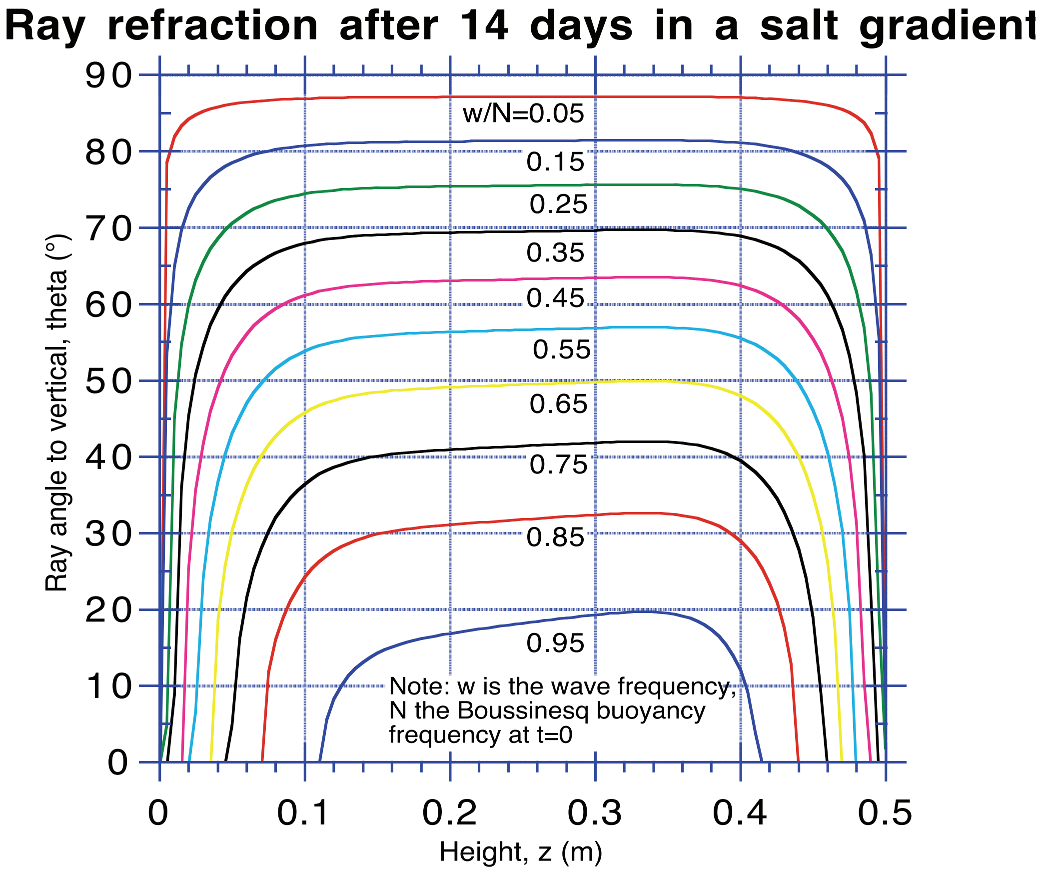


Fig. A2.2: The effect of diffusion on the buoyancy frequency N.

The drop-off in N at the top and bottom of the tank refracts rays as described by the ray theory in section 2.2. This refraction is shown in Fig. A2.3 for a number of frequencies. Note that the higher frequencies are more severely refracted and more confined than the lower frequencies, and that there is also weak refraction through the constant-gradient region, since this differs from the exponential density variation (A1.26) required for a constant N. Overall, for frequencies not too large compared with N, the refraction is very slight away from the top and bottom boundaries.



Ray refraction after 14 days in a salt gradient

Note: w is the wave frequency, N the Boussinesq buoyancy frequency at t=0

Fig. A2.3: Ray refraction in a diffused salinity gradient.

Appendix 3: LabVIEW data acquisition program

This appendix presents the data acquisition program I wrote to collect and analyse the analog signals produced by the polarimeter and the conductivity bridge. The package used was LabVIEW version 3 for Windows, by National Instruments. This provides a very comprehensive environment for the construction of “virtual instruments” or VI's (in fact, “VIEW” stands for “Virtual Instrument Engineering Workshop”). The National Instruments data acquisition board in the PC could be controlled by the VI software. The programming ‘language’ used is intuitive but rather unusual and requires some explanation.

The programming environment consists of a vast array of modules (called “sub-VI's”) analogous to procedures or subroutines, whose functions range from simple arithmetic to data acquisition, analysis, hardware control, file handling and other complex functions. Standard programming constructs such as loops, decision structures and local variables are also provided. The distinguishing feature of LabVIEW programming is that it is entirely graphical: all sub-VI's and programming constructs are represented by icons, and lines drawn between the icons represent exchange of information between these objects (analogous to passing parameters between procedures in Pascal). Thus programs look like circuit diagrams showing signal connections between components.

Illustrations of this can be found on the next few pages, after the “front panel”. Program constructs such as loops and case statements are represented by frames surrounding parts of the diagram - some have been labelled to make the structure more comprehensible. Printouts are shown displaying the contents of all the IF constructs in either state, as well as both steps (0 and 1) in the sequence construct. The wiring convention used is that sub-VI's accept inputs at their left and produce outputs at the right. “Wires” carrying different data types (such as integers, arrays or strings) are shown with a different pattern or thickness.

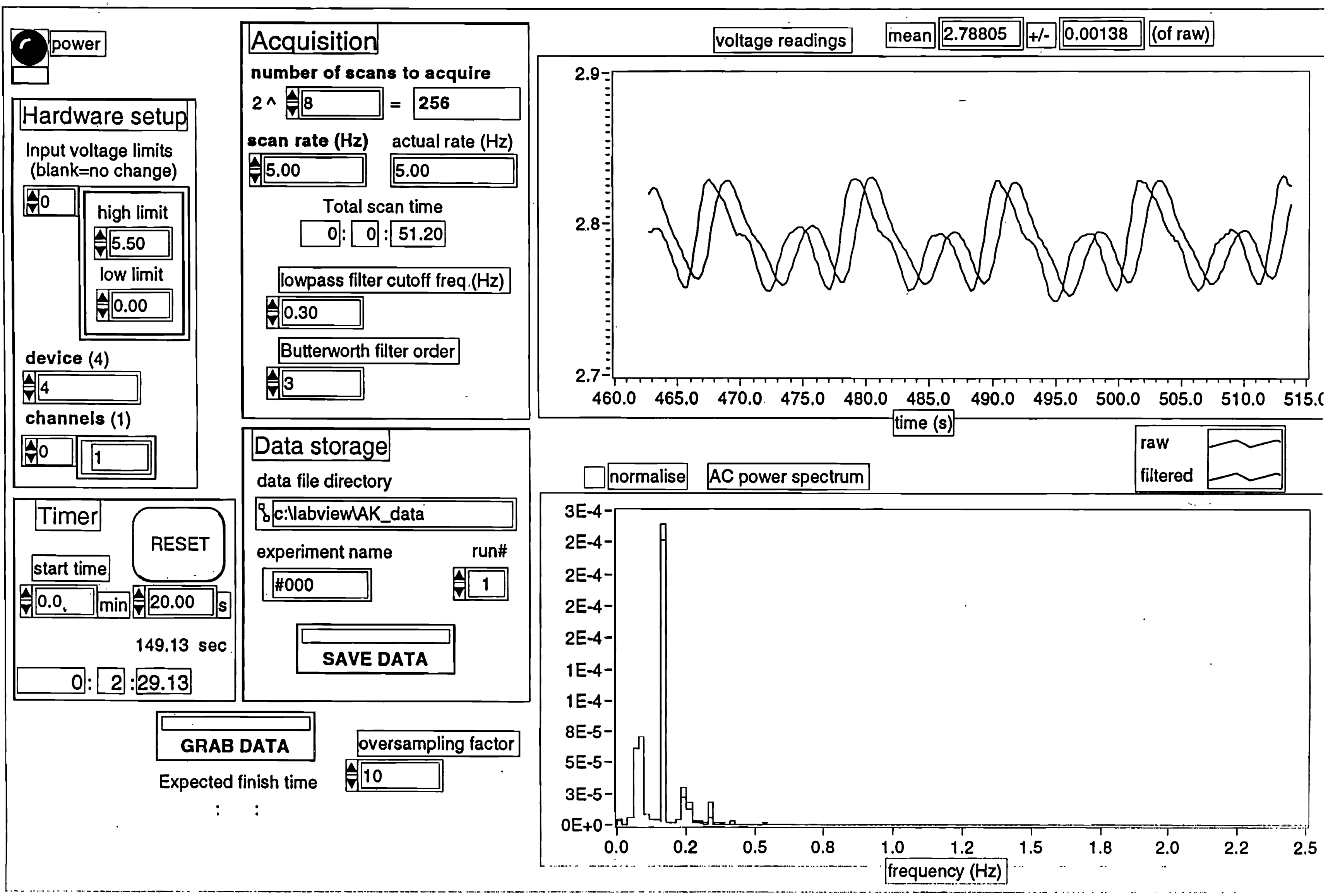
The program executes by the “dataflow” process, whereby a sub-VI will produce an output as soon as it receives all the inputs it needs. Thus the execution is not a sequential, linear process as it is with most text-based languages, and many steps may be executing simultaneously (at least conceptually, as it is running on a single-processor machine). The sequence construct is used to dictate an order to the processing done inside it - in my case, to acquire and filter the data before calculating its power spectrum.

