DETERMINING FISH NUMBER DENSITY BY A STATISTICAL ANALYSIS OF BACKSCATTERED SOUND

BY

QUINTON SMITH

SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER
OF SCIENCE IN ENGINEERING

UNIVERSITY OF CAPETOWN
APR 1988

SYNOPSIS

This thesis involves work done on the statistical analysis of the intensity of high frequancy sound pulses backscattered from fish shoals in order to determine the number density of fish. The abundance of fish species can be estimated from acoustic back-scattering signals without requiring knowledge of the target strength of fish nor calibration of sonar equipment. To do the statistical analysis a statistical model of the backscattered intensity is needed. The model is then used to analyse the sampled backscattered intensity in order to determine the fish number density.

The model used in this research assumes that fish are identical and independent scatterers of sound and that they are randomly distributed in the water. A generalised model of the statistics of the backscatter from a shoal of fish is developed to answer the question of how the estimates of fish number density vary for different scatterer types and different distributions of fish. An examination is made of all assumptions underlying the statistical model with a view to extending the model still further. The model is applied to various volume distributions including the Poisson distribution and to Ricean,Rayleigh and constant amplitude scattering statistics. The model is extended to include the non-ideal theoretical beam pattern of both circular and square transducers.

An investigation of the number of statistically independent measurements needed in the statistical analysis of data arising from the model is done in order to determine the standard deviation in the resulting estimate of number

density obtained from the model. An expression is derived for the standard deviation in the number density in terms of the statistical parameters of the model (such as beamshape, scatterer nature and distribution) and the number of statistical independent measurements used in the analysis of data. A method of dealing with fish shoals which have variations in density within them is developed. A theoretical comparison of this method with the well established method of echo-integration is made.

The scope of this study is limited to stationary statistics (which ignores beam divergence) although suggestions are made on extending the model to non-stationary statistics. The model is verified by an analysis of the backscattered sound intensity from which the fish number density is estimated and compared to the estimate obtained by the echo integration method. In the analysis of data a Poisson distribution and Rayleigh amplitude statistics were assumed. It was not within the scope of this thesis to verify these assumptions or to determine the actual fish distribution and scattering statistics in the shoals under investigation. However it is argued that these assumptions are approximately correct. The data was collected at night time on dispersed anchovy and redeye shoals of low density as the method works best on low densities. The statistical estimate was made using a method which applied to shoals containing density variations. An estimate of standard deviation was obtained for the statistical methods' estimate of number density for comparison with echo integration estimates of number densities. Good agreement between the statistical and echo integration method estimates of number density were obtained once standard deviation and scatterer statistics were taken into account.

This work suggests a method of estimating fish numberdensity (which does not require target strengths) to be used in fish counting surveys.

ACKNOWLEDGEMENTS

I would like to thank the following people for their help and encouragement.

Professor P.N.Denbigh my supervisor

Mr I.Hampton (of the Sea Fisheries Research Institute, Capetown) for advice and help with the collection of data and providing the facilities for the analysis of data.

Professor J.F.W. Bell for advice on thesis presentation.

Prof G.B.Brundritt of the Oceanography Dept,U.C.T.

Leslie Shackleton and Roger Krohn, Liaison officers of the Benguela Ecology Program.

Jonathan Weintroub for producing some of the experimental data used in this thesis.

Support for this work was provided by the Council for Scientific and Industrial Research Foundation for Research and Developement and the Benguela Ecology Program of the CSIR.

CONTENTS

pg

1. INTRODUCTION 1-1

2. THEORY

2.1 Revision of Previous Theoretical Work 2-1

2.1.1 The Work of Pusey on Single Interval Statistics of Light Scattered by Identical Independent Scatterers 2-1

2.1.2 Extension of Pusey's Work to the Underwater Acoustic Scattering Problem 2-2

2.1.3 Using the Method on Fish at Sea 2-4

2.2 Examinations of Assumptions Underlying the Model 2-5

2.2.1 Scatterer Nature 2-6

2.2.2 Fish Distribution 2-6

2.2.3 Multiple Scattering 2-7

2.2.4 Non-uniform Illumination of Scatterers 2-7

2.2.5 Backscatter, Reverberation and Ocean Noise 2-8

2.2.6 Experimental Conditions 2-8

2.2.7 Non-stationarity Effects 2-8

2.2.8 Peculiar Sensitivities 2-8

2.3 Derivation of Generalised Model 2-9

2.3.1 Application of Generalised Model to Poisson Volume Distribution and Various Amplitude Statistics 2-12

2.3.1.1 Constant Scatterer Amplitude 2-12

2.3.1.2 Rayleigh Amplitude Statistics 2-13

2.3.1.3 Ricean Statistics 2-13

2.3.2 Application of Generalised Model to Various Volume Distribution Statistics 2-14

2.3.2.1 Poisson Distribution 2-14

2.3.2.2 Binomial Distribution 2-15

2.3.2.3 Uniform Distribution 2-15
2.3.2.4 Constant Density 2-16
2.4 Effects of Beam Pattern 2-17
2.4.1 The Circular Piston Transducer 2-17
2.4.2 The Square Piston Transducer 2-17
2.5 The Effect of the Number of Meausurements on the Estimate of Number Density 2-18
2.6 Method of Dealing with Inhomogeneities in the Fish Shoal 2-20
2.6.1 Choice of Reference Density Levels 2-22
2.6.2 Choice of Density Levels with Regard to Standard Deviation of Estimates 2-23
2.6.3 The Calculation of Overall Density for a Shoal Consisting of a Number of Different Density Bands 2-25
2.6.4 Ensamble Averaging 2-26
2.7 Comparison of Theory of Echo Integration with that of Statistical Method 2-27
2.7.1 Theory of Echo Integration 2-27
2.7.2 Comparison with the Theory of the Statistical Method 2-28

3. EXPERIMENTAL DATA ANALYSIS (using data collected in 1985) 3-1
3.1 Examination of the Data 3-1
3.2 Results for Density Analysis 3-2
3.3 Effect of Varying Number of Density Bands 3-7
3.4 Effect of Varying Averaging Block Size 3-9
3.5 Comparison with Echo Integration 3-11

4. AFRICANA CRUISE DATA(June 1987) 4-1
4.1 Experiment I 4-1

4.2 Experiment II 4-3

5. CONCLUSION 5-1

6. REFERENCES 6-1

7. APPENDICES A-1

APPENDIX 1-Derivation of the Equations for $<I_N^2>$, $<I_N^3>$,$<I_N^4>$ A-1

A1.1 Equation for $<I_N^2>$ A-1

A1.2 Equation for $<I_N^3>$ A-2

A1.3 Equation for $<I_N^4>$ A-3

APPENDIX 2-Derivation of the Moments of a Ricean PDF A-4

APPENDIX 3-Calculation of the Moments for Various Distributon PDF's A-4

A3.1 Poisson PDF Moments A-5

A3.2 Binomial PDF Moments A-5

A3.3 Uniform PDF Moments A-5

APPENDIX 4-Calculation of $<b^4>/<b^2>^2$ for Different Transducers A-5

A4.1 Circular Piston Transducer A-5

A4.2 Rectangular Piston Transducer A-6

APPENDIX 5-Errors in the Estimates of Normalised Moments Due to a Finite Number of Measurements A-7

A5.1 Determining the Standard Deviation in the Second Normalised Moment A-7

A5.2 Derivation for Variance When Dealing With Constant Amplitude Scatterers and a Poisson Distribution A-9

A5.3 Derivation for Variance when Dealing with Rayleigh Amplitude Statistics and Poisson Distribution A-9

A5.4 Determining the Standard Deviation in <N> A-10
A5.5 Determining the Standard Deviation in <I> for the Purpose of Block Averaging A-10
A5.6 Determining the Standard Deviation in <N> for a Non-homogenous Density Shoal A-11
A5.7 Determining Standard Deviations Which Include the Effect of Beamshape A-12
APPENDIX 6-Listing of Program Used to Analyse Data A-14
APPENDIX 7-"Experiments with an Acoustic Technique for Estimating Fish Number Density, not Requiring Fish Target Strength" by P.N.Denbigh and Q.Smith Paper presented at the International Symposium on Fisheries Acoustics, June 22-26,1987,Seattle,Washington,U.S.A. A-17

1.0 INTRODUCTION

This thesis involves work done on the statistical analysis of the intensity of high frequancy sound pulses backscattered from fish shoals in order to determine the number density of fish. The abundance of fish species needed to be estimated from acoustic back-scattering signals more accurately than was possible using the well established echo integration method. The main drawback of the echo integration method is due to the difficulty in determining the target strength of the fish. The target strength has to be determined by doing a trawl on the fish shoal, weighing and measuring the length of the fish in order to determine the average length and weight of fish which is needed to evaluate the fishes target strength. This has to be done for every measurement of fish density and is time consuming and labour intensive and the target strength relation to fish length and weight is still not known very accurately. The advantage of the statistical method now under investigation is that it does not require knowledge of target strength, nor does it require calibration of the sonar. The method is based on work originally done by Pusey on scattering of light from molecules in a dilute solution. The concentration of molecules in the solution could then be determined by the statistical analysis of the light scattered back from the solution.(Pusey et al,1974) The method was extended by Wilhelmij to pulses of high frequancy sound scattered back from scattering arrays underwater. Wilhelmij simulated a fish shoal using polystyrene spheres suspended in a water tank. The statistical properties needed for the analysis are the first and second moment of the backscattered intensity. These are related to the average number of fish <N> in a resolution cell of the sonar by the equation $\langle N\rangle=1/(\langle I^2\rangle/\langle I\rangle^2-2)$ The volume of the resolution cell can be determined from the beamwidth of the transducer and the length of the sound pulse. Knowing the volume of the resolution cell and the average number of fish per resolution cell the fish number density (number of fish per m^3) can be determined.(Wilhelmij and Denbigh,1984;Wilhelmij 1983) The equation for <N> is derived from a mathematical model of the fish shoal which assumes that the fish are randomly distributed throughout the volume in which they occur and that this random distribution can be modelled by the Poisson or other probability density function. It is further assumed that each fish scatters with a random amplitude. When a fish shoal was simulated in the laboratory using polystyrene spheres suspended in a water tank by Wilhelmij the method proved accurate at estimating the number density of polystyrene spheres in the volume in which they were suspended provided the number of spheres in the resolution cell was less than 10. This implies that what is needed is a low density of the spheres or small resolution cells which could mean sonar transducers with narrow beams and the use of short pulse lengths of sound. The lower the density the more accurate the results were found to be. It was also found that a great many statistically independent samples of the backscattered intensity were needed for an accurate estimate (Wilhelmij and Denbigh,1984) Weintroub collected data at sea on actual fish shoals and it was found that the mathematical model had to be modified to deal with the scatter from fish. The scatter from fish was assumed to be Rayleigh (Weintraub and Denbigh,1985). Even after this modification they did still not achieve good agreement between the statistical and echo integration methods.In addition they proposed that the analysis had to take into account density variations within the fish shoal. It is the purpose of this thesis to account for these density variations and to further modify the mathematical model to deal with actual fish to improve the accuracy of the estimates of fish number density obtained by the statistical method. In addition to further improve the accuracy of the statistical method more realistic assumptions about the beamshape will be made. The thesis will also determine the standard deviation by an analytical formula more accurate than the approximate formula used by Wilhelmij and Denbigh(1984). Previously accurate estimates of standard deviation had to be obtained by lengthy computer simulation(Weintraub 1986). The analytical formula offers greater utility. An investigation of assumptions underlying the model will be made and a more generalised model will be developed to enable it to be adapted to various assumptions.

1.1 Mathematical Background for the Statistical Analysis of the Backscattered Sound Intensity.

The statistical properties of the backscattered sound intensity needed are the average and expected value of the intensity and intensity squared namely $E\{I\}$ and $E\{I^2\}$ otherwise known as the first and second moment.The notation $<I>$ and $<I^2>$ for the first and second moment is most often used in this thesis. To find $<I>$ and $<I^2>$ the backscattered intensity has to be sampled and then $<I>=\Sigma I_i/k$ and $<I^2>=\Sigma I_i^2/k$ where I_i are the samples of intensity, k=total number of samples. It is assumed that fish are randomly distributed in the volume of water and that this distribution can be described by a Poisson pdf. Thus there is a random number N of fish in a scattering volume and the probability of having k fish in any scattering volume is $P\{N=k\}=\exp(-<N>)<N>^k/k!$ $k=0,1,2,3....$ (Papoulis,pg 101) where $<N>=E\{N\}=\Sigma kP\{N=k\}$ (Papoulis, pg 145) Each fish will scatter the sound pulse with an amplitude that will vary from fish to fish depending on the fish orientation or size. The fish size and orientation will vary randomly thus the amplitude return from any fish is a random variable. The amplitude backscatter from any one fish has two main components 1) The sound reflected from the swimbladder (if the fish has one) 2) the sound reflected from the rest of the body(Clay and Heist,1984). The swimbladder is assumed to reflect sound with a relatively constant amplitude irrespective of fish size or orientation.Sound reflection from the rest of the body depends on fish orientation and changes as the fish swims and is thus random. Thus the fish scatters with both constant and randomly varying amplitudes. This can be modelled by the Ricean probability density function $w_R(a)=[2a/\sigma_d]\exp[-(a^2+\sigma_c)/\sigma_d]I_0[2a\sqrt{\sigma_c}/\sigma_d]$ (Clay and Heist,1984) where a=sound amplitude σ_c =average constant power σ_d =average distributed power. Define $\gamma=\sigma_c/\sigma_d$ .For small fish (or fish with no swimbladder) γ is small and can be taken to be approximately zero and then fish scatter sound with amplitude modelled by Rayleigh pdf. For large fish with large swimbladder γ is large and the sound scattered can be modelled by constant amplitude scatter.(Clay and Heist,1984) Thus each scattering volume has a random number of fish each scattering in a random manner. All the signals from each fish in the scattering volume add up to form the backscattered signal. Since the number of fish in each scattering volume is random the backscattered amplitude from any sample volume is the sum of a random number of random variables. This is called a "Random Sum" (Papoulis,pg248). Thus the backscattered amplitude $A=\sum_{i=1}^{\hat{N}} a_i$ where a_i is the scatter from the i'th fish and $\hat{N}$ is the number of fish in a scattering volume. Note that A is also a random variable. For the analysis in practise will mean taking the expected values of a non-homogenous filtered Poisson process.(Parzen pg156) I is called a filtered generalised or compound Poisson process if the statistics are stationary and it is a sum of a random number of independant identical distributed random variables(Parzen pg128). The scope of this thesis is limited to compound Poisson processes and the mathematical model used is based on a compound Poisson process.

2. THEORY

2.1 Revision of Previous Theoretical Work

2.1.1 The Work of Pusey on Single Interval Statistics of Light Scattered by Identical Independent Scatterers(Pusey et al, 1974)

Pusey worked on light scattered by an arbitrary no. of molecular scatterers found in dilute solutions.The number of scatterers in the solution could be determined from the moments of the probability density function of the return signal.These scatterers were assumed to be independant, uncorrelated, identical and non-interacting. They were assumed to be spherically symmetrical and small compared to the wavelength of incident light so that rotational effects were negligible.They thus scatter with uniform amplitudes and random phase $\hat{x}=1.\exp(\hat{j\theta})$ where $\hat{x}$ is a random number.

The scatterers were assumed to lie in a uniformly illuminated scattering volume which was much smaller than the sample volume. Coherence time is the time it takes one particle to move a distance 1/k where k is the scattering vector and is thus associated with fluctuations of the number of scatterers within the scattering volume. To ignore this fluctuation one needs the scattering volume large compared to 1/k. It is assumed that the total number of scatterers in the sample volume was constant.The scatterers were assumed to be distributed with a constant mean number density per scattering volume throughout the sample volume.

The assumptions of particle independance and constant mean number density implies that the number of scatterers within a scattering volume vary in a Poisson distribution throughout the sample volume(Pusey 1974,pg 535). That is according to the Poisson law $P(x=N)=\exp(-\langle N\rangle)*\langle N\rangle^{k}/k!$. Where $\langle N\rangle$ is the mean number of scatterers per scattering volume, k=0,1,2,3... and n is the number of scatterers within a particular scattering volume.

It is known by the Central limit thereom that the sum of many independant random contributions produces Gaussian statistics(Pg 266 Papoulis).Thus the returned backscatter from a large number of independant random scatterers within a scattering volume will be Gaussian. However if the number of scatterers within the scattering volume is small (approx <10) the statistics will deviate from Gaussian. The amount of deviation from Gaussian statistics will give an indication of the number of scatterers within the scattering volume. This deviation can be determined by calculating the normalised moments of the intensity of the backscattered signal. The normalised moment can be used to calculate the average number of scatterers within the volume. The Real and Imaginary components of the scattered field will be Gaussian thus the amplitude will have a Rayleigh pdf. Then the pdf of the amplitude A will be $p(A)=(A/\sigma^2)\exp(-A/2\sigma^2)$ where $I=A^2$ is the intensity and the average value of intensity $\langle I\rangle=2\sigma^2$. The intensity has an exponential pdf of the form $p(I)=(1/\langle I\rangle)\exp(-1/\langle I\rangle)$.

The n'th moment is defined as $\langle I^n\rangle=\int_{\infty}^{\infty} I^n p(I)\,dI$ where p(I) is the pdf of I

$= \sum I_i^n P(I=I_i)$ if I is of the discrete type

Equations have been derived (by Pusey et al, 1974) for $\langle N\rangle$ the mean number density in terms of the normalised moments. The first four are

$\langle I\rangle/\langle I\rangle=1$

$\langle I^2\rangle/\langle I\rangle^2=2+1/\langle N\rangle$

$\langle I^3\rangle/\langle I\rangle^3=6+9/\langle N\rangle+1/\langle N\rangle^2$

$\langle I^4\rangle/\langle I\rangle^4=24+72/\langle N\rangle+34/\langle N\rangle^2+1/\langle N\rangle^3$

Only the first two moments are required to calculate $\langle N\rangle$.

2.1.2 Extension of Pusey's work to the Underwater Acoustic Scattering Problem (Wilhelmij and Denbigh 1984, Wilhelmij 1983)

The experiments of Pusey with light are analogous to the problem of scattering from underwater sonar. Instead of a light beam there is an acoustical beam incident on underwater scatterers. The scattering volume in underwater acoustics is the resoltion cell volume. The scatterers must be in the farfield of the beam. The returned scattering signal is envelope detected and sampled to give discrete intensity points from which the normalised moments are calculated. The sonar beam diverges introducing non-stationarity into the statistics and an every increasing cell volume size. Ensamble averaging techniques are needed to handle this. In this technique samples at the same depth are used to calculate the second normalised moment at each depth. Then the the normalised moment at all depths are averaged together to produce an overall normalised moment which is used to calculate the number density.
The resolution cell volume is determined by the pulse length and beamwidth and equals $\Omega r^2 c\tau/2$ and at a range $r=ct/2$ the number per cell $\langle N\rangle(t)=.125\Omega\rho c^3\tau t^2$ (This is for an ideal conical beam while the beamshape actually has a main lobe

surrounded by side lobes.)
The number density using ensamble averaging is then calculated according to the formula $\rho = 1/((\langle I^2\rangle/\langle I\rangle^2_{av}-2).125c\Omega\tau t_1 t_m)$
Where $\langle I^2\rangle/\langle I\rangle^2_{av}$ is the average of the second normalised moments for each depth, $\Omega = 2\pi(1-\sin(\theta/2))$ is the solid angle associated with the 3dB beamwidth $= 2\theta$ and t_1 and t_m are the start and stop sampling times within a ping corresponding to the start and stop depths in water. The pulse length is τ. The speed of sound= c

There are two important criteria for the success of the method. One is the requirement already dealt with which requires a low number of scatterers per cell. Not only is it true that a low number are needed to produce deviations from Rayleigh statistics but it was also found experimentally that the lower the number per cell the more accurate the estimate of number density turns out to be.

When applied to acoustical backscatter this criterion requires a narrow beamwidth and small pulse length to make the resolution cell as small as possible giving the smallest number per cell possible for a given scatterer number density per meter cubed.A small number per cell also reduces multiple scattering which increases with increasing density.The resolution cell volume increases with depth thus to achieve a low number per cell the scatterers must be as close to the transducer as possible. The method thus works best for low scatterer number densities.

The second criterion is that there must be enough statistically independant samples for the accurate calculations of normalised moments. This is because the expressions for the moments are only approximate for discrete samples and accuracy improves with more samples. The return signal thus has to be sampled with at least one sample per cell since there is one independant sample per cell.It is also true that for a certain underwater depth range the shorter the pulse length the more resolution cells there will be within the same depth range.Thus a shorter pulse length provides more independant sample points.Collecting more envelope records also increases the number of independant points if the transducer beamwidth of one envelope does not overlap with another.

The pulse length must not be to short otherwise it will be distorted and the method will produce incorrect values for the number of scatterers per cell.

There is thus great advantage in using the shortest pulse length possible, having the transducer as close to the scatterers as possible and having as narrow a beamwidth as possible.

It was found that for a given number of independant samples the calculation of $\langle N \rangle$ using the second normalised moment was more accurate than those involving higher order moments. The error in $\langle N \rangle$ for a given error in second normalised moment or a given number of independant samples is approximately given by the expression $|(sd\langle N \rangle)/\langle N \rangle| = (sd(\langle I^2 \rangle/\langle I \rangle^2))\langle N \rangle \simeq 2\langle N \rangle/\sqrt{M}$
where M is the number of independant sample points used to calculate the second normalised moment. Computer simulations showed that the actual number of independant points for a given accuracy was higher than the value predicted by this expression.(Weintroub and Denbigh 1986, Weintroub 1986)

2.1.3 Using the Method on Fish at Sea
(Weintroub and Denbigh 1986, Weintroub 1986)

At sea the scatterers are fish instead of polystyrene spheres and the transducers were mounted on the ship hull. Transducers of frequancies 120 khz and 38 khz were used and pulse lengths of .14,.3,.6,.9 mS were used.
For experimentation at sea shoals of anchovy were used. Daytime aggregations proved unsuitable because during the day the fish concentrated themselves in dense shoals occupying a small volume. At night however due to limited visibility these shoals spread out into broad layers sometimes several miles long and wide which had a much lower number density and occupied a large enough volume to provide many sample envelopes. The results using the statistical method were compared to the results using the echo integration method. The echo integration results were calculated from the volume backscatter(VBS) and target strength(TS) using the formula

$= \frac{10^{.1(VBS-TS_g+TVG)}}{\bar{w}}$ where $TS_g = -10.9\log\bar{l} - 50.9$, TVG is the time varying gain correction factor, $\bar{w}$=mean weight of the fish and $\bar{l}$=mean length of the fish. Measurements of w and l were obtained by trawling.

The result of the comparison showed the statistical method gave much lower estimates of number density than the integration method. Part of the reason was that the fish were not uniform scatterers because they did not have the same cross section from all directions. This lead to the assumption that the statistics of a signal scattered from a single fish were Rayleigh and not uniform. This assumption changed the formula for $\langle N \rangle$ from

$\langle N \rangle = 1/(\langle I^2 \rangle / \langle I \rangle^2 - 2)$

to

$\langle N \rangle = 2/(\langle I^2 \rangle / \langle I \rangle^2 - 2)$

giving number densities twice as great.

These estimates were however still approximately three times lower than those given by the echo integration technique. Reasons for this descrepancy were reasoned to be too few statistically independant points and wrong assumption of constant mean number densities. It was clear that assumptions underlying the model had to be re-examined.

When the pulse length and hence resolution cell volume was varied during the experiment the number of fish per cell were shown to vary proportionally. This showed that the method definately was working.

An examination of assumptions underlying the model is done next in order to find the underlying reasons for the discrepencies discovered above.

2.2 Examinations of Assumptions Underlying the Model

The question of accounting for the difference between echo integration and statistical density estimates now arises.What is causing the estimates to be lower than they should be and why was the experiment succesful with polystyrene spheres and not with fish?The answer lies in considering the differance between ship/sea/fish and laboratory apparatus/water tank/polystyrene spheres and altering the model to take into account these differences.Some of the major differences looked at in this chapter are:

i)the nature of the scattering from individual fish
ii)the nature of fish distribution within the shoal
iii)multiple scattering off other fish
iv)non-uniform illuminated scatterers and uneveness in insonification
v)backscatter,reverberation and ocean noise due to seaweed etc.
vi)various experimental differances between laboratory work and ship measurements eg. in sampling,bandwidth,ship movement etc.
vii) non-stationarity in the statistics due to diverging sonar beam,attenuation with depth,and the variation of the number of scatterers with time

viii) sensitivity in one or more of the above factors to which the echo moment is more prone than the echo integration method

These factors are now dealt with in greater detail.

2.2.1 Scatterer Nature

The assumptions about the nature of the scatterers need to be reconsidered. Originally it was assumed that scatterers were independant and identical however fish in a shoal may not be identical in size or species or in orientation towards the transducer.

Furthermore the behaviour of fish swimming together may make them correlated scatterers and not independant ones since they interact and swim in the same direction. In addition, molecular particles were assumed to be spherically symmetrical and/or small compared to the incident wavelength, so that rotation effects were negligible however fish are directional scatterers -this will reduce the effective number of scatterers and give an estimate of number density which is too low. Particles were originally considered to have uniform random phase and constant scattered amplitudes. This was extended to Rayleigh scatterers with random phase. The random phase assumption may have to be investigated too. According to Clay and Heist(1984), scatter from individual fish obey Ricean statistics, which are affected by fish movement and average size. The effect of assuming Ricean statistics in the model is investigated and it is found that estimates of number density lay between those for constant and Rayleigh scatterer amplitudes.

To aid in the above investigations, a model has been derived for general scatterer statistics, which still assumes independant scatterers, enabling the effect of assumptions about scatterer statistics to be investigated.When normalised moments of intensity are plotted against number density for different scatterer probability density functions it was seen to vary for different pdf's. Thus number density is sensitive to fish scatter pdf which is dependant on fish movement,size,species etc.

2.2.2 Fish Distribution

Another major assumption questioned is the nature of fish distribution. Up to now a poisson distribution has been assumed.Later a formula for generalised fish distributions is derived, so that it can be seen how estimates of the number density are affected by variations in the fish distribution.Various possible distributions will be looked at to see how they affect the estimate of number density.If the number density is found to be sensitive to changes in fish distribution it could give estimates which are too high or too low, depending on the nature of the distribution. Thus the assumption of a Poisson distribution must be carefully considered. This assumption was partially verified by R.L.Swarts in his master's thesis, entitled "Covarience function of the acoustic backscatter from a school of fish"(Ehrenburg 1971,Ref 4). It may however not be entirely correct, since the Poisson model assumes identical, spatially independant(and therefore uncorrelated) point scatterers with constant uniform mean number density and no theoretical limit to the number of fish which can fit into a cell. If correlated(ie not distributed in a totally random manner) then the scatterer field will deviate from Rayleigh statistics. In practise there will be a limit to the no of fish which can physically occupy the resolution cell volume. The Poisson model also assumes non-interacting scatterers. If scatterers interact this will cause deviations from the Poisson model. If the scatterer positions are not such that a path difference exceeding the wavelength of the incident wave is introduced, errors will result.This will happen when fish are too close together since two fish will be detected as one.