The contents of all sub-VI’s I wrote (and the sub-VI's inside *them* in some cases) are shown on later pages, together with their connection terminals. All other icons represent VI’s which are part of the LabVIEW package - these are discussed in detail in the manuals.

The user-interface to the program is provided by the “front panel”, which mimics the controls and indicators of an electronic instrument whose circuitry is the program. A wide variety of different controls and indicators are provided, and can be arranged to form the front panel in the same way as one constructs pictures using an object-oriented drawing program. These indicators and controls also appear as icons in the circuit diagram and can be wired in to provide the user with interactive input and output from the program.

The front panel of the data acquisition program I constructed is shown on the next page. It allows the user to adjust the hardware settings and data acquisition parameters (such as sampling frequency and number of samples - the total sampling time display is updated automatically) and to reset the timer to any given time, allowing synchronisation with an external timebase (accurate to better than 500 ms). The lowpass Butterworth filter order (rolloff) and cutoff frequency can be adjusted, as can the oversampling factor. Oversampling reduces high-frequency random noise by taking a number of samples in each user-defined sampling period and averaging over these samples to produce each data point. The “GRAB DATA” button initiates data acquisition and displays the expected finish time. Once acquired, the raw and filtered data and their one-sided AC power spectra are displayed and the user is given the opportunity to save all the data in a spreadsheet-readable file of the chosen name.

Front Panel



power

Hardware setup

Input voltage limits (blank=no change): 0

high limit: 5.50

low limit: 0.00

device (4): 4

channels (1): 0 1

Timer

start time: 0.0 min 20.00 s

149.13 sec

0: 2 :29.13

Expected finish time : :

oversampling factor: 10

Acquisition

number of scans to acquire 2 ^ 8 = 256

scan rate (Hz): 5.00

actual rate (Hz): 5.00

Total scan time 0: 0 : 51.20

lowpass filter cutoff freq.(Hz): 0.30

Butterworth filter order: 3

Data storage

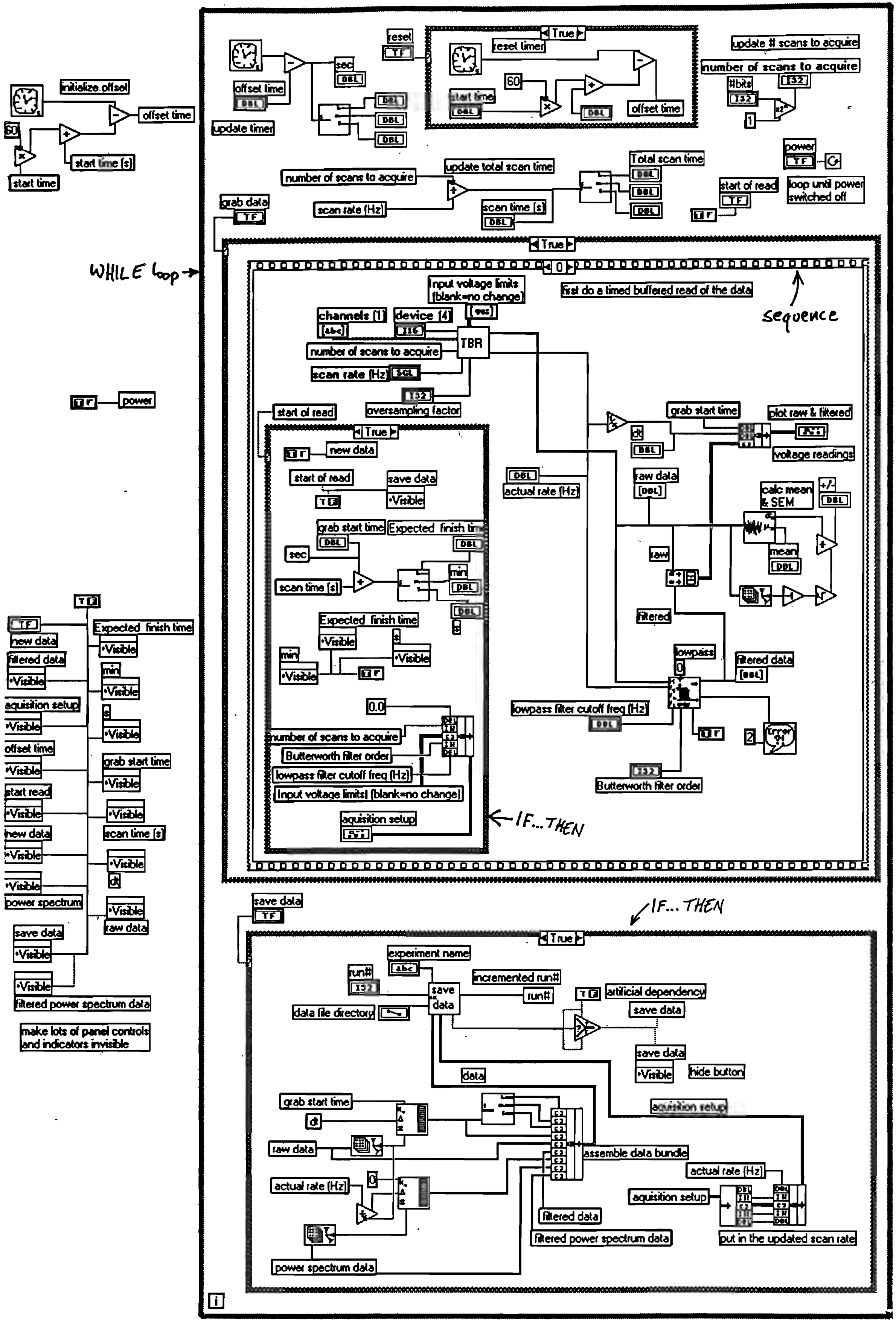
data file directory: c:\labview\AK_data

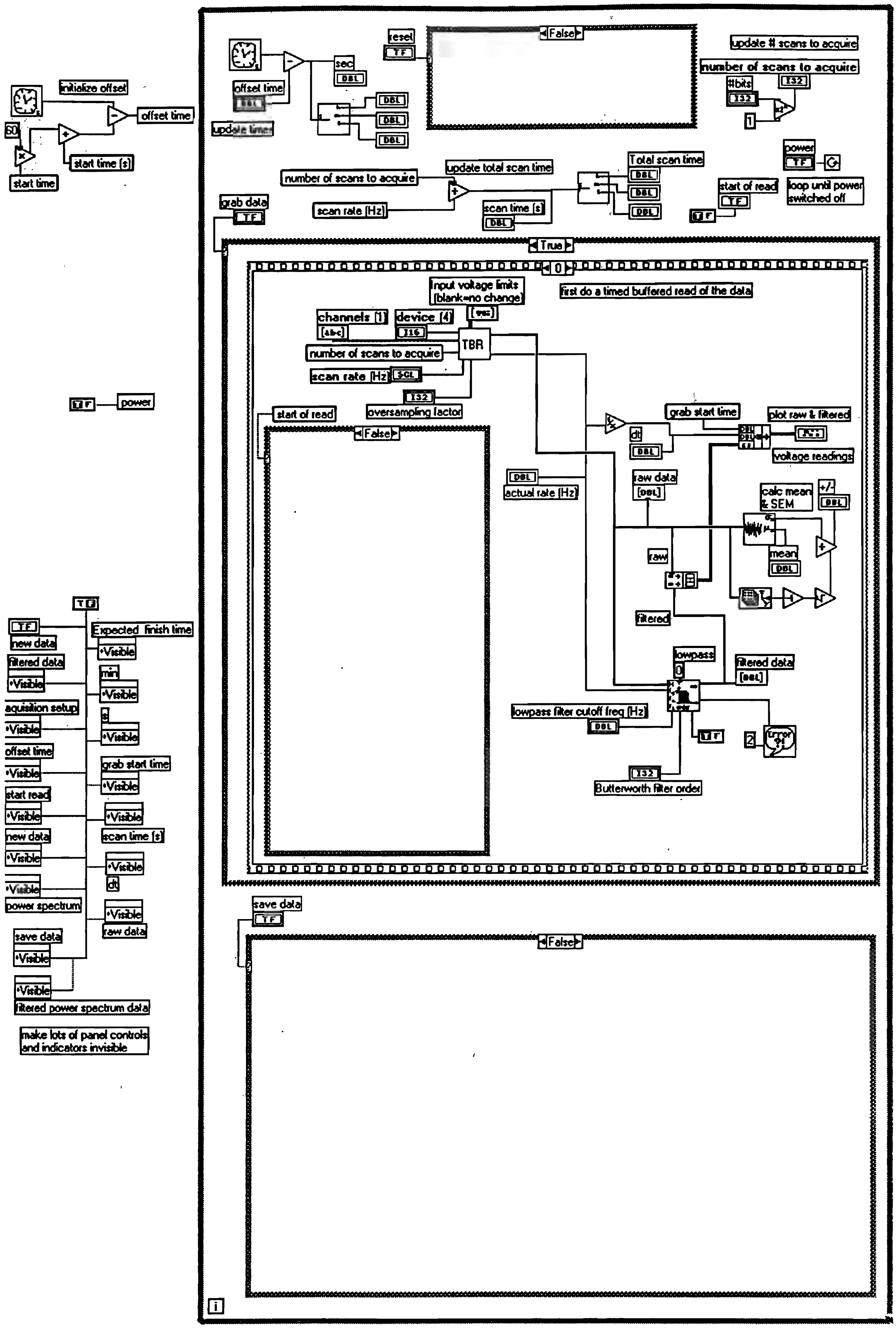
experiment name: #000

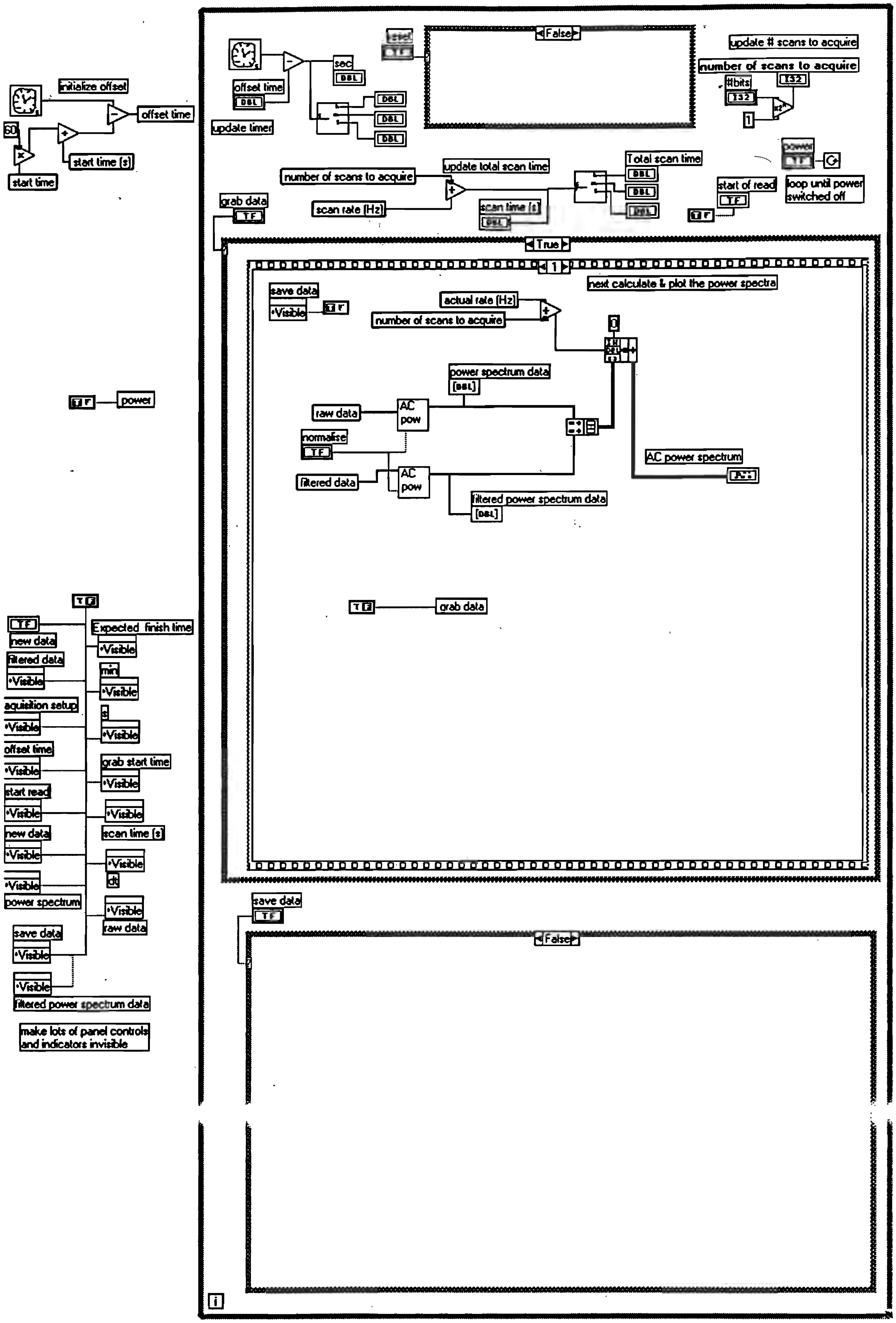
run#: 1

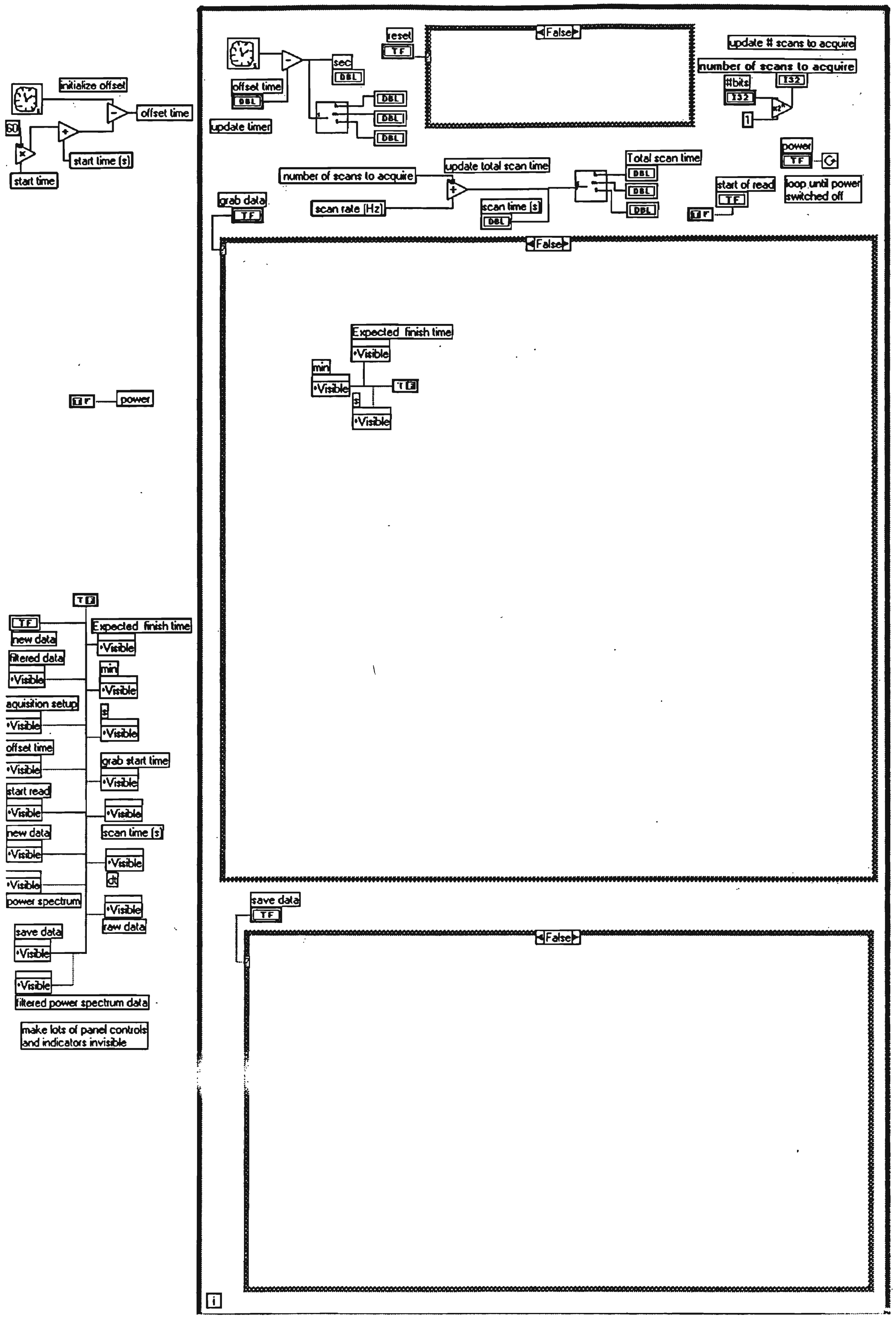
voltage readings mean 2.78805 +/- 0.00138 (of raw)