Another factor concerning distribution is the possibility that fish shoals have a mean density which is not constant throughout the shoal. Thus the assumption of constant mean density which is implicit in the Poisson model does not hold. As an attempt to overcome this problem the fish shoal could be divided into a number of density regions each with approxiamately constant mean number density.This will only work if the assumption that there is a poisson distribution with constant mean number density within each density region.

2.2.3 Multiple Scattering

There is also the problem of multiple scattering from one fish onto another.Multiple scattering occurs at high fish densities and will give a still higher(and erroneous) estimate of fish number density.This will cause errors particularly when the number of fish per cell approaches ten.Multiple scattering increases with increasing fish size,number of fish per cell and depth of scatterers.

2.2.4 Non-uniform illumination of Scatterers

One also needs to have uniformly illuminated scattering volume but fish in different positions of a sonar beam are illuminated with different intensities. Non-uniform illumination due to position in the beam becomes more critical at greater depths.There is also uneveness in insonification due to scatterering of the beam by fish nearer the transducer.This will also be a greater problem at greater depths and densities.

2.2.5 Backscatter, Reverberation and Ocean Noise

Backscatter from other objects, reverberation and ocean noise will also produce errors in the estimate. The effect of these factors on the model will not be investigated in this thesis.

2.2.6 Experimental Conditions

Sampling intervals need to be short (and Bandwidth as high as possible) compared to the fluctuation time of the scattered intensity and one needs as many independant samples as possible

2.2.7 Non-Stationarity Effects

There is also non-stationarity due to a diverging transducer beam and attenuation and the variation of the number of scatterers with time(fish movement relative to boat).Ways of including this effect in the model will be looked at.

2.2.8 Peculiar Sensitivities

The statistical method may be more sensitive than echo integration methods to the seven factors above.The statistical method may have problems in areas which the other two methods do not have, since it may be sensitive to things that the other two methods are not. These "things" are more likely to be those that effect the statistics of the backscatter, such as noise etc. Echo integration integrates noise and thus reduces the effect of noise but there is no similar smoothing when looking at backscatter statistics since this would remove the statistical information. Hence backscatter statistics would be more sensitive to ship movement and rolling and reverberation noise etc. than echo integration.

2.3 Derivation of Generalised Model

It is necessary to derive a more generalised statistical model than has been done before using a different method than that used by Pusey(1974) or Denbigh and Weintroub(1986). This is to have a model which holds for any scatterer type and any volume distribution. It is necessary to obtain generalised expressions for the second, third and fourth normalised intensities which are $<I^2>/<I>^2, <I^3>/<I>^3, <I^4>/<I>^4$ respectively. This derivation uses a new and different approach than previous derivations. This was necessary as previous derivations could not be used to yield generalised expressions for any scatterer type and any scatterer distribution throughout the volume. Previous work on Rayleigh scatterer pdf could not go beyond the second normalised moment. Using the new approach it is possible to calculate higher order moments for Rayleigh scatterers as well.

The number of fish in any cell is assumed to be an integral amount so that all fish are completely in the cell. In any resolution cell there will be a random number of fish N in the cell determined by the distribution pdf $P_{<N>}(N)$. Each fish will scatter acoustic waves in a random manner with random amplitude C_i and random phase θ_i. The in phase and quadrature components will add up to produce a resulting signal for the cell. For a cell containing N fish the amplitude is given by

(sum of the in phase components)+j(sum of the quadrature components)

Thus $A_N=(\sum_{i=1}^{N} C_i\cos\theta_i)+j(\sum_{i=1}^{N} C_i\sin\theta_i)=\sum_{i=1}^{N} C_i\exp(j\theta_i)$

The intensity $I_N=|A_N|^2=|\sum_{i=1}^{N} C_i\exp(j\theta_i)|^2=|\sum_{i=1}^{N}\exp z_i|^2=(\sum_{i=1}^{N}\exp z_i)(\sum_{i=1}^{N}\exp \bar{z}_i)$

where $z_i=\ln C_i+j\theta_i=x_i+jy_i$

Taking average values of both sides $<I_N>=<(\sum_{i=1}^{N}\exp z_i)(\sum_{i=1}^{N}\exp \bar{z}_i)>=<\sum_{i=1}^{N}\exp(z_i+\bar{z}_i)>$

for i=k, $z_i=z_k$ and $z_i+z_k=x_i+jy_i+x_i-jy_i=2x_i$

for $i\neq k$, $<\exp(z_i+\bar{z}_k)>=0$

Thus $<I_N>=<\sum_{i=1}^{N}\exp(2x_i)>=<\sum_{i=1}^{N} C_i^2>=\sum_{i=1}^{N}<C_i^2>$ since N and C_i are independent random variables.

Further if all fish are identical and independent $<C_1^2>=<C_2^2>=<C_3^2>=.....$

Letting all $<C_i^2>=<a^2>$ then $<I_N>=\sum_{i=1}^{N}<C_i^2>=N<C_i^2>=N<a^2>$ for each cell. All resolution cells have a discrete random number of fish in them given by $P_{<N>}(N)$. Summing all the intensity returns $<I_N>$ from all cells in the volume gives a total average value of intensity of $<I>=\sum_{i=1}^{N}<I_N>P_{<N>}(N)=\sum_{i=1}^{N}N<a^2>P_{<N>}(N)=<a^2>\sum_{i=1}^{N}NP_{<N>}(N)=<a^2><N>$ where $<N>$ is the average value of fish in a cell and $<N>=\sum_{i=1}^{N}NP_{<N>}(N)$

In the same way $I_N^n=(\sum_{i=1}^N \exp z_i)^n(\sum_{i=1}^N \exp \bar{z}_i)^n$

To evaluate $(\sum_{i=1}^N \exp z_i)^n$ use the Multinomial Expansion (pg 4 Spiegel)

$$(\sum_{i=1}^N x_i)^n = \xi n!/(n_1!n_2!n_3!\ldots\ldots n\ !)x_1^{n_1}x_2^{n_2}\ldots\ldots x_N^{n_N} = \xi M_{N,n}x_1^{n_1}x_2^{n_2}\ldots x_N^{n_N}$$

where $n_1,n_2,\ldots n_N \varepsilon\{0,1,2,\ldots\ldots n\}$ and $\sum_{i=1}^N n_i=n$

Hence $(\sum_{i=1}^N \exp z_i)^n = \xi M_{N,n}(\exp z_1)^{n_1}(\exp z_2)^{n_2}\ldots\ldots(\exp z_N)^{n_N}$

$$= \xi M_{N,n}\exp(n_1z_1+n_2z_2+n_3z_3+\ldots\ldots+n_Nz_N)$$

Now $<(\sum_{i=1}^N \exp z_i)^n(\sum_{i=1}^N \exp \bar{z}_i)^n> = \xi M^2_{N,n}<\exp(2n_1x_1+2n_2x_2+\ldots\ldots+2n_Nx_N)>$

since terms like $z_i+\bar{z}_i=x_i+jy_i+(x_i-jy_i)=2x_i$ and other terms have average value=0

(Henceforth $\Sigma=\sum_{i=1}^N$ and ξ is used for the Multinomial Expansion

Thus $<I_N^n>=\xi M_{N,n}<\exp(2n_1x_1+2n_2x_2+\ldots\ldots+2n_Nx_N)>$

$$= \xi M^2_{N,n}<\exp(2n_1x_1).\exp(2n_2x_2)\ldots\ldots\exp(2n_Nx_N)>$$

but $\exp(2n_ix_i)=(\exp^2x_i)^{n_i}=(C_i^2)^{n_i}$

therefore $<I_N^n>=\xi M^2_{N,n}<(C_1^2)^{n_1}><(C_2^2)^{n_2}>\ldots\ldots<(C_N^2)^{n_N}>$ if all fish are independant

since $E(x_1x_2x_3\ldots\ldots x_N)=E(x_1)E(x_2)E(x_3)\ldots\ldots E(x_N)$ if $x_1,x_2,x_3\ldots\ldots x_N$

are independant random variables (Pg 211 Papoulis)

and if all fish are identical $<C_1^2>=<C_2^2>=\ldots\ldots=<C_N^2>=<a^2>$ since they will have

the same average amplitude.

therefore $<I_N^n>=\xi M^2_{N,n}\prod_{i=1}^N<(a^2)^{n_i}>$

expressions for n=2,3,4 are obtained in appendix 1.

From appendix 1

$$<I_N^2>=N<a^4>+2<a^2>^2N(N-1)$$

$$\langle I_N^3\rangle=N\langle a^6\rangle+3^2\langle a^2\rangle\langle a^4\rangle N(N-1)+6\langle a^2\rangle^3 N(N-1)(N-2)$$

$$\langle I_N^4\rangle=N\langle a^8\rangle+(16\langle a^2\rangle\langle a^6\rangle+18\langle a^4\rangle^2)N(N-1)+72\langle a^2\rangle^2\langle a^4\rangle N(N-1)(N-2)$$

$$+24\langle a^2\rangle^4 N(N-1)(N-2)(N-3)$$

Summing $\langle I_N^2\rangle$ for all cells in the sample volume with distribution pdf $P_{\langle N\rangle}(N)$

gives $\langle I^2\rangle=\sum_{N=1}^{\infty}\langle I_N^2\rangle P_{\langle N\rangle}(N)$

$$=\langle a^4\rangle\sum_{N=1}^{\infty} NP_{\langle N\rangle}(N)+2\langle a^2\rangle^2(\sum_{N=1}^{\infty} N(N-1)P_{\langle N\rangle}(N))$$

$$=\langle a^4\rangle\langle N\rangle+2\langle a^2\rangle^2(\sum_{N=1}^{\infty} N^2P_{\langle N\rangle}(N)+\sum_{N=1}^{\infty} NP_{\langle N\rangle}(N))$$

$$=\langle a^4\rangle\langle N\rangle+2\langle a^2\rangle^2(\langle N^2\rangle-\langle N\rangle)$$

$$=(\langle a^4\rangle-2\langle a^2\rangle^2)\langle N\rangle+2\langle a^2\rangle^2\langle N\rangle$$

Therefore $\langle I^2\rangle/\langle I\rangle^2=((\langle a^4\rangle-2\langle a^2\rangle^2)\langle N\rangle)/(\langle a^2\rangle^2\langle N\rangle^2)+(2\langle a^2\rangle^2\langle N^2\rangle)/(\langle a^2\rangle^2\langle N\rangle^2)$

$$=((\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle+2\langle N^2\rangle/\langle N\rangle^2$$

This is a completely general formula for the second normalised moment for all random,independant,identical scatterers distributed with a discrete pdf in the sample volume.

Similarly $\langle I^3\rangle=\sum_{N=1}^{\infty}\langle I_N^3\rangle P_{\langle N\rangle}(N)$

$$=\langle a^6\rangle\langle N\rangle+9\langle a^2\rangle\langle a^4\rangle(\langle N^2\rangle-\langle N\rangle)+6\langle a^2\rangle^3(\langle N^3\rangle-3\langle N^2\rangle+2\langle N\rangle)$$

Therefore

$$\langle I^3\rangle/\langle I\rangle^3=(\langle a^6\rangle/\langle a^2\rangle^3-9\langle a^4\rangle/\langle a^2\rangle^2+12)/\langle N\rangle^2+(9\langle a^4\rangle/\langle a^2\rangle^2-18)\langle N^2\rangle/\langle N\rangle^3+6\langle N^3\rangle/\langle N\rangle^3$$

and

$$\langle I^4\rangle=\langle a^8\rangle\langle N\rangle+(16\langle a^2\rangle\langle a^6\rangle+18\langle a^4\rangle^2)(\langle N^2\rangle-\langle N\rangle)+72\langle a^2\rangle^2\langle a^4\rangle(\langle N^3\rangle-3\langle N^2\rangle+2\langle N\rangle)$$

$$+24\langle a^2\rangle^4(\langle N^4\rangle-6\langle N^3\rangle+11\langle N^2\rangle-6\langle N\rangle)$$

Therefore

$$\langle I^4\rangle/\langle I\rangle^4=(\langle a^8\rangle/\langle a^2\rangle^4-16\langle a^6\rangle/\langle a^2\rangle^3-18\langle a^4\rangle^2/\langle a^2\rangle^4+144\langle a^4\rangle/\langle a^2\rangle^2-144)/\langle N\rangle^3$$

$$+(16\langle a^6\rangle/\langle a^2\rangle^3+18\langle a^4\rangle^2/\langle a^2\rangle^4-216\langle a^4\rangle/\langle a^2\rangle^2+264)\langle N^2\rangle/\langle N\rangle^4$$

$$+(72\langle a^4\rangle/\langle a^2\rangle^2-144)\langle N^3\rangle/\langle N\rangle^4+24\langle N^4\rangle/\langle N\rangle^4$$

2.3.1 Application of Generalised Model to Poisson Volume Distribution and Various Scatterer Amplitude Statistics

In this section to see the effect of various scatterer types on the result a Poisson distribution is assumed and then various scatterer amplitudes are compared.