☐ normalise

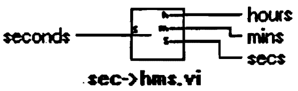

initialize offset
offset time
60
start time [s]
start time
WHILE loop
power
Expected finish time
new data
filtered data
•Visible
aquisition setup
offset time
start read
power spectrum
save data
filtered power spectrum data
min
s
grab start time
scan time [s]
dt
raw data
make lots of panel controls and indicators invisible
offset time
update timer
sec
reset
True
reset timer
start time
update # scans to acquire
number of scans to acquire
#bits
1
number of scans to acquire
update total scan time
Total scan time
scan rate [Hz]
scan time [s]
start of read
power
loop until power switched off
grab data
0
Input voltage limits [blank=no change]
first do a timed buffered read of the data
sequence
channels [1]
device [4]
TBR
number of scans to acquire
scan rate [Hz]
oversampling factor
start of read
new data
save data
Expected finish time
scan time [s]
Butterworth filter order
lowpass filter cutoff freq [Hz]
Input voltage limits [blank=no change]
aquisition setup
IF...THEN
0.0
actual rate [Hz]
grab start time
plot raw & filtered
voltage readings
raw data
calc mean & SEM
+/-
mean
raw
filtered
lowpass
filtered data
lowpass filter cutoff freq [Hz]
Butterworth filter order
2
error
IF...THEN
save data
experiment name
run#
incremented run#
save data
data file directory
artificial dependency
hide button
data
aquisition setup
grab start time
dt
raw data
actual rate [Hz]
assemble data bundle
actual rate [Hz]
aquisition setup
filtered data
filtered power spectrum data
put in the updated scan rate
power spectrum data
i


initialize offset
offset time
60
start time [s]
start time
power
Expected finish time
Visible
new data
filtered data
Visible
min
Visible
aquisition setup
s
Visible
Visible
offset time
grab start time
Visible
Visible
start read
Visible
Visible
new data
scan time [s]
Visible
Visible
dt
Visible
power spectrum
Visible
save data
raw data
Visible
Visible
filtered power spectrum data
make lots of panel controls and indicators invisible
reset
False
offset time
sec
update times
update # scans to acquire
number of scans to acquire
#bits
I32
1
number of scans to acquire
update total scan time
Total scan time
scan rate [Hz]
scan time [s]
start of read
power
loop until power switched off
grab data
True
0
Input voltage limits (blank=no change)
first do a timed buffered read of the data
channels [1]
device [4]
TBR
number of scans to acquire
scan rate [Hz]
SGL
I32
oversampling factor
start of read
False
grab start time
plot raw & filtered
dt
voltage readings
actual rate [Hz]
raw data
calc mean & SEM
+/-
mean
raw
filtered
lowpass
0
filtered data
lowpass filter cutoff freq [Hz]
2
Error
I32
Butterworth filter order
save data
False
i

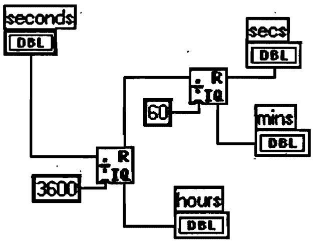

initialize offset
60
offset time
start time (s)
start time
power
reset
False
sec
offset time
update timer
update # scans to acquire
number of scans to acquire
#bits
1
power
number of scans to acquire
update total scan time
Total scan time
scan rate (Hz)
scan time (s)
grab data
start of read
loop until power switched off
True
1
next calculate & plot the power spectra
save data
Visible
actual rate (Hz)
number of scans to acquire
0
power spectrum data
raw data
AC pow
normalise
filtered data
AC pow
AC power spectrum
filtered power spectrum data
grab data
Expected finish time
Visible
new data
filtered data
Visible
min
Visible
aquisition setup
Visible
s
Visible
offset time
Visible
grab start time
Visible
start read
Visible
Visible
scan time (s)
new data
Visible
Visible
dt
Visible
power spectrum
Visible
raw data
save data
Visible
Visible
filtered power spectrum data
make lots of panel controls and indicators invisible
save data
False
i


initialize offset
offset time
60
start time (s)
start time
power
new data
filtered data
Visible
aquisition setup
Visible
offset time
Visible
start read
Visible
new data
Visible
Visible
power spectrum
save data
Visible
Visible
filtered power spectrum data
Expected finish time
Visible
min
Visible
s
Visible
grab start time
Visible
Visible
scan time (s)
Visible
dt
Visible
raw data
make lots of panel controls and indicators invisible
reset
False
sec
offset time
update timer
update # scans to acquire
number of scans to acquire
#bits
I32
I32
1
update total scan time
number of scans to acquire
scan rate (Hz)
scan time (s)
Total scan time
start of read
power
loop until power switched off
grab data
False
Expected finish time
Visible
min
Visible
s
Visible
save data
False
i

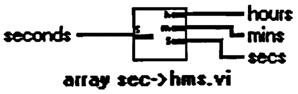
Connector Pane



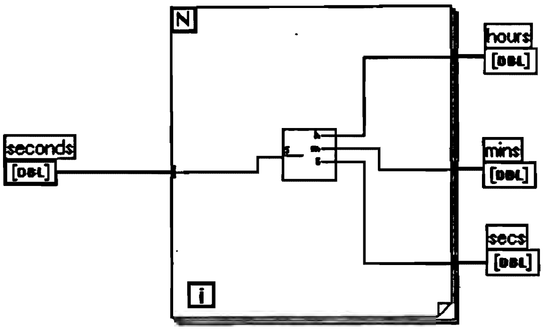
Block Diagram



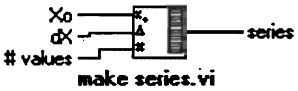
Connector Pane



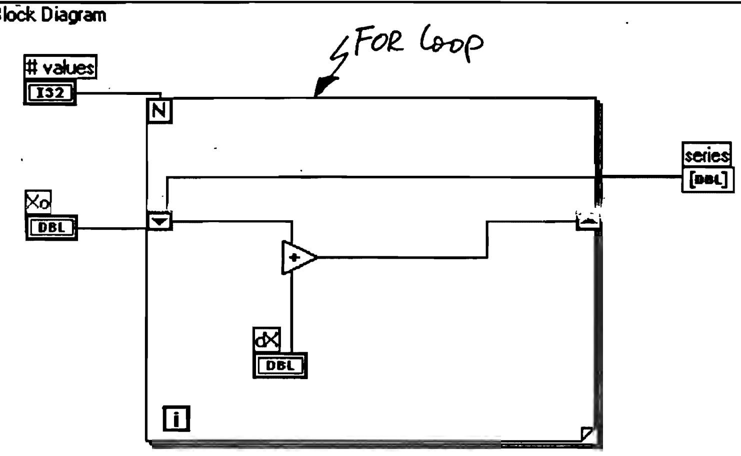
Block Diagram



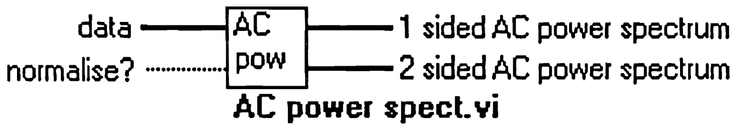
Connector Pane



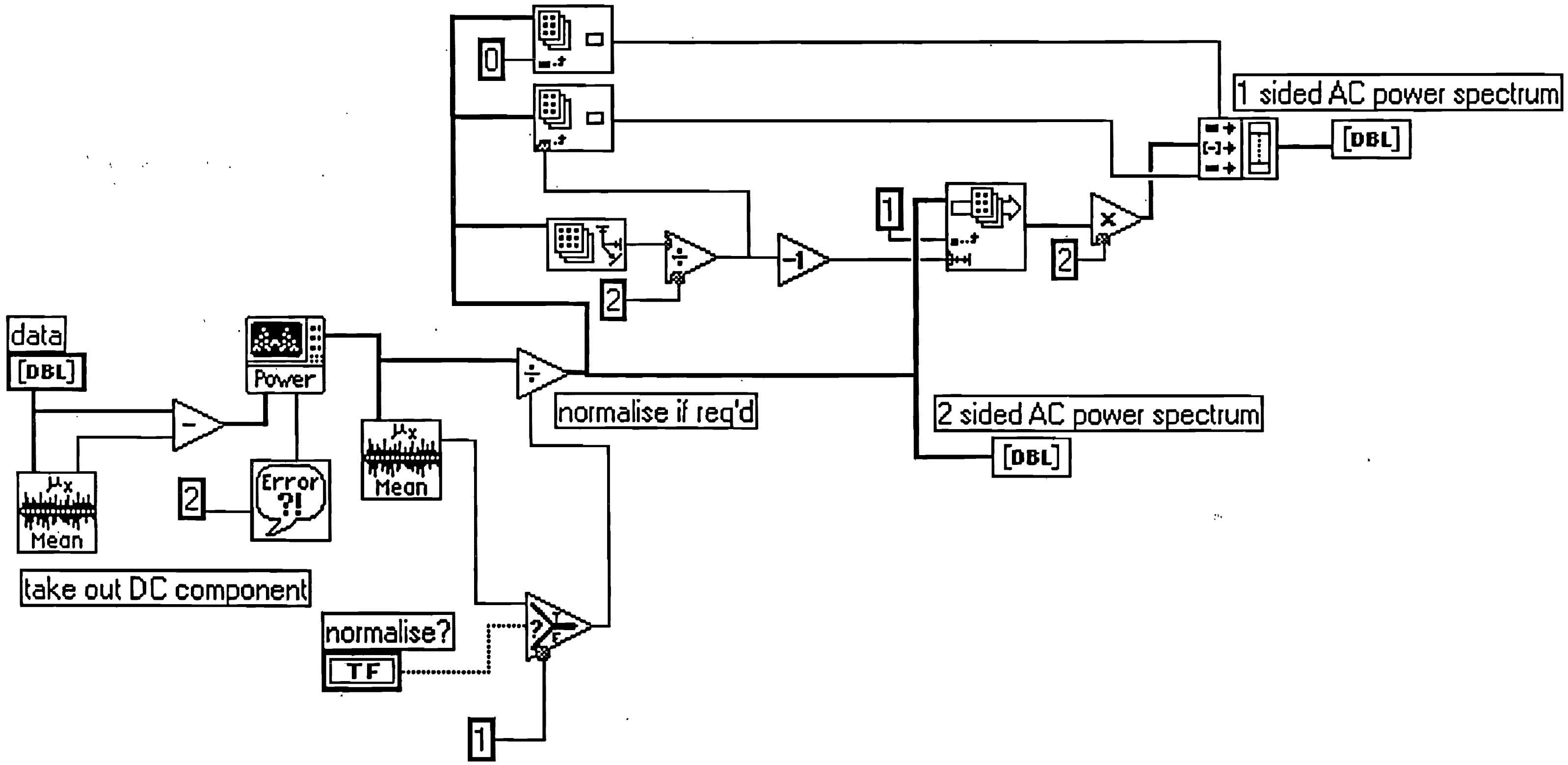
Block Diagram



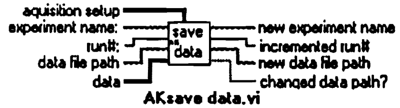
Connector Pane



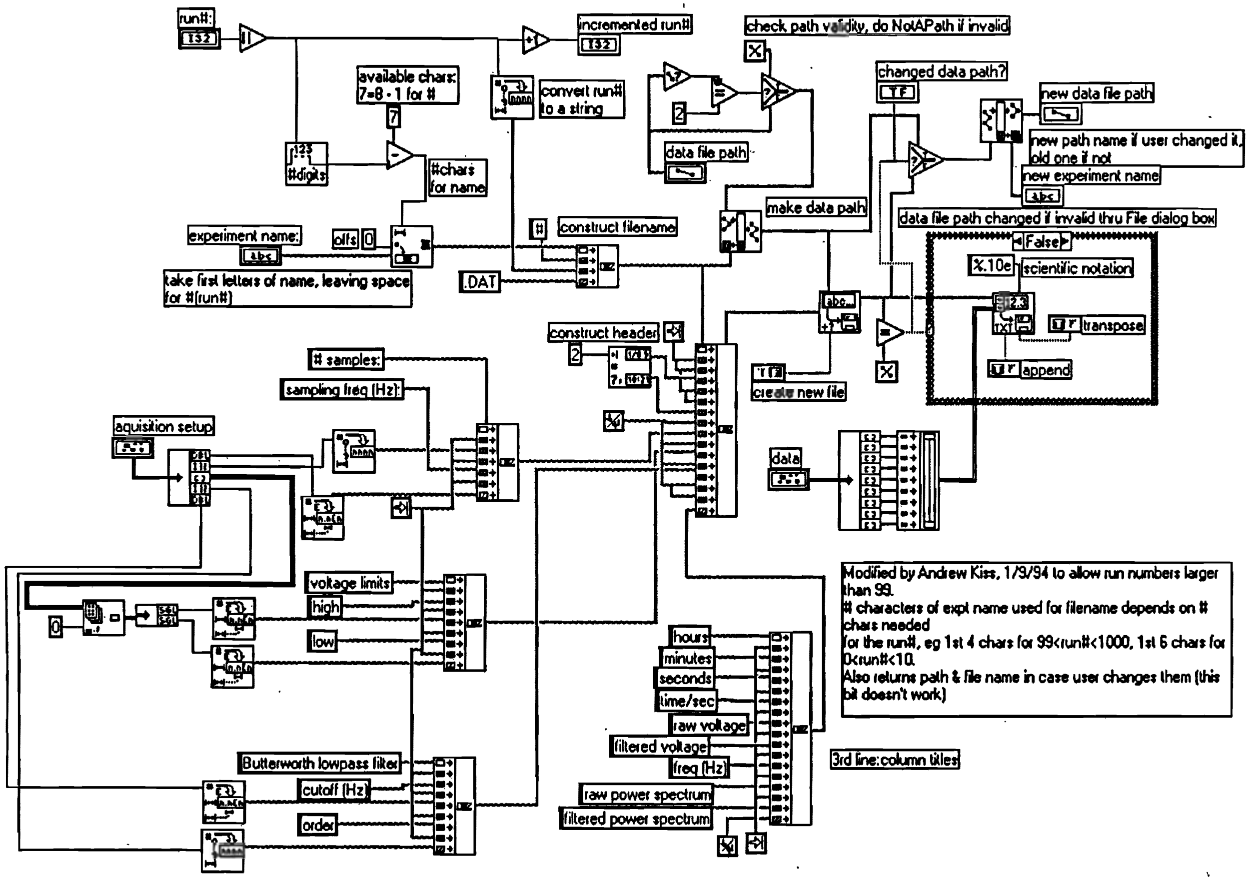
Block Diagram



Connector Pane

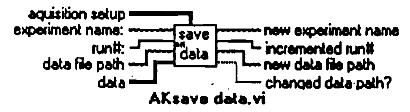


Block Diagram

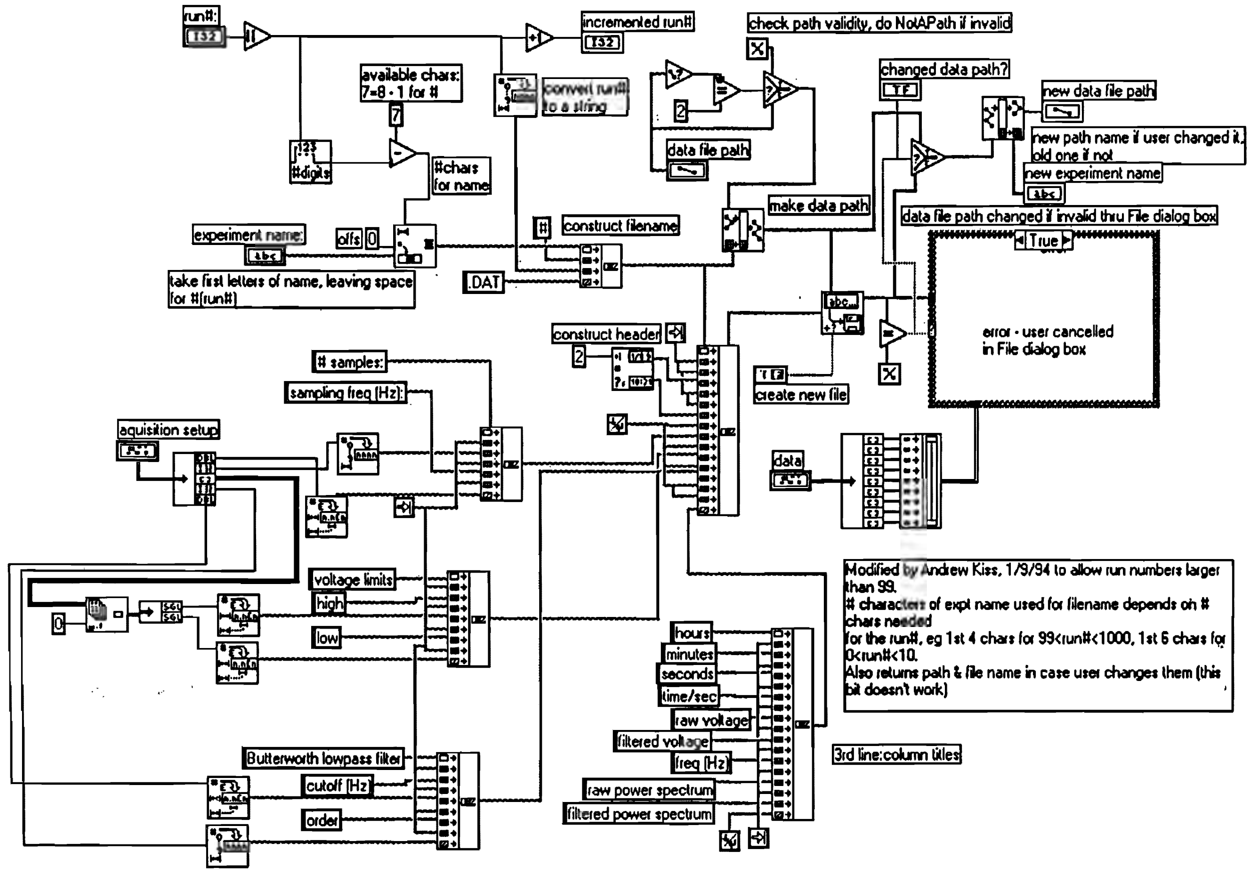


Modified by Andrew Kiss, 1/9/94 to allow run numbers larger than 99.
characters of expt name used for filename depends on # chars needed
for the run#, eg 1st 4 chars for 99<run#<1000, 1st 6 chars for 0<run#<10.
Also returns path & file name in case user changes them (this bit doesn't work)