For a Poisson distribution $P(x=N)=\exp(-\langle N\rangle)\langle N\rangle^{k}/k!$ $k=0,1,2\ldots\ldots$

with $\langle N^2\rangle=\langle N\rangle^2+\langle N\rangle$

$$\langle N^3\rangle=\langle N\rangle^3+3\langle N\rangle^2+\langle N\rangle$$

$$\langle N^4\rangle=\langle N\rangle^4+6\langle N\rangle^3+7\langle N\rangle^2+\langle N\rangle$$ see appendix 3

Applying this to the generalised formula yields

$$\langle I^2\rangle/\langle I\rangle^2=(\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle\ +2\langle N^2\rangle/\langle N\rangle^2$$

$$=(\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle\ +2(\langle N\rangle^2+\langle N\rangle)/\langle N\rangle^2$$

$$=(\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle-2$$

Therefore $\langle N\rangle=(\langle a^4\rangle/\langle a^2\rangle^2)/(\langle I^2\rangle/\langle I\rangle^2-2)$

Similarly $\langle I^3\rangle/\langle I\rangle^3=(\langle a^6\rangle/\langle a^2\rangle^3)/\langle N\rangle^2+9(\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+6$

$$\langle I^4\rangle/\langle I^4\rangle=(\langle a^8\rangle/\langle a^2\rangle^4)/\langle N\rangle^3+(16\langle a^6\rangle/\langle a^2\rangle^3+18\langle a^4\rangle^2/\langle a^2\rangle^4)/\langle N\rangle^2$$

$$+(72\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+24$$

2.3.1.1 Application to Constant Scatterer Amplitude

Constant scatterer amplitude was dealt with by Pusey(1974) and experimentally tested on polystyrene spheres for acoustics by Willhelmij(1983).

For a constant scatterer amplitude $\langle a^{2k}\rangle/\langle a^{k}\rangle^2=1$ $k=1,2,3\ldots\ldots$

Thus the above equations become $\langle I^2\rangle/\langle I\rangle^2=1/\langle N\rangle+2$ and $\langle N\rangle=1/(\langle I^2\rangle/\langle I\rangle^2-2)$

$$\langle I^3\rangle/\langle I\rangle^3=1/\langle N\rangle^2+9/\langle N\rangle+6$$

$$\langle I^4\rangle/\langle I\rangle^4 = 1/\langle N\rangle^3+34/\langle N\rangle^2+72/\langle N\rangle+24$$

These equations agree with those Pusey derived by a different method.

2.3.1.2 Rayleigh Amplitude Statistics

For Rayleigh pdf $P(a)=a/\sigma^2\exp(-a^2/2\sigma^2)$ and $\langle a^{2k}\rangle=2^k k!\alpha^{2k}$ for k=1,2,3....

Thus the equations become $\langle I^2\rangle/\langle I\rangle^2= 2!(1/\langle N\rangle+1)$ and $\langle N\rangle=2/(\langle I^2\rangle/\langle I\rangle^2-2)$

$$\langle I^3\rangle/\langle I\rangle^3= 3!(1/\langle N\rangle^2+3/\langle N\rangle+1)$$

$$\langle I^4\rangle/\langle I\rangle^4= 4!(1/\langle N\rangle^3+7/\langle N\rangle^2+6/\langle N\rangle+1)$$

Thus for Rayleigh amplitudes the estimate of number of fish per cell is now twice as great.The number of fish per cell that can be accurately measured is greater for Rayleigh than for Constant amlitude scatter since from the graph the second normalised moment for Rayleigh scatterers tend to the limit of 2! at much higher densities than constant amplitude scatterers do. It is when the second normalised moment tends to 2! that small errors in the calculation of second normalised moment can have a large effect on <N> (Wilhelmij 1983).

2.3.1.3 Ricean Statistics

According to Clay and Heist(1984) acoustic backscatter from individual fish obey Ricean statistics.This is because backscatter is made up of a concentrated component and distributed component.The concentrated component is caused by the fish bladder and the distributed component is caused by the fish body which swishes to and fro.This distributed component introduces noise onto the concentrated component which is analogous to noise on a sine wave.This problem was dealt with by Rice(1944,1945)Clay and Heist give the pdf of the envelope as

$$w_R(a)=(2a/\sigma_d)\exp(-(a^2+\sigma_c)/\sigma_d)I_0((2a\sqrt{\sigma_c})/\sigma_d)$$

$$=(2a(1+\gamma)/\sigma_{bs})\exp(-(1+\gamma)a^2+\gamma\sigma_{bs})/\sigma_{bs})I_0(2a\sqrt{(\gamma(1+\gamma)}/\sqrt{\sigma_{bs}}))$$

where σ_c=power associated with the concentrated component

σ_d=power associated with the distributed component

and $\sigma_{bs}=\sigma_c+\sigma_d$ $\gamma=\sigma_c/\sigma_d$

From appendix 2 $\langle a^2\rangle=\sigma_d(\gamma+1)$

$$<a^4>=\sigma_d^2(\gamma^2+4\gamma+2)$$

$$<a^6>=\sigma_d^3(\gamma^3+9\gamma^2+18\gamma+6)$$

$$<a^8>=\sigma_d^4(\gamma^4+16\gamma^3+72\gamma^2+96\gamma+24)$$

Thus the equations become

$<I^2>/<I>^2=((\gamma^2+4\gamma+2)/(\gamma+1)^2)/<N>\ +2$ and $<N>=((\gamma^2+4\gamma+2)/(\gamma+1)^2)/(<I^2>/<I>^2-2)$

$$<I^3>/<I>^3=((\gamma^3+9\gamma^2+18\gamma+6)/(\gamma+1)^3)/<N>^2+9((\gamma^2+4\gamma+2)/(\gamma+1)^2)/<N>+6$$

$$<I^4>/<I>^4=((\gamma^4+16\gamma^3+72\gamma^2+96\gamma+24)/(\gamma+1)^4)/<N>^3$$

$$+((16(\gamma+1)(\gamma^3+9\gamma^2+18\gamma+6)+18(\gamma^2+4\gamma+2)^2)/(\gamma+1)^4)/<N>^2$$

$$+72((\gamma^2+4\gamma+2)/(\gamma+1)^2)/<N>+24$$

It is interesting to note that for $\gamma=0$ the above equations yield the equations for Rayleigh scatter .For $\gamma=\infty$ the equations reduce to constant scatterer equations.In particular for the equation

$<N>=(<a^4>/<a^2>^2)/(<I^2>/<I>^2-2)$ with $<a^4>/<a^2>^2=(\gamma^2+4\gamma+2)/(\gamma+1)^2$

For small fish or fish with no swimbladder $\sigma_c>>\sigma_d$ and $\sigma_{bs}\cong\sigma_d$ then $\gamma=0$ and $<a^4>/<a^2>^2=2$ as for Rayleigh amplitude statistics
For fish with large swimbladders and for polystyrene spheres $\sigma_c<<\sigma_d$
$\sigma_{bs}=\sigma_c$ when $\gamma=\infty$,$<a^4>/<a^2>^2=1$ as for constant amplitude statistics
Thus for Ricean scatter $1<(<a^4>/<a^2>^2)<2$
See graphs for comparison of different scatterer statistics ie for $\gamma=0,\gamma=\infty$ (fig 2.3.1)

2.3.2 Application of generalised model to various distribution statistics
To illustrate the various distributions the expression for the second normalised moments of intensity
$<I^2>/<I>^2=(<a^4>/<a^2>^2-2)/<N>+2<N^2>/<N>^2$ is used.

2.3.2.1 Poisson distribution
For a Poisson distribution $<N^2>/<N>^2=(<N>^2+<N>)/<N>^2=1+1/<N>$
This distribution has already been dealt with in section 2.3.1 . For the Poisson distribution the expression for second normalised moment becomes
$<I^2>/<I>^2=(<a^4>/<a^2>^2)/<N>+2$

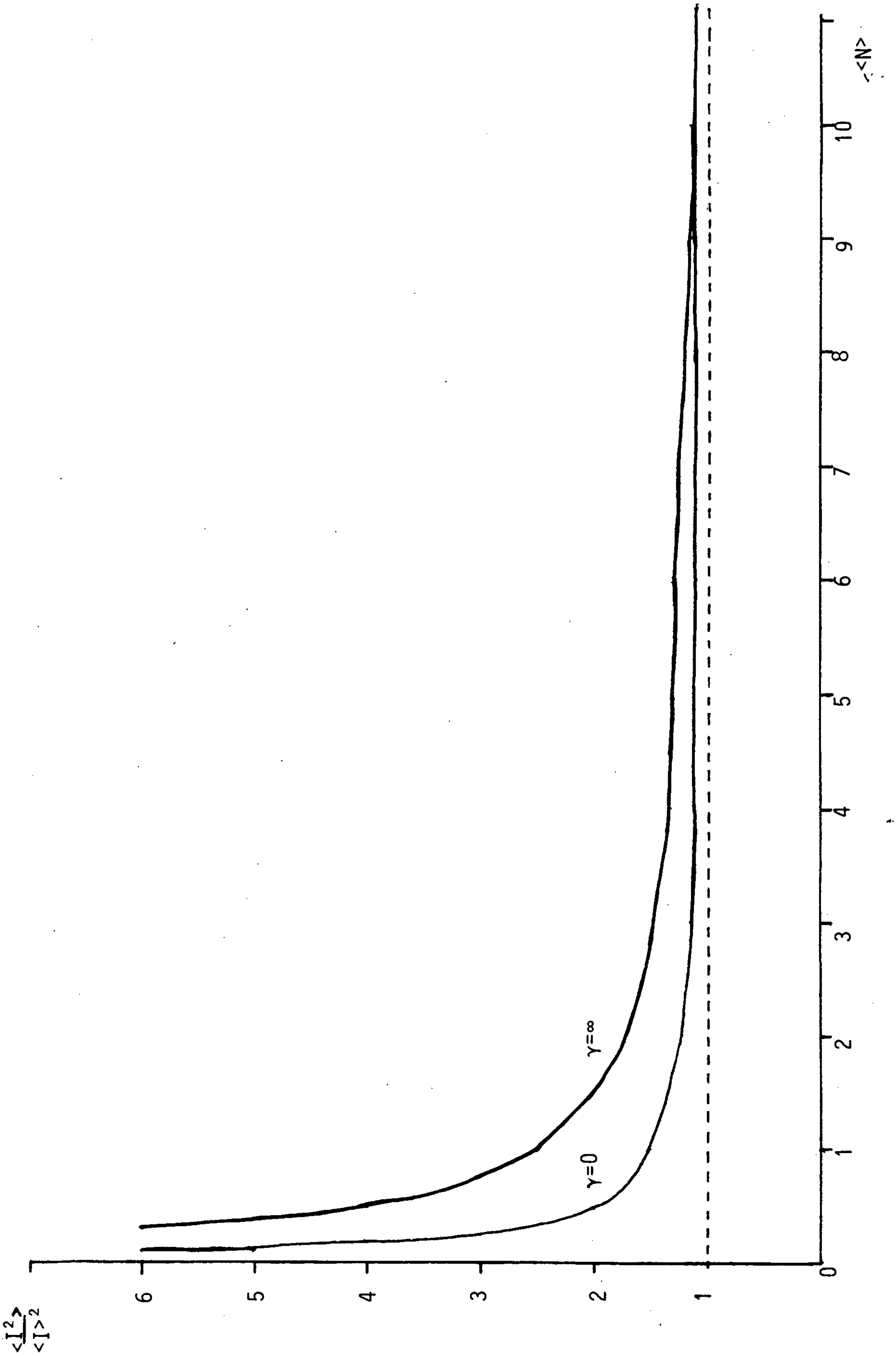

Fig 2.3.1 Graph of scatterer statistics for γ =0 and γ=∞

2.3.2.2 Binomial distribution

If there is a limit to the number of fish that can fit in a cell the Binomial distribution would be more appropriate.

From appendix 3.2 for a Binomial Distribution $\langle N^2\rangle=\langle N\rangle^2-\langle N\rangle\langle N\rangle/n+\langle N\rangle$

Thus $\langle N^2\rangle/\langle N\rangle^2_B=(\langle N\rangle^2-\langle N\rangle\langle N\rangle/n+\langle N\rangle)/\langle N\rangle^2=1-1/n+1/\langle N\rangle\;=\langle N^2\rangle/\langle N\rangle^2_P-1/n$

So $\langle I^2\rangle/\langle I\rangle^2_B=\langle I^2\rangle/\langle I\rangle^2_P-2/n$ Thus $(\langle I^2\rangle/\langle I\rangle^2)_B\leq(\langle I^2\rangle/\langle I\rangle^2)_P$

and $\langle I^2\rangle/\langle I\rangle^2=(\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle+2(1-1/n+1/\langle N\rangle)=(\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle-2/n+2$

where n is the maximum number of fish that can be found in a cell.

Therefore $\langle N\rangle=(\langle a^4\rangle/\langle a^2\rangle^2)/(\langle I^2\rangle/\langle I\rangle^2+2/n-2)$

Notice that $\langle N\rangle_B<\langle N\rangle_P$ for the same experimental value of second normalised moment.

2.3.2.3 Uniform distribution

If the fish in a shoal is evenly spread out then a uniform distribution is more appropriate.