Connector Pane

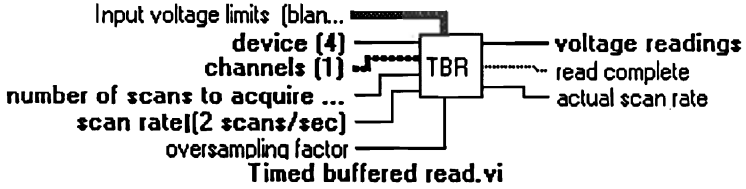


Block Diagram



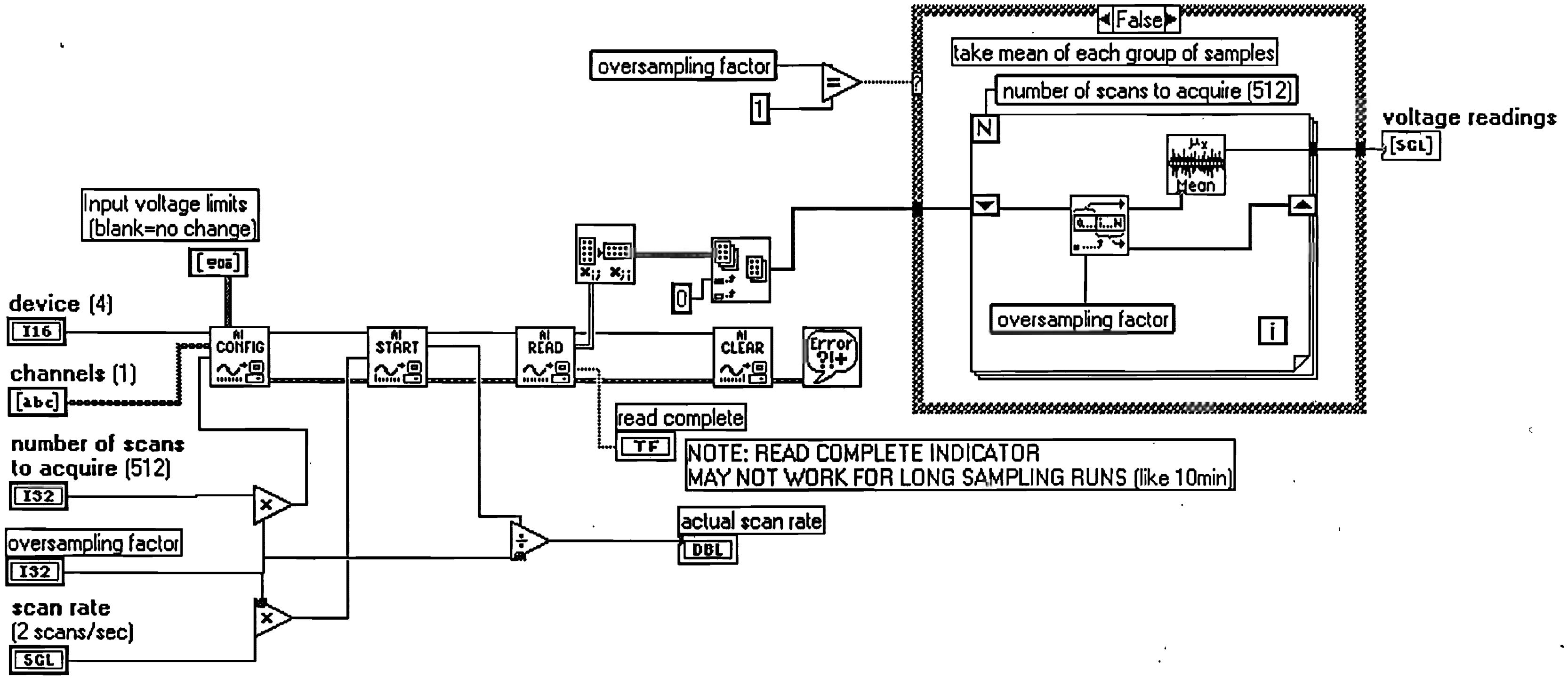
Modified by Andrew Kiss, 1/9/94 to allow run numbers larger than 99.
characters of expt name used for filename depends on # chars needed
for the run#, eg 1st 4 chars for 99<run#<1000, 1st 6 chars for 0<run#<10.
Also returns path & file name in case user changes them (this bit doesn't work)