From appendix 3.3 for a Uniform Distribution $\langle N^2\rangle=4\langle N\rangle^2/3+\langle N\rangle/3$

For a uniform distribution $\langle N^2\rangle/\langle N\rangle^2_U=(4\langle N\rangle^2/3+\langle N\rangle/3)/\langle N\rangle^2=4/3+1/(3\langle N\rangle)$

$$=1+(\langle N^2\rangle/\langle N\rangle^2)_P/3$$

$$=1+(\langle N^2\rangle/\langle N\rangle^2)_P-2(\langle N^2\rangle/\langle N\rangle^2)_P/3$$

$$=1+(\langle N^2\rangle/\langle N\rangle^2)_P-2(1+1/\langle N\rangle)/3$$

$$=(\langle N^2\rangle/\langle N\rangle^2)_P+1/3-2/(3\langle N\rangle)$$

Thus $\langle I^2\rangle/\langle I\rangle^2_U=(\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle+2(\langle N^2\rangle/\langle N\rangle^2)_U$

$$=(\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle+2(\langle N^2\rangle/\langle N\rangle^2)_P+2\;1/3-2/(3\langle N\rangle)$$

$$=\langle I^2\rangle/\langle I\rangle^2_P+2(1/3-2/(3\langle N\rangle))$$

and $\langle I^2\rangle/\langle I\rangle^2=(\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle+2(4/3+1/(3\langle N\rangle))$

$$=(\langle a^4\rangle/\langle a^2\rangle^2-4/3)\langle N\rangle+8/3$$

Therefore $\langle N\rangle =(\langle a^4\rangle/\langle a^2\rangle^2-4/3)/(\langle I^2\rangle/\langle I\rangle^2-8/3)$

$(\langle I^2\rangle/\langle I\rangle^2_U)>(\langle I^2\rangle/\langle I\rangle^2_P)$ for $1/3-2/(3\langle N\rangle)>0$ ie for $\langle N\rangle > 2$

$(\langle I^2\rangle/\langle I\rangle^2_U)<(\langle I^2\rangle/\langle I\rangle^2_P)$ for $\langle N\rangle < 2$

2.3.2.4 Constant Density

For a constant density throughout shoal each cell will have the same number of fish in it.The number of fish N in any cell is fixed. Thus for any cell the probability that the number of fish $\hat{x}$ equals N is absolutely certain or $P(\hat{x}=N)=1$ and $P(\hat{x}\neq N)=0$. Thus $\langle N\rangle=E(\hat{x})=0.P(\hat{x}=0)+1.P(\hat{x}=1)+.....N.P(\hat{x}=N)=N$

and $\langle N^2\rangle=E(x^2)=0^2.P(\hat{x}=0)+1^2.P(\hat{x}=1)+......+N^2.P(\hat{x}=N)=N^2=\langle N\rangle^2$

Therefore $\langle N^2\rangle/\langle N\rangle^2=\langle N\rangle^2/\langle N\rangle^2=1$ and $\langle I^2\rangle/\langle I\rangle^2=(\langle a^4\rangle/\langle a^2\rangle^2-2)/\langle N\rangle+2$

Therefore $\langle N\rangle=(\langle a^4\rangle/\langle a^2\rangle^2-2)/(\langle I^2\rangle/\langle I\rangle^2-2)$

A comparison of different distributions of Rayleigh scatterers is shown in Fig2.3.2 It is also possible to do combinations of Poisson and Constant(or any other) density distributions. This is justifiable since the Poisson distribution has only been partially verified for fish shoals by R.L.Swarts(Ehrenburg 1971,Ref 4)

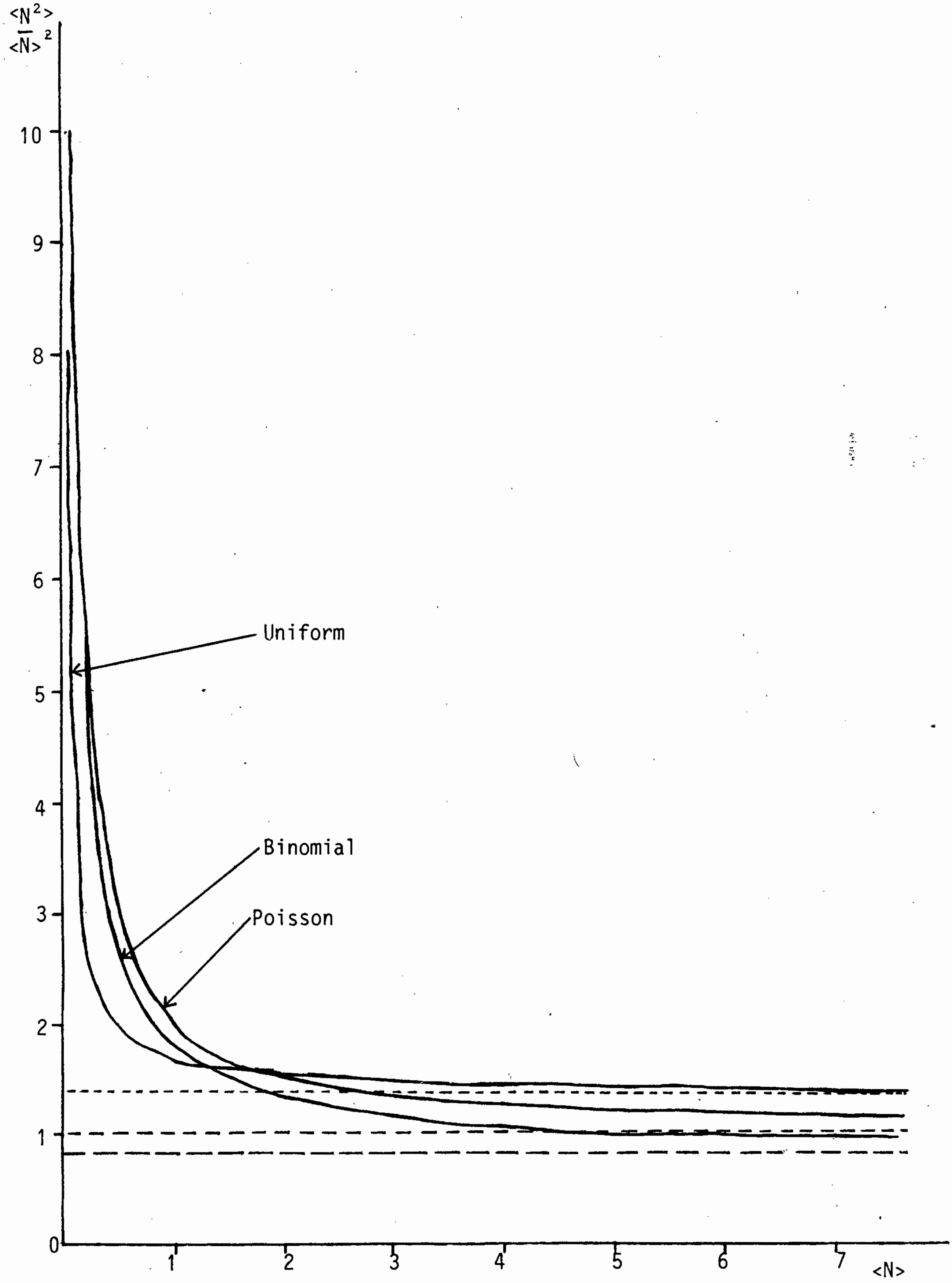

Fig 2.3.2 Graph of second normalised moment of N versus <N>

2.4 Effects of Beam Pattern

Up to now an ideal beamshape has been assumed. For a non-ideal beamshape the resultant amplitude $\hat{e}=\hat{a}.D^2(\theta,\phi)$ (Peterson et al 1976,Pg 619) .
Putting $b=D^2(\theta,\phi)$,$\hat{e}$ becomes the product of two random functions ie. $\hat{e}=\hat{a}.\hat{b}$.
If one assumes that $\hat{a}$ and $\hat{b}$ are independant random variables then
$E(e^n)=E(a^n).E(b^n)$ or $<e^n>=<a^n><b^n>$ (pg 211 Papoulis).
Therefore $<e^4>/<e^2>^2=(<a^4>/<a^2>^2)(<b^4>/<b^2>^2)$
For Rayleigh scatter $<a^4>/<a^2>^2=2$ for Ricean $<a^4>/<a^2>^2=(\gamma^2+4\gamma+1)/(\gamma+1)^2= 0$
Hence $<N>_e=<e^4>/<e^2>^2/(<I^2>/<I>^2-2)=(<a^4>/<a^2>^2)(<b^4>/<b^2>^2)/(<I^2>/<I>^2-2)$
Where $<N>_e$ is the corrected value of number per cell
The quantity $<b^4>/<b^2>^2$ is calculated for a hemisphere ie for a resolution cell with solid angle 2π.Hence $<N>_e$ is the number of fish in a cell of solid angle 2π with size $2\pi r_g^2(c\tau/2)=\pi r_g^2 c\tau$ where r_g is the geometric mean of the start and stop depths.
$<N>_e=(<b^4>/<b^2>^2)((\gamma^2+4\gamma+1)/(\gamma+1)^2)/(<I^2>/<I>^2-2)=(<b^4>/<b^2>^2)<N>$
Hence the number of fish per cubic meter
$\rho =((<b^4>/<b^2>^2)/(\pi r_g^2 c\tau))<N>=<N>/(\pi r_g^2 c\tau/(<b^4>/<b^2>^2))$
equivalent cell volume$=\pi r_g^2 c\tau/(<b^4>/<b^2>^2)$ as the beamshape can be regarded as modifying the cell volume.

2.4.1 The circular piston transducer

For a circular piston transducer of diameter d
$D(\phi)=2J_1(\pi(d/\lambda)\sin\phi)/(\pi(d/\lambda)\sin\phi)$ where λ is the wavelength of the signal.
The transducer used has effective diameter d=8.28cm .The transducer is a 120 khz so $\lambda=c/f=1.25$cm where c=1500 m/s therefore $d/\lambda=8.28/1.25=6.624$.
From appendix 4 for $d/\lambda=6.624$ $<b^4>/<b^2>^2=245$(This agrees well with the value $<b^4>/<b^2>^2=200$ given for $d/\lambda=6$ in Ehrenburgs thesis(Ehrenburg 1973,APPB)) so the equivalent cell volume size $=\pi r_g^2 c\tau/245=.0128 r_g^2 c\tau$.
Thus $\rho=<N>/(.0128 r_g^2 c\tau)$ $=<N>/(19.2 r_g^2 \tau)$ **for c=1500**

2.4.2 The square piston transducer

For a square piston transducer of dimensions W,H
$D(\theta,\phi)= \sin(\pi(W/\lambda)\sin\phi)/(\pi(W/\lambda)\sin\phi) . \sin(\pi(H/\lambda)\sin\theta)/(\pi(H/\lambda)\sin\theta)$
where W is the width and H the height of the piston.
The 38khz transducer used has $W/\lambda= 6$ and $H/\lambda= 6$ from appendix 4 this yields $<b^4>/<b^2>^2=237$ .Thus equivalent cell volume$=\pi r_g^2 c\tau/237$ $=.0133 r_g^2 c\tau$
Thus the fish number density $\rho=<N>/(.0133 r_g^2 c\tau)=<N>/(19.95 r_g^2 \tau)$

2.5 The effect of the number of measurements on the estimate of number density

As the return echo amplitude is sampled to produce one independant sample per cell there is only a limited number of independant samples for the calculation of the second normalised moment of intensity.This limited number of independant measurements will produce an error in the estimate of <N>.
The standard deviation for both constant amplitude and rayleigh scattering have been derived in appendix 5.This derivation does not assume an infinite <N> and is thus more accurate than previous derivations of standard deviations by Wilhelmij and Denbigh (1984).
From appendix 5 the formula for Rayleigh scatter and a Poisson distribution is
$\mathrm{sd}\ \langle I^2\rangle/\langle I\rangle^2 = ((4(2/\langle N\rangle^3+17/\langle N\rangle^2+10/\langle N\rangle+1)/M)$
$\mathrm{sd}\ \langle N\rangle = \langle N\rangle^2\sqrt{((2/\langle N\rangle^3+17/\langle N\rangle^2+10/\langle N\rangle+1)/M)}$
Where sd is an abreviation for the standard deviation of...
For constant amplitude scatter and a Poisson distribution the formula are
$\mathrm{sd}\ \langle I^2\rangle/\langle I\rangle^2 = \sqrt{((4+20/\langle N\rangle+13/\langle N\rangle^2+1/\langle N\rangle^3)/M)}$
$\mathrm{sd}\ \langle N\rangle = \langle N\rangle^2\sqrt{((1/\langle N\rangle^3+13/\langle N\rangle^2+20/\langle N\rangle+4)/M)}$
These predictions of standard deviation can be compared with the simulations done by Weintroub (1986).
On page 4-15 Weintroub produced a table of results of a simulation of the return echo from a fish shoal with <N>=2 assuming Rayleigh scatterers.
From the formula above $\mathrm{sd}\ \langle I^2\rangle/\langle I\rangle^2 = \sqrt{(42/M)} = 6.48/\sqrt{M}$
$\mathrm{sd}\ \langle N\rangle = 12.96/\sqrt{M}$
Table 2.5.1 compares the simulated results with those calculated by formula for calculations done with different values of M.

M	Mean est of <N>	Simulated sd <N>	Calculated sd <N>	Difference	Difference as a percentage of <N>=2
100	3.1	1.85	1.296	.554	27.00 %
200	2.59	0.90	0.916	.016	0.80 %
400	2.36	0.63	0.648	.018	0.90 %
800	2.2	0.46	0.458	.002	0.10 %
1000	2.11	0.41	0.409	.001	0.05 %
1400	2.08	0.31	0.346	.036	1.80 %
2000	2.03	0.28	0.289	.009	0.45 %
2500	2.09	0.21	0.259	.049	2.45 %
3000	2.05	0.19	0.236	.046	2.30 %

Table 2.5.1 Comparison of standard deviation from calculation by formula with simulation for <N>=2 for various values of M

In addition there is a simulation for <N>=4 and M=1500 on pg A-7 which yields a sd $<I^2>/<I>^2$=.12652 and a sd <N>=.98796 .The calculated values of standard deviation for <N>=4 and M=1500 are sd $<I^2>/<I>^2$= (18.375/1500)=.1106 and sd <N>=.8854 .The difference between the simulated and calculated values of sd <N>=.102 or 2.5 percent of <N>=4.

Weintroub also does simulations for different values of <N> for Rayleigh scatterers on pg 4-13 .

Table 2.5.2 compares the simulated and calculated sd for M=500 and M=1000 for Rayleigh scatterers.

True <N>	Simul sd <N>	Calcul sd <N>	Diff	Diff as per of <N>
.1	.014	.0275	.0135	13.5 %
.5	.078	.114	.036	7.2 %
1.0	.197	.245	.048	4.8 %
2.0	.616	.579	.037	1.8 %
4.0	2.04	1.53	.51	12.7 %

Table 2.5.2(a) M=500

True <N>	Simul sd <N>	Calcul sd <N>	Diff	Diff as per of <N>
.1	.011	0.019	.008	8 %
.5	.068	0.081	.013	2.6 %
1.0	.150	0.173	.023	2.3 %
2.0	.369	0.409	.04	2.0 %
4.0	1.14	1.08	.06	1.5 %

Table 2.5.2(b) M=1000

Table 2.5.3 does the same but for Constant amplitude scatterers using simulation from pg 4-12 of Weintroub(1986)

True <N>	Simul sd <N>	Calcul sd <N>	Diff	Diff as per of <N>
.1	.007	.022	.015	15 %
.5	.071	.114	.043	8 %
1.0	.205	.275	.07	7 %
2.0	1.51	.774	.736	18 %
4.0	58.0	2.24	55.7	1394 %

Table 2.5.3 (a) M=500

True <N>	Simul sd <N>	Calcul sd <N>	Diff	Diff as per of <N>
.1	.006	.0158	.0098	9.8 %
.5	.053	.080	.027	5.4 %
1.0	.157	.195	.038	3.8 %
2.0	.507	.526	.019	.95 %
4.0	5.49	1.584	3.9	97 %

Table 2.5.3 (b) M=1000

Thus the simulated values are quite close to the calculated values for larger values of M and for larger values of <N>. It can be concluded that the formulae are quite accurate for $M > 100$ and $<N> > .5$ for Rayleigh scatterers. It is also fairly accurate for constant amplitude scatterers though for $<N>=4$ comparison is difficult since the simulated values of sd are unreliable. It is also true that the calculated values are closer to the simulated values for Rayleigh than for Constant amplitude scatter.
Thus for Rayleigh scatterers it is possible to have higher values of <N> within a resolution cell than for constant amplitude scatterers for the same accuracy in the result. Thus data can be collected on Rayleigh scatterers at higher densities than for constant scatterers.

2.6 Method of Dealing With Inhomogeneities in The Fish Shoal

According to the Poisson distribution assumptions there is a constant mean number density in the shoal but in practise varying densities occur within a fish shoal.To discover these density variations it is neccesary to determine the average density in a certain region and compare it with nearby regions. The average intensity is given by $<I>=<a^2><N>$ .Thus the number per cell <N> and therefore the density is directly proportional to the average intensity within a region.If the average intensity varies for different regions then the shoal does not have a constant mean number density.Since there could be density variations with depth and with distance travelled ie both horizontally and vertically in a shoal the intensity data points needed to calculate <I> must include both points in a vertical and horizontal direction to determine the density at a certain point within the shoal.For this reason a block of data points is chosen.(see fig 2.6.1)

The length and width of this data block must be chosen to reflect density variation with depth and length.This may involve trial and error but density variation can be guaged from the echo chart recording to give some idea how much density variation there is with depth and length to make the dimensions X and Y easier to choose. Preferably X and Y dimensions of the block must be as small as possible to reflect as much density variation as possible. The block can however not be made too small as there will be too few statistically independant data points within it. This will result in an error in the equation $<I>=<a^2><N>$ since <I> needs to be calculated with sufficient independant data points to prevent errors.The number of independant points needed and the percentage error is derived in appendix 4. The table summarising the results is reproduced here in table 2.6.1.

<N>	% s.d. in <I> for M=100
.1	45 %
.5	22 %
1.0	17 %
2.0	14 %
4.0	12 %

Table 2.6.1(a)

M	% s.d in <I> for <N>=2
50	20 %
100	14 %
200	10 %
300	8 %
500	6 %
1000	4 %

Table 2.6.1(b)

The number of independant points in depth is determined by the number of resolution cells there are in the distance Y.Each resolution cell contributes one independant point. The number of independant points with length is determined by the number of pings within the distance X since the points in one ping are independant of the points in adjacent pings.The block is situated around a data point.The average intensity <I> then gives a measure of the average density surrounding the data point. The average intensity <I> is calculated using the approximate expression $\langle I\rangle=\sum_{i=1}^{M} I_i/M$

The average intensity surrounding each data point collected can thus be calculated. Data points falling in regions of like density are grouped together to evaluate the second normalised moment of intensity since then the criterion that only points of constant mean number density are used is met.Because density varies with depth and length (ie density variations are spatial) points of like density occur in two dimensional regions or bands of constant density when looking at a vertical transect of the fish shoal(fig 2.6.2 for a hypothetical example).All points within the same region are used in the calculation of second normalised moment.The regions will vary from those of high density to those of low density as in Fig2.6.2

The choice of the number of regions and which regions to include points in is to a certain extent arbitrary. The criterion which most needs to be met is that there must be enough statistically independant points within the region and the density of the highest region must have less than 10 fish per cell for the accurate determination of the second normalised moment.Apart from that as many regions as possible are chosen. Another criterion which has to be met is that the sd<I> for the block must be much less than the difference between the upper

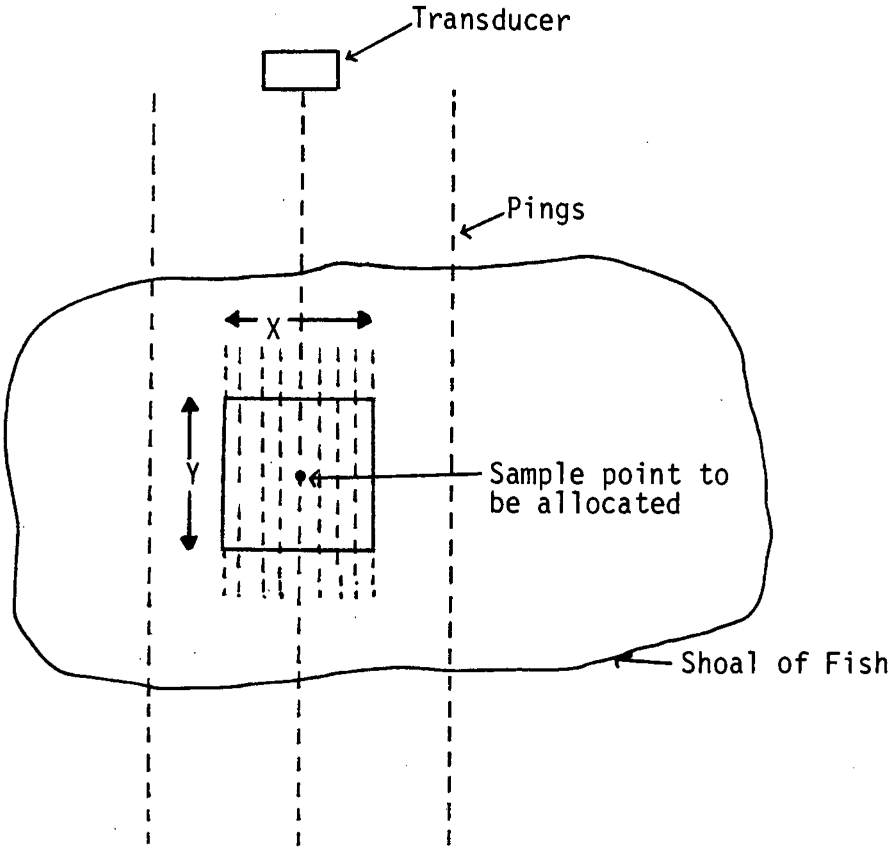

Fig 2.6.1 Method of determining average densities around sample point for allocation to a density band

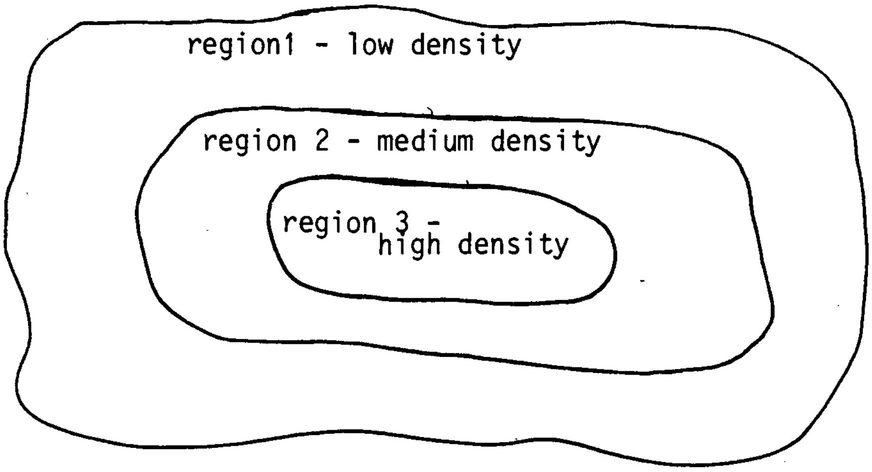

Fig 2.6.2 Density regions within a fish shoal

and lower limits of $<I>$ which are allocated to the region. This will be dealt with in section 2.6.2

2.6.1 Choice of Reference Density Levels

To decide which density regions or bands the points are to be allocated a number of reference levels of average intensity (hence density since $<I>=<a^2><N>$) are set .The region between two adjacent levels is called a density band. If $<I>$ for a block surrounding a point falls between two adjacent reference levels then its intensity and its square are summed to other points falling in the same density band.

Finally the sums are used to calculate the second normalised moment for each band using the approximation $<I^2>/<I>^2=M\sum_{i=1}^{M} I_i^2/(\sum_{i=1}^{M} I_i)^2$ .The number per cell for each band is then calculated using the formula $<N>=2/(<I^2>/<I>^2-2)$ for Rayleigh scatterers.

The width of the density band is an arbitrary choice and difficult to establish criterion for. Ideally it should reflect the density variations within a region and be chosen so that an approximately equal no of points fall in each band. In practise it is easiest to choose equal widths for each band and to arrange the bands symmetrically about the mean value of intensity for the whole shoal to ensure that all bands have roughly an equal number of points. In practise this works reasonably well. The mean value of intensity = I_{av}

For n bands there will be (n+1) reference levels. The first reference level will be at 0.Equidistant reference levels are assumed to be symmetrical about I_{av} ie. equal number of levels on both sides of I_{av} the last level being at $2I_{av}$ .Since there are n equidistant bands all bands must be included within $2I_{av}$. Each level will thus be seperated by $2I_{av}/n$.

The first level L_0 will be at 0 therefore $L_0=0$

the second level L_1 will be at $2I_{av}/n$ therefore $L_1=2I_{av}/n$

the third level L_2 will be at $2I_{av}/n+2I_{av}/n=(2I_{av}/n)2$ therefore $L_2=(2I_{av}/n)2$

hence $L_k=(2I_{av}/n)k$

and $L_n=(2I_{av}/n)n=2I_{av}$

Hence the formula used to allocate reference levels is $L_k=(2I_{av}/n)k$ where n is the total number of bands or regions L_k is the k'th reference level and k= 0,1,2.......n

The second normalised moment $<I^2>/<I>^2=M\sum_{i=1} I_i^2/(\sum_{i=1} I_i)^2$ for each region k

M =number of points in density band k

and for Rayleigh scatterers and Poisson distribution the average number per cell $<N>=2/(<I^2>/<I>^2-2)$

2.6.2 Choice of Block Window Size and Density Levels With Regard to Standard Deviation of Estimates

The s.d. in $\langle I \rangle$ is half the minimum seperation which needs to exist between different intensity levels ie is the minimum width of the intensity band. Thus choosing fewer independant points in a'block forces one to choose fewer density bands which are wider in width.

$\mathrm{s.d.} \ll (L_{k+1} - L_k)/2$ for a minimum error to prevent mis-allocation of intensity points.

When dealing with a number of levels where $\langle I_k \rangle$ is an intensity in the middle of band k often there is a roughly linear variation in density then $\langle I_2 \rangle = 3\langle I_1 \rangle, \langle I_3 \rangle = 5\langle I_1 \rangle, \ldots\ldots \langle I_k \rangle = (2k-1)\langle I_1 \rangle$ and hence $\langle N_k \rangle = (2k-1)\langle N_1 \rangle$.