Connector Pane

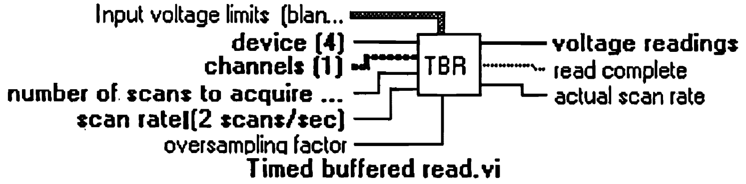


Timed buffered read.vi

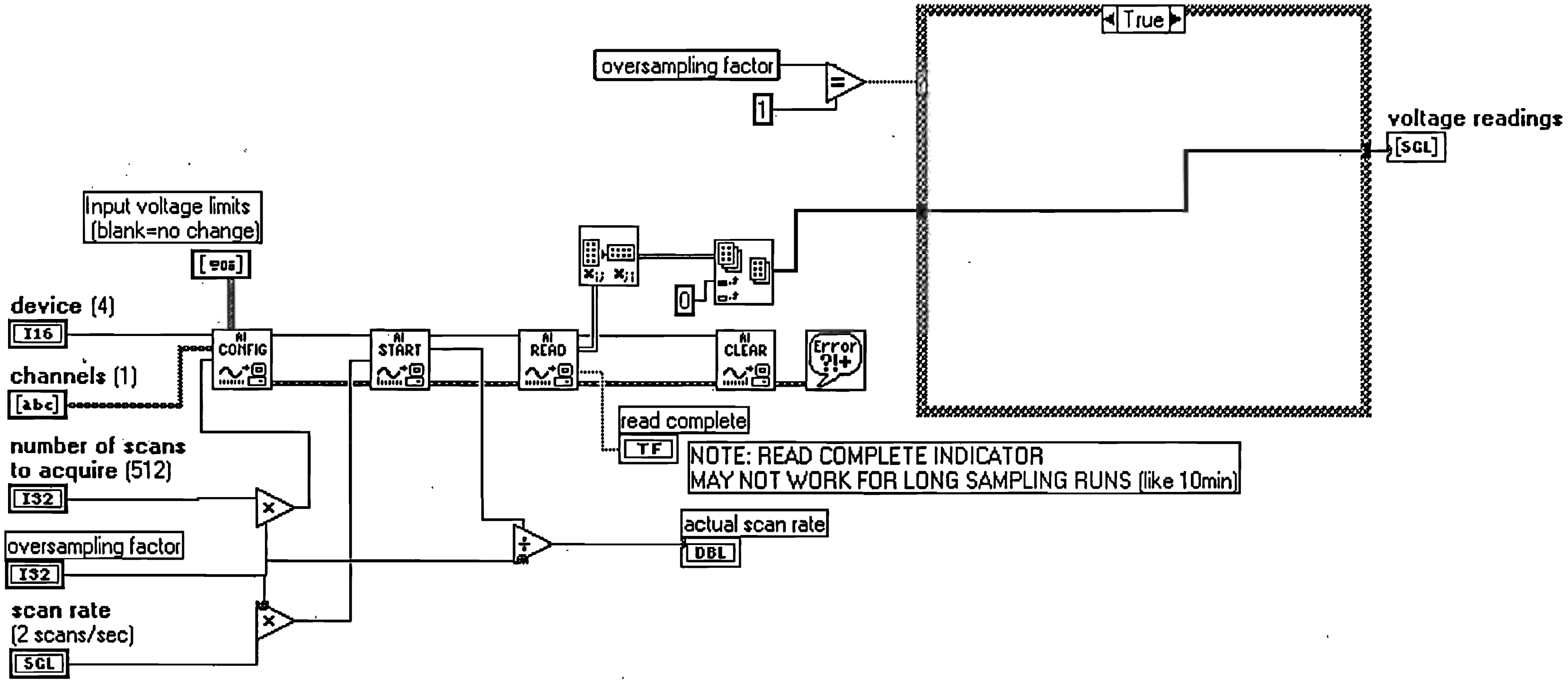
Block Diagram



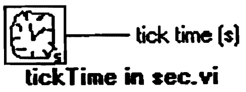
Connector Pane



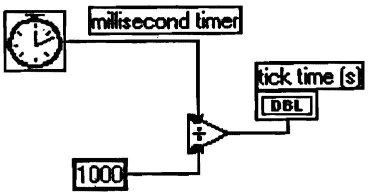
Block Diagram



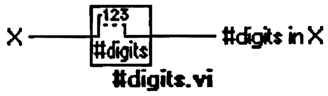
Connector Pane



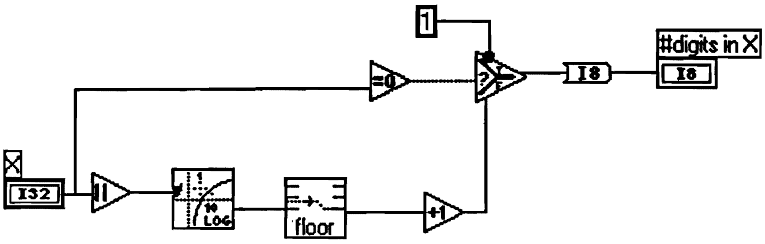
Block Diagram



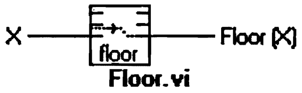
Connector Pane



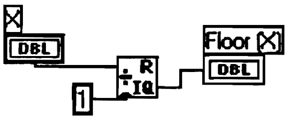
Block Diagram



Connector Pane



Block Diagram



Appendix 4: Second harmonic generation

The wavemakers were found to produce secondary waves at twice the driving frequency ω when ω was less than $N/2$. The mechanism by which these second-harmonic waves were produced was puzzling. Second-harmonic wave generation has been widely reported in the literature (for example, Mowbray & Rarity, 1966; Kistovich *et al.*, 1990), but to my knowledge its origin has not been discussed.

Although the wavemaker motion was not exactly sinusoidal due to the finite size of the stroke relative to the crank length, Fourier analysis of the calculated wavemaker motion showed that the amplitude of the second harmonic should have been about 0.4% as large as that of the fundamental. Wave generation efficiency could not have increased by a factor of 100 between the fundamental and second harmonic frequencies, so the anharmonicity of the crank motion was insufficient to account for the strength of the second harmonic wave (40% that of the fundamental).

Second-harmonic generation did not depend strongly on the wavemaker size or shape (several wavemakers were tried, from 1 cm to 10 cm high, with profiles ranging from square to semicircular), or on the flow around the wavemaker (it made no difference if there was a gap of 5 mm between the wavemaker and the wall). However the second harmonic wave amplitude decreased more rapidly than that of the fundamental when the stroke of the wavemaker was reduced, suggesting two possible causes for the phenomenon:

[1] The dependence on wave amplitude suggested that the second harmonic may have arisen as a result of some nonlinear process. In particular, the following mechanism would explain the general results. Consider the up- and down-going wave beams from the source to both have frequencies equal to that of the wavemaker, but to consist of a range of different wavelengths (recall that wave frequency and wavelength are not coupled by the dispersion relation). As the beams were only about 1.5 wavelengths wide, their spatial Fourier transforms would not have been very monochromatic, and such a collection of wavelengths is plausible. The up- and down-going wave beams overlap in the vicinity of the wavemaker, and a resonant triad could form in this region (see section 2.4). If we use subscripts 1 and 2 for the two beams and 3 for a second harmonic, then a possible resonant triad is given by

$$\omega_3=\omega_1+\omega_2=2\omega_1 \qquad \text{(A4.1)}$$

and

$$\mathbf{k}_3=\mathbf{k}_1+\mathbf{k}_2 \qquad \text{(A4.2).}$$

Note that (A4.1) gives the second harmonic frequency. To result in a travelling wave, $\mathbf{k}_3$ must satisfy the dispersion relation (2.1.7) for the frequency $2\omega_1$ - this is why a range of wavelengths is needed in the two primary beams. As Fig. A4.1 shows, such a condition can be satisfied if the interacting wavevectors from the two beams have the correct ratio of magnitudes.

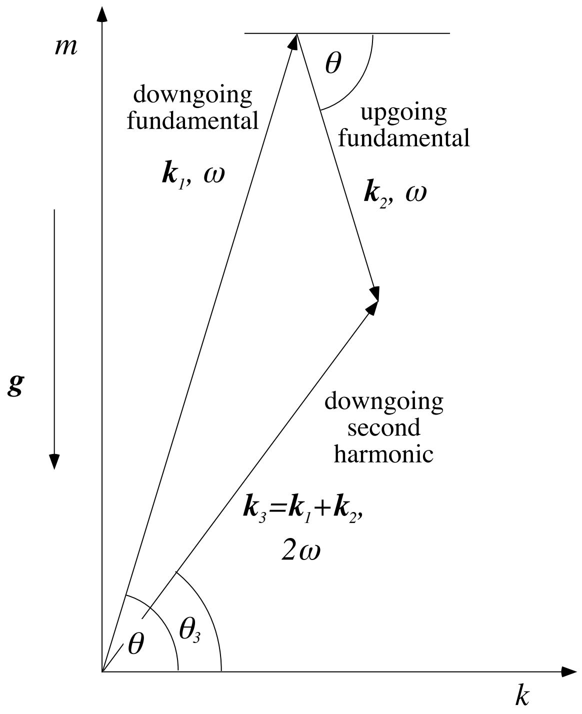


Fig. A4.1: Wavevectors for a resonant triad.

[2] The second possible mechanism: The wavemaker has a sinusoidal vertical motion with frequency ω. In regions in which the slope of the wavemaker does not change sign over the length of one stroke, the resultant horizontal fluid displacements are also periodic, with the same frequency (and higher harmonics of that frequency if the wavemaker is curved). Near the point at which the tangent to the wavemaker is vertical, horizontal fluid displacements change sign *twice* over one cycle of the wavemaker, resulting in horizontal motions with a frequency 2ω which are strong compared to those of frequency ω. Horizontal motions are important in a stratified fluid because they are not inhibited by restoring buoyancy forces. The horizontal motion with frequency 2ω may result in a second-harmonic wave beam. The region which generates this second harmonic wave becomes larger as the stroke is increased, thus increasing the relative amplitude of the second harmonic as observed.

References

APEL, J. R. 1987 *Principles of Ocean Physics.* Academic Press.

BALL, F. K. 1963 Energy transfer between external and internal gravity waves. *J. Fluid Mech.* **19**, 465-478.

BARENBLATT, G. I. 1987 *Dimensional Analysis*. Gordon and Breach Science Publishers.

BATCHELOR, G. K. 1970 *An Introduction to Fluid Dynamics*. Cambridge University Press.

BOOKER, J. R. & BRETHERTON, F. P. 1967 The critical layer for internal gravity waves in a shear flow. *J. Fluid Mech.* **27**, 513-539.

BRETHERTON, F. P. 1966 The propagation of groups of internal gravity waves in a shear flow. *Q. J. R. Met. Soc.* **92**, 466-480.

BRISCOE, M. G. 1975 Introduction to collection of papers on oceanic internal waves. *J. Geophys. Res.* **80**, 289-290.

BROUTMAN, D. 1984 The focusing of short internal waves by an inertial wave. *Geophys. Astrophys. Fluid Dyn.* **30**, 199-225.

BROUTMAN, D. 1986 On internal wave caustics. *J. Phys. Oceanogr.* **16**, 1625-1635.

BROUTMAN, D. & GRIMSHAW, R. 1988 The energetics of the interaction between short small-amplitude internal waves and inertial waves. *J. Fluid Mech.* **196**, 93-106.

BROUTMAN, D. & MCINTYRE, M. E. 1994 Classifying encounters between short oceanic internal waves and inertial waves. (unpublished draft).

BROUTMAN, D. & YOUNG, W. R. 1986 On the interaction of small-scale oceanic internal waves with near-inertial waves. *J. Fluid Mech.* **166**, 341-358. https://doi.org/10.1017/S0022112086000186

CURLE, N. & DAVIES, H. J. 1968 *Modern Fluid Dynamics*. D. Van Nostrand Co.