It is obvious that for larger $\langle I \rangle$ the s.d. tends to be larger because of the increase in $\langle I \rangle$ to which it is proportional.The fact that the percentage s.d. decreases with increasing $\langle I \rangle$ will tend to counteract this slightly.
Since the formula for $\mathrm{s.d.} = \langle I \rangle \sqrt{(2/\langle N \rangle + 1))}$ it follows that

$$\mathrm{s.d.}_k/\mathrm{s.d.}_1 = \langle I_k \rangle / \langle I_1 \rangle \sqrt{((2/\langle N_k \rangle + 1)/(2/\langle N_1 \rangle + 1))}$$

$$= (2k-1)\sqrt{((2/((2k-1)\langle N_1 \rangle) + 1)/(2/\langle N_1 \rangle + 1))}$$

$$= (2k-1)\sqrt{((2/(2k-1) + \langle N_1 \rangle)/(2 + \langle N_1 \rangle))}$$

$$= \sqrt{((2 + \langle N_1 \rangle (2k-1)^2)/(2 + \langle N_1 \rangle))}$$

Alternatively $\mathrm{sd}_k/\mathrm{sd}_1 = (2k-1)\sqrt{((2/\langle N_k \rangle + 1)/(2/(\langle N_k \rangle/(2k-1)) + 1))}$

$$= (2k-1)\sqrt{((2 + \langle N_k \rangle)/(2(2k-1) + \langle N_k \rangle))}$$

It is obvious that $\mathrm{sd} \geq \mathrm{sd}_1$ since $(2k-1) > 1$ for $k > 1$ therefore the higher bands will always have the highest sd and as $\langle N_1 \rangle$ increases the ratio $\mathrm{sd}_k/\mathrm{sd}_1$ increases still further.
Table 2.6.2 gives the magnitude of s.d. for higher intensities relative to the s.d. for lower intensities for the hypothetical case where $\langle I_k \rangle = (2k-1)\langle I_1 \rangle$ and $\langle N_k \rangle = (2k-1)\langle N_1 \rangle$ and assuming $\langle N_1 \rangle = 1$.

k	sd_k/sd_1
1	1
2	$\sqrt{5}=2.2$
3	$\sqrt{(35/3)}=3.4$
4	$\sqrt{21}=4.5$
5	$\sqrt{33}=5.7$
6	$\sqrt{(143/3)}=6.9$

Table 2.6.2

This shows approximately how much wider the higher level must be than the lowest one in order to keep the statistical errors the same for each level. It shows that a great number of bands is not a practicle possibility and that the k'th band needs to be approximately k times wider than the first band to keep the statistical errors the same in both bands.

Thus the requirement that $sd_k<<(L_{k+1}-L_k)/2$ is most stringent for the higher density bands so one should at least have $sd_n<<(L_{n+1}-L_n)/2$
Hence one needs to determine the value of sd_n.To do this it is neccessary to know $<N_n>$ and $<I_n>$ which would have to be guessed or estimated from a trial calculation of number densities .Once sd_n has been found it could then be decided to make the width of all other bands the same as the n'th band or make all bands of such a width that they would have their own $sd_k<<(L_{k+1}-L_k)/2$.This means that the higher density bands will be wider than the lower density band ie $L_{k+2}-L_{k+1}>L_{k+1}-L_k$ and the simpler scheme with $L_k=(2I_{av}/n)k$ would no longer apply.

As a rough guide again asume roughly linear variations in density and $sd_k/sd_1=\sqrt{((2+<N_1>(2k-1)^2)/(2+<N_1>))}$ and thus $sd_k>sd_{k-1}>.........>sd_1$
Choosing the width of each density band to be proportional to the sd of intensity in that band ie $L_{k+1}-L_k=c.sd_k$ so

$$L_1=L_0+c.sd_1$$
$$L_2=L_1+c.sd_2$$

$$L_{k+1}=L_k+c.sd_k$$

therefore

$$\begin{aligned}L_{k+1}&=L_0+c.sd_1+c.sd_2.........+c.sd_k\\&=L_0+c.\sum_{i=1}^{k}sd_i\\&=L_0+c.sd_1\sum_{i=1}^{k}sd_i/sd_1\\&=L_0+c.<I_1>\sqrt{(2/<N_1>+1)}\sum_{i=1}^{k}((2+<N_1>(2i-1)^2)/(2+<N_1>))\end{aligned}$$

since $sd_1=<I_1>\sqrt{(2/<N_1>+1)}$
$<I_1>$ and $<N_1>$ can be estimated from a trial run and usually $L_0=0$.The value $c.sd_1$ can be made a fraction of the average intensity of the whole shoal.

Alternatively $L_{k+1}=L_k+c.sd_k\sum_{i=1}^{k}sd_i/sd_k$

and $sd_i/sd_k=\sqrt{((2/<N_i>+1)/(2/<N_k>+1))}$

$=\sqrt{(((2/<N_k>)(2k-1)/(2i-1)+1)/(2/<N_k>+1))}$

since $<N_i>=(2i-1)<N_1>=((2i-1)/(2k-1))<N_k>$

hence $sd_k/sd_1=\sqrt{((2(2k-1)/(2i-1)+<N_k>)/(2+<N_k>))}$

therefore $L_{k+}=L_k+c.sd_k\sum_{i=1}^{k}((2(2k-1)/(2i-1)+<N_k>)/(2+<N_k>))$

For the analysis of data in this thesis equidistant levels have been chosen. This means that the sd<I> will be higher for higher density bands and that there will be greater chance of data points being misallocated in higher density bands than in lower density bands.

2.6.3 The Calculation of Overall Density for a Shoal Consisting of a Number of Different Density Bands

Consider the volume occupied by the fish shoal ie the sample volume to be made up of a number of equal sized cells or volumes ie the scattering volume. Hence the unit volume is the cell which is assumed constant throughout. Each region has a certain number of data points in it M_k and a average number per cell $<N_k>$ associated with it.Each sample point is surrounded by a certain volume v which is constant for all sample points in the sample volume if it is assumed that boat speed,pulse length,sampling rate are all constant.

Thus the total volume in density region k $V_k=M_kv$

Hence forth $\Sigma\equiv\sum_{i=1}^{n}$

and the total volume of all regions together $V_T=\Sigma V_k=\Sigma M_kv=v\Sigma M_k$

The total number of fish in region k=number per unit volume x volume of region

$$=<N_k>V_k=<N_k>vM_k$$

Total number of fish in shoal$=\Sigma<N_k>V_k=\Sigma<N_k>vM_k=v\Sigma<N_k>M_k$

Therefore the average number per unit volume $<N_k>$ for the whole shoal or sample volume is given by the

total number of fish in shoal/total volume of shoal$=v\Sigma<N_k>M_k/(v\Sigma M_k)$

Therefore $<N_{av}>=\Sigma<N_k>M_k/\Sigma M_k$

From appendix 5.6 $sd<N_{av}>=\surd(\Sigma M_k^2(sd<N_k>)^2)/(\Sigma M_k)$

2.6.4 Ensamble averaging

Over wider depth ranges ensamble averaging has to be used to remove the effect of beam divergence. This entails calculating $<I^2>/<I>^2$ for each subdivision of depth range (within a density region) which in the data used in this thesis is assumed narrow enough so that ensamble averaging need not be used. Thus the greater depth range is divided into smaller depth ranges which are narrow enough so that the effect of beam divergence can be ignored. (The number of points within each smaller depth range of a density region is a subset of the total number of points in that region and so may not always be sufficient to calculate $<I^2>/<I>^2$ for each smaller depth range in a region.) The value of $<I^2>/<I>^2$ for all smaller depth ranges within a region of constant density are averaged together to produce an average value of $<I^2>/<I>^2_{avk}$ for the whole density region.

If $<I^2>/<I>^2_{y,k}$ =second normalised intensity moment for y'th depth level and k'th density region. Where $<I^2>/<I>^2_{y,k}=(\Sigma I^2/M_{yk})/(\Sigma I/M_{yk})^2=M_{yk}(\Sigma I/(\Sigma I)^2)_{yk}$ and M_{yk} is the number of data points in the smaller depth range y in a region k. $<I^2>/<I>^2_{avk}$ for each density band k =average value of all values of $<I^2>/<I>^2_{yk}$ for each smaller depth range.

$<I^2>/<I>^2_{avk}=(\sum_y<I^2>/<I>_{yk})/\sum_y 1=\sum_y M_{yk}(\Sigma I^2/(\Sigma I)^2)_{yk}$

Therefore $<I^2>/<I>^2_{avk}=(\sum_y M_{yk}(\Sigma I^2/(\Sigma I)^2)_{yk})/\sum_y 1$

where Σ is the sum of all points at same smaller depth range and in same density region and $\sum_y$ is the sum of all moments in each density region at same smaller depth range.

As non-stationary statistics are outside the scope of this thesis ensamble averaging is not used in the analysis of data. The assumption that the depth range is narrow enough is questionable and may lead to errors in the analysis. If the work of this thesis can be extended to non stationary statistics it may be able to further improve the estimates of number density by taking beam divergence into account.

2.7 Comparison of Theory of Echo Integration with that of Statistical Method

2.7.1 Theory of Echo Integration

The derivation from first principles of the equation used for echo-integration is adapted from H.Bodolt(1969).

Let the beam pattern of the transducer be $b(\theta,\phi)$ and the axial intensity at unit distance be I_0 (The source level $SL=10\log I_0$).

Thus the intensity at 1 meter in the direction (θ,ϕ) is $I_0 b(\theta,\phi)$

Thus the incident intensity on the scattering volume dV will be $I_0 b(\theta,\phi)\exp(-\beta r)/r^2$ where β is the attenuation constant

The intensity of the sound back-scattered by dV will be $I_0 b(\theta,\phi)\exp(-\beta r)s_v/r^2 dV$ where s_v is the volume backscattered strength defined as the ratio of sound intensity scattered back in a direction towards the source by a unit volume to the intensity of the incident plane wave.(Both intensities being referred to a distance 1 metre from the unit volume)

At the transducer the reverberation from dV will have the intensity $I_0 b^2(\theta,\phi)\exp(-2\beta r)s_v/r^4 dV$ and will produce a mean-squared voltage output of $R^2 I_0 b^2(\theta,\phi)\exp(-2\beta r)s_v/r^4 dV$ where R is the voltage response of the receiver.

Integration(summation) over the volume yields

$V^2_{rms}=R^2 I_0\exp(-2\beta r)s_v/r^4\int b^2(\theta,\phi)dV$

The elemental volume is $dV=r^2 c\tau/2 d\Omega$

Therefore $V^2_{rms}=R^2 I_0\exp(-2\beta r)s_v c\tau/(2r^4)\int b^2(\theta,\phi)d\Omega$

$\int b^2(\theta,\phi)d\Omega$ is the equivalent beamwidth of the transducer and will be called Ψ

Taking the average value of both sides

$\overline{V}^2_{rms}=R^2 I_0\exp(-2\beta r)\overline{s}_v c\tau\Psi/(2r^4)\equiv$ backscattered intensity(Weintroub 1986)

Taking logarithms

$10\log\overline{V}^2_{rms}=20\log R+10\log I_0-(20\log r+2\alpha r)+10\log(c\tau/2)+10\log\Psi+10\log\overline{s}_v$ (dB)

where $\alpha=-\beta\log_{10}e$ and $s_v=I_{scat}/I_0 \;=\sigma_f\rho$ since the intensity scattered back from dV will be a product of the number of fish in dV(ie the density ρ) and the intensity scattered back from a single fish $=I_0\,\sigma_f$ where σ_f is the target strength of an individual fish.Thus $s_v=\sigma_f\rho$ .Taking average values $\overline{s}_v=\overline{\sigma}_f\overline{\rho}$

Thus $\overline{\rho}=\overline{s}_v/\overline{\sigma}_f=10^{.1(10\log(\overline{s}_v/\overline{\sigma}_f))}=10^{.1(VBS-TS)}$ were $VBS=10\log_{10}\overline{s}_v$ and $TS=10\log_{10}\overline{\sigma}_f$

This is the method used to calculate the echo-integration estimates of number density later.It should also be noted that density variations are automatically averaged out when using the echo integration method since the average of s_v will be the average of the backscatter from all density regions.ie if there were two densities $\rho_1$and ρ_2 then the $s_v=\sigma_f(\rho_1+\rho_2)=\rho_{av}\sigma_f$ where $\rho_{av}$is the overall average density.Thus the overall average density is directly obtained from the backscattered signal. In the same way the backscatter s_v is insensitive to two different target strengths within the shoal.

2.7.2 Comparison with the theory of the Statistical Method

The derivation of the statistical model has already been explained in section 2.3 for stationary statistics. It must be noted that the echo integration method does contain the effects of beam divergence and attenuation with distance. For the statistical model it was always assumed that the scattering layer was thin enough to ignore beam divergence and attenuation with distance was negligible.These assumptions will have to be checked very carefully as the $1/r^4$ effect is a very basic and universal effect. Both beam divergence and attenuation will have to be incorporated in the statistical model to make it more realistic. When making comparisons between the theory of the two methods it is seen how this can be done.

In the integration method $\bar{s}_v=I_{scat}/I_0 \ =\bar{\sigma}_f\bar{\rho}$ while in the statistical method $<I>=<a^2><N>$.By comparing these two equations it can be seen that

$<a^2>\equiv\bar{\sigma}_f$;$<N>\equiv\bar{\rho}$;$<I>\equiv\overline{V}^2_{rms}\propto\bar{s}_v$

Thus to include the effects of non-stationarity it is neccessary to multiply $<a^2><N>$ by $\exp(-2\beta r)/r^4$ for the first moment of intensity $<I>$. A similar expression could also be derived for the second moment $<I^2>$which also includes non-stationarity effects. It will then be possible to obtain a second normalised moment of intensity which includes the non-stationarity effects.

If the fish are in an infinately thin scattering layer at constant depth r then the second normalised moment of intensity remains unchanged but in practise the scattering layer is not infinately thin and hence it's width and depth will be of importance in calculating the second normalised moment of intensity and will yield a different expression for the second normalised moment of intensity.

It can also be seen that the statistical model is actually an extension of the echo-integration model. Whereas the echo-integration method uses only the first normalised moment the statistical method uses the first,second and higher order moments.

As noted previously the target strength is equivalent to $\langle a^2 \rangle$ .The target strength is directly related to the weight and length of the fish caught by the formula $TS_g=-10.9\log\overline{l}-50.9$ where TS_g is the target strenth per gram of fish whereas for the statistical method $\langle a^2\rangle=\sigma_d(\gamma+1)=\sigma_d(\sigma_c/\sigma_d+1)=\sigma_c+\sigma_d$. The two quantities $\langle a^2\rangle$and $\overline{\sigma}_f$may not be the same when evaluated for comparison of the two methods and may be responsible for differences when comparing estimates produced by the two methods. In addition γ could only be estimated since determining γ was outside the scope of this thesis since it would have involved extensive experimentation on the scatterer statistics from individual live fish of various species.

Both methods incorporate the effects of beam pattern. Furthermore any other effects such as multiple scattering can be included in the statistical model in the same way that they are included in the echo integration method since the former is an extension of the latter. In addition for comparison the mean squared error for the statistical method can be calculated and compared to that for the echo integration method which was calculated by Ehrenberg and Lytle(1972). The mean squared error is defined as $\overline{e^2}=E((N-E(N))^2)$ .It could then be seen how the accuracy of the methods compare at different densities.

3. EXPERIMENTAL DATA ANALYSIS(USING DATA COLLECTED IN 1985)

The data in this chapter is mostly on night time anchovy shoals since at night the shoal disperses and has a low density. During the day the shoal is concentrated in a small high density volume that the ship quickly passes over. This is too high a density for the method to work and too little data is collected to give enough statistically independant points. It is for these reasons that night time data is preferred.

3.1 Examination of the data(Histograms, Contour plots)

The data collected by Weintroub had to be examined for regional density variations to see if there was any density variation within the shoal. This was done by block averaging which has already been outlined in section 2.6 which involves taking the average of surrounding points and allocating that point to a density band depending on the average intensity around that point.

A two dimensional plot of the allocated band for each sample point in the pings was made with each ping vertical thus giving an echo chart type representation of the fish shoal which showed the density band to which each sample point was allocated. In the plot it could be seen that sample points did have differing values of surrounding intensity and hence there was not a constant density throughout the shoal. In addition the points allocated to the same density band were found to be grouped together in contour like regions within the shoal.Thus density was seen to vary with depth and along track. Whith density regions There may be a difference between <I> and the intensity at the middle of the band. This may indicate that not all points have been correctly assigned to the appropriate band or the difference may be due to the standard deviation in <I>. Since these two values are approximately the same in most cases either could be used to calculate the sd<I> for each band. For this analysis I have used <I> to calculate sd<I> . The problem of all formulas of standard deviation when applied in this way is that they must be calculated in terms of the true values of <N> and <I> but all that is known are the estimates of <N> and <I> from the data which are inaccurate due to the statistical errors in them.

The use of <N> and <I> from the statistical analysis of data will introduce an error into the calculation of standard deviation. Hence the standard deviation itself is only approximately known but it is good to have an estimation or rough figure for standard deviation all the same.If the errors in <N> are not too large then they shouldn't make too much difference to the calculation of sd<N>.Otherwise one would have to estimate <I> and <N> for each density band and use these values to calculate the sd<N>. <I> for each band could then be taken to be the intensity in the middle of the band and <N> could be the estimated value of <N> assuming that <N> goes up in the ratio 1:3:5:7 for the density bands.

For a hypothetical Intensity point with intensity in the middle of the band would require a sd<I> /width of band >.5 for it to be misallocated to another band .

If it is assumed that a point has intensity equal to the mean ie η then the greatest possible value of <I> by calculation due to finite M is $\eta + \sigma$ and its smallest is $\eta - \sigma$ where σ =sd<I> Therefore $L_{k+1} - L_k > 2\sigma$

Of course there is still a probability that some points will fall into another band but this is low.This probability could be calculated.A band would have to be infinately wide to have all points falling in it. The misallocation of points will cause a further error in <N> for each band-it may be possible to calculate the error in <N> due to misallocation if the width of the band and sd<I> is known.

3.2 Results for density analysis

After the discovery of variations in density in the shoals the analysis method given in section 2.6 was applied. A fortran program was written to implement this method. The program is listed in Appendix 6.

The averaging block chosen spanned 9 pings and 19 sample points

Since each cell gives an independant point the number of independant points = 9*19/no.of samples per cell=171/no of samples per cell

Exp	pulse length	samples per cell	points in block	independant samples in block	<I>	width of band
1	.14	11.2	171	15	1275	637
2A	.31	12.4	171	14	2821	1410
2B	.68	27.2	171	6.28	3875	1937
3A	.68	27.2	171	6.28	811	405
3B	.86	34.4	171	4.97	1103	551
4	.8	32	171	5.34	3240	1550
5	.8	16	171	10.7	2330	1550
6A	.68	5.44	171	31.4	1465	696
6B	.86	6.88	171	24.8	1707	890

Table 3.2.1-Table giving some experimental data used in the ensuing calculations

Here-after the following notation is used sd_{wb}=sd<I>*100/width of band IP=statistically independant points.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.39	451	250	64%	160	25%	.094
2	.93	172	860	46%	395	62%	.384
3	1.93	195	1408	37%	521	81%	.884
4	2.91	186	2338	33%	784	123%	1.583

$<N_{av}>$=1.249 +/- .349

Table 3.2.2-Table of analysis for experiment 1. The sd_{wb} is greater than the maximum 50% for levels 2,3,4 which means there is a great deal of misallocation of sample points in the partitioning of densities . This suggests that the values of <N> for levels 2,3,4 are in error due to misallocation.In addition the sd<N> are quite high for bands 2,3,4 when compared with level 1 due to the low number of independant points in those bands.

Band	$<N>$	IP	$<I>$	$\%sd<I>$	$sd<I>$	sd_{wb}	$sd<N>$
1	.39	226	223	64%	142	22%	.133
2	.97	88	1020	45%	460	72%	.568
3	2.19	98	1526	35%	544	85%	1.48
4	2.17	91	2190	36%	784	123%	1.51

$<N_{av}>=1.167+/-.414$

Table 3.2.3-Odd pings from experiment 1. Here again sd_{wb} is too high for 2,3,4 and since the number of independant points for odd pings is half the number for all pings together the $sd<N>$ are high.

Band	$<N>$	IP	$<I>$	$\%sd<I>$	$sd<I>$	sd_{wb}	$sd<N>$
1	.40	225	277	63%	175	27%	.136
2	1.05	85	694	44%	306	48%	.629
3	1.73	97	1290	38%	489	77%	1.084
4	4.04	95	2481	31%	783	123%	3.575

$<N_{av}>=1.454+/-.717$

Table 3.2.4-Even pings from exp1. The discrepancy in estimates of $<N>$ between even and odd pings and all pings can be seen to be due to a reduced number of independant points in the analysis of even and odd pings of exp 1 which increases the standard deviation.Otherwise it can be seen that the results are remarkably consistant.(**Denbigh and Smith,1987**)

Band	$<N>$	IP	$<I>$	$\%sd<I>$	$sd<I>$	sd_{wb}	$sd<N>$
1	.24	175	424	82%	349	25%	.096
2	.37	86	2423	68%	1656	117%	.202
3	.67	46	3644	53%	1959	139%	.511
4	1.92	74	7942	38%	3034	215%	1.424

$<N_{av}>=.647 +/-.290$

Table 3.2.5 Results for exp2A. These results are highly inaccurate because of the large sd_{wb} and large $sd<N>$ The large $sd<N>$ is due to the low number of independant points and the large sd_{wb} due to the low number of independant points in the averaging block.(Note exp 2A and 2B are from the same shoal but are done with different pulse lengths).

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	2.38	66	969	54%	523	27%	2.01
2	1.46	165	2906	61%	1773	91%	.67
3	1.98	118	4843	56%	2712	140%	1.17
4	1.16	90	6780	65%	4407	227%	.685

$<N_{av}>$=1.676+/-.523

Table 3.2.6 Results for exp 2B. Results are inconclusive because of high standard deviation

Band	< N >	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.744	400	154	37%	56.8	14%	.196
2	1.25	346	534	31%	165	40%	.385
3	1.41	263	1171	29%	349	86%	.513
4	.729	208	1946	37%	722	178%	.265

$<N_{av}>$=1.031+/-.174

Table 3.2.7 Results for Exp 3A These results are more reliable only level 3 and 4 have high sd_{wb} and the sd<N> are not too large.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	2.36	65	326	61%	199	36%	2.0
2	3.58	49	741	56%	415	75%	4.15
3	2.40	44	1288	61%	782	142%	2.48
4	1.60	62	2701	67%	1818	330%	1.22

$<N_{av}>$=2.424+/-1.250

Table 3.2.8 Results for exp 3B. These results are inconclusive because of the large sd_{wb} and large sd<N> due to the increase in pulse length over 3A there are less independant points all round.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	4.30	63	1136	40%	458	29%	4.79
2	5.23	209	2315	39%	907	58%	3.56
3	3.05	159	3877	42%	1663	107%	1.83
4	3.49	90	5738	41%	2399	154%	2.95

$<N_{av}>$=4.153+/-1.71

Table 3.2.9 Results for exp 4. Results are inconclusive because of the large sd_{wb} for bands 2,3,4 and large sd<N>.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd$_{wb}$	sd<N>
1	1.03	90	579	53%	304	19%	.59
2	1.984	231	1643	43%	711	46%	.84
3	1.432	141	2462	47%	1166	75%	.71
4	2.01	128	4644	43%	2005	129%	1.15

<N_{av}>=1.71+/-.46

Table 3.2.10 Results for exp 5. Results are inconclusive in bands 3 and 4 due to large sd$_{wb}$.This is why there seems to be a drop in <N> going from band 2 to 3 when you would epect it to increase.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd$_{wb}$	sd<N>
1	1.85	298	446	26%	115	16%	.677
2	3.95	430	1003	22%	219	31%	1.622
3	5.51	397	1728	21%	360	52%	2.80
4	5.8	279	2894	20%	599	86%	3.62

<N_{av}>=4.311+/-1.188

Table 3.2.11 Results for Exp 6A . Results are good except in band 4 because it has a large sd$_{wb}$.This is why <N> is not appreciably larger in band 4 than band three.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd$_{wb}$	sd<N>
1	4.86	184	658	24%	157	17%	3.39
2	8.69	465	1170	22%	260	29%	5.40
3	7.24	263	2161	22%	490	55%	5.32
4	4.96	198	3335	24%	793	89%	3.37

<N_{av}>=7.05+/-2.71

Table 3.2.12 Results for exp 6B. Results are inconclusive because of the large sd<N> in all levels and the large sd$_{wb}$ in level 3,4 .Note 6B is for the same shoal as 6A but with an increased pulselength .The negative effects of increasing pulselength can clearly be seen.

It can be noted in all these experiments that missalocation due to high sd$_{wb}$ results in the estimate of <N> not increasing as one goes to higher density bands.

Note although each level has a large standard deviation the final result has a small standard deviation(perhaps because the final result includes more independant points)

To see the effect of block size and number of independant points on accuracy of calculations 171 block points and four levels were chosen for all experiments in order to compare the effect of different pulse lengths.
With encouraging self consistancy and good agreement of 1 and 6 it was decided to get more data with many more independant points.
It was done this way to tie up the theory of errors in result due to number of independant points and block size with the practise of choosing block size.

3.3 Effect of varying number of density bands

Too few density levels will not seperate out the data into homogenous density regions. Too many will leave regions with too few independant points and require very large block sizes to ensure that a point is not allocated to a wrong density level because of the error due to a limited number of independant points in calculating it.
In expt 1 have varied the number of density bands to see if there is an optimum amount of density levels. Too few levels would not be able to reflect the density variations within a shoal since there would be density variations within the density bands. However there is a limit on the number of density bands that can be used because this will limit the number of independant points in each band thus increasing the errors in the estimate of number density due to high sd<N> . It also requires larger block sizes to keep the sd<I> down to less than the width of band.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.680	1006	967	51%	495	8.8%	.1118

$<N_{av}>$=.680+/-.111

Table 3.3.1 Results for 1 band.Despite the low sd<N>,<N> is incorrect because different density regions have not been taken into account.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.428	624	419	61%	258	9%	.087
2	2.091	382	1862	36%	672	24%	.703

$<N_{av}>$=1.059+/-.272

Table 3.3.2 Results for 2 bands. Here again there is not enough density bands.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.414	518	279	62%	1165	18%	.093
2	1.397	233	1224	40%	752	52%	.536
3	2.683	255	2129	34%	639	77%	1.209

$\langle N_{av} \rangle$=1.216+/-.333

Table 3.3.3 Results for 3 bands. Now there are enough density bands but it can be noticed that as the number of bands increase the sd_{wb} increases. It is too large for band 3.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.39	451	250	64%	160	25%	.094
2	.93	172	860	46%	395	62%	.384
3	1.93	195	1408	37%	521	81%	.884
4	2.91	186	2338	33%	784	123%	1.583

$\langle N_{av} \rangle$=1.249+/-.349

Table 3.3.4 Results for 4 levels . The sd_{wb} is too large for bands 2,3,4

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.536	398	211	56%	118	10.5%	.138
2	.661	155	585	52%	303	27%	.276
3	1.312	146	1184	41%	486	43%	.628
4	2.545	160	1717	34%	592	52%	1.419
5	2.923	147	2385	33%	799	71%	1.794

$\langle N_{av} \rangle$=1.335 +/- .364

Table 3.3.5 Results for 5 levels .The sd_{wb} is too high in band 4,5 and the sd<N> for each band is becoming too high since each band now has less independent points.Thus the more bands the less independant points per band.It can also be seen that <N> does not increase as regularly when going to a higher band. This is due to the increasing sd<N>due to fewer independant points per band than for 4 levels.(Estimates of <N> are no longer as accurate as before).

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.648	290	154	52%	80	11.4%	.197
2	.531	161	423	56%	238	34%	.214
3	.773	89	547	49%	268	38%	.431
4	1.417	83	1199	40%	480	68%	.914
5	1.279	94	1216	41%	503	71%	.757
6	3.000	101	1587	33%	529	75%	2.24
7	3.428	102	2249	32%	731	104%	2.703
8	1.513	247	3399	39%	1337	189%	.574

Table 3.3.6 Here too many bands have resulted in high sd_{wb} and