DWIGHT, H. B. 1961 *Tables of Integrals and Other Mathematical Data*. 4th Ed. Macmillan.

ECKART, C. 1960 *Hydrodynamics of Oceans and Atmospheres*. Pergamon Press.

GARRETT, C. & MUNK, W. 1972 Space-time scales of internal waves. *Geophys. Fluid. Dyn.* **3**, 225-264.

GARRETT, C. & MUNK, W. 1975 Space-time scales of internal waves: a progress report. *J. Geophys. Res.* **80**, 291-297.

GILL, A. E. 1982 *Atmosphere-Ocean Dynamics*. Academic Press.

GUENTHER, R. D. 1990 *Modern Optics*. John Wiley & Sons.

HANAZAKI, H. 1993 On the nonlinear internal waves excited in the flow of a linearly stratified Boussinesq fluid. *Phys. Fluids A.* **5** No. 5, 1201-1205.

HASSELMANN, K. 1966 Feynman diagrams and interaction rules of wave-wave scattering processes. *Rev. Geophys.* **4** No. 1, 1-33.

HAYES, W. D. 1970 Kinematic wave theory. *Proc. Roy. Soc. Lond. A* **320**, 209-226.

HENYEY, F. S. & POMPHREY, N. 1983 Eikonal description of internal wave interactions: a non-diffusive picture of "induced diffusion". *Dyn. Atmos. Oceans.* **7**, 189-219.

HEAD, M. J. 1983 The use of miniature four-electrode conductivity probes for high-resolution measurement of turbulent density or temperature variations in salt-stratified water flows. *PhD thesis, University of California, San Diego*.

HOLDER, D. W. & NORTH, R. J. 1963 *Schlieren Methods*. HMSO.

HOLLOWAY, G. 1980 Oceanic internal waves are not weak waves. *J. Phys. Oceanogr.* **10**, 906-914.

HOLLOWAY, G. 1982 On interaction time scales of oceanic internal waves. *J. Phys. Oceanogr.* **12**, 293-296.

IVANOV, A. V. 1989 Generation of internal waves by an oscillating source. *Izvestiya, Atmos. Ocean. Phys.* **25**, No. 1, 61-64.

IZERGIN, V. L. 1992 Equations for the modulations of a nonlinear packet of internal gravity waves in a shear flow. *Izvestiya, Atmos. Ocean. Phys.* **28**, No. 6, 496-499.

KISTOVICH, A. V., NEKLYUDOV, V. I. & CHASHECHKIN, YU. D. 1990 Nonlinear two-dimensional internal waves generated by a periodically moving source in an exponentially stratified medium. *Izvestiya, Atmos. Ocean. Phys.* **26**, No. 10, 771-776.

KLEINE, H. & GRÖNIG, H. 1991 Color schlieren methods in shock wave research. *Shock Waves.* **1**, 51-63.

LAMB, H. 1932 *Hydrodynamics*, 6th Ed. Cambridge University Press.

LAMBERT, R. B. JR. & DAVEY, M. 1974 Continuously direct-reading polarimeter for density contrast measurements in optically active solutions. *Rev. Sci. Instrum.* **45** No.12, 1531-1536.

LANDAU, L. D. & LIFSHITZ, E. M. 1969 *Mechanics.* 2nd Ed. Pergamon Press.

LANDAU, L. D. & LIFSHITZ, E. M. 1959 *Fluid Mechanics.* Pergamon Press.

LEBLOND, P. & MYSAK, L. 1978 *Waves in the Ocean.* Elsevier Scientific Pub. Co.

LIDE, D. R., ED. *et. al.* 1991 *CRC Handbook of Chemistry and Physics*, 72nd Ed. CRC Press.

LIGHTHILL, M. J. 1960 Studies on magneto-hydrodynamic waves and other anisotropic wave motions. *Phil. Trans. Roy. Soc. A* **252** 397-430.

LIGHTHILL, M. J. 1978 *Waves in Fluids.* Cambridge University Press.

MACASKILL, C. & BROUTMAN, D. 1992 A numerical investigation of the interaction of short internal waves with inertial waves. *11th Australasian Fluid Mech. Conf., Hobart.*

MCCOMAS, C. H. & BRETHERTON F. P. 1977 Resonant interaction of oceanic internal waves. *J. Geophys. Res.* **82** No. 9, 1397-1412.

MCEWAN, A. D. 1971 Degeneration of resonantly-excited standing internal gravity waves. *J. Fluid Mech.* **50**, 431-448.

MCEWAN, A. D. 1973 Interactions between internal gravity waves and their traumatic effect on a continuous stratification. *Boundary-Layer Meteor.* **5**, 159-175.

MCEWAN, A. D., MANDER D. W. & SMITH, R. K. 1972 Forced resonant second-order interaction between damped internal waves. *J. Fluid Mech.* **55**, 589-608.

MERZKIRCH, W. 1974 *Flow Visualization.* Academic Press.

MOWBRAY, D. E. 1967 The use of schlieren and shadowgraph techniques in the study of flow patterns in density stratified liquids. *J. Fluid Mech.* **27**, 595-608.

MOWBRAY, D. E. & RARITY, B. S. 1966 A theoretical and experimental investigation of the phase configuration of internal waves of small amplitude in a density stratified liquid. *J. Fluid Mech.* **28**, 1-16.

MÜLLER, P., HOLLOWAY, G., HENYEY, F. & POMPHREY, N. 1986 Nonlinear interactions among internal gravity waves. *Rev. Geophys.* **24**, No. 3, 493-536.

OFFICER, C. B. 1974 *Introduction to Theoretical Geophysics.* Springer-Verlag.

OSTER, G. 1965 Density gradients. *Sci. Amer.* **213,** No. 2, 70-76.

PHILLIPS, O. M. 1960 On the dynamics of unsteady gravity waves of finite amplitude, part 1. *J. Fluid Mech.* **9**, 193-217.

PHILLIPS, O. M. 1966 *The Dynamics of the Upper Ocean*, 2nd Ed. Cambridge University Press.

PHILLIPS, O. M. 1981 Wave interactions - the evolution of an idea. *J. Fluid Mech.* **106**, 215-227.

RAYLEIGH, LORD 1883 Investigation of the character of the equilibrium of an incompressible heavy fluid of variable density. *Proc. Lond. Math. Soc.* **14**, 170-177.

RENOUARD, D. P., TOMASSON, G. G. & MELVILLE, W. K. 1993 An experimental and numerical study of nonlinear internal waves. *Phys. Fluids A.* **5** No. 6, 1401-1411.

RENOUARD, D. P. & GERMAIN, J.-P. 1994 Experimental study of long nonlinear internal waves in a rotationg fluid. *Ann. Geophysicae.* **12**, 254-264.

RUDDICK, B. R. & SHIRTCLIFFE, T. G. L. 1979 Data for double diffusers: physical properties of aqueous salt-sugar solutions. *Deep-Sea Res.* **26A**, 775-787.

SAMODUROV, A. S. 1974 Plane nonlinear internal gravity waves in a stratified fluid. *Izvestiya, Atmos. Ocean. Phys.* **10**, No. 8, 904-906.

STEVENSON, T. N. 1968 Some two-dimensional internal waves in a stratified fluid. *J. Fluid Mech.* **33**, 715-720.

STEVENSON, T. N. 1973 The phase configuration of internal waves around a body moving in a density stratified fluid. *J. Fluid Mech.* **60** 759-767.

THOMAS, N. H. & STEVENSON, T. N. 1972 A similarity solution for viscous internal waves. *J. Fluid Mech.* **54**, 495-506. https://doi.org/10.1017/S0022112072000837

THORPE, S. A. 1966 On wave interactions in a stratified fluid. *J. Fluid Mech.* **24**, 737-751.

THORPE, S. A. 1968 On the shape of progressive internal waves. *Phil. Trans. Roy. Soc. A.* **263**, 563-614.

THORPE, S. A. 1975 The excitation, dissipation and interaction of internal waves in the deep ocean. *J. Geophys. Res.* **80**, No. 3, 328-338.

THORPE, S. A. 1989 The distortion of short internal waves produced by a long wave, with application to ocean boundary mixing. *J. Fluid Mech.* **208**, 395-415.

TRITTON, D. J. 1977 *Physical Fluid Dynamics.* Van Nostrand.

TURNER, J. S. 1973 *Buoyancy Effects in Fluids.* Cambridge University Press.

VON SCHWIND, J. J. 1980 *Geophysical Fluid Dynamics for Oceanographers.* Prentice-Hall.

WASHBURN, E. W., Ed. *et. al.* 1927 *International Critical Tables of Numerical Data, Physics, Chemistry and Technology, Vol II.* McGraw-Hill.

WITHAM, G. B. 1960 A note on group velocity. *J. Fluid Mech.* **9**, 347-352.

YIH, C.-S. 1965 *Dynamics of Nonhomogeneous Fluids.* Macmillan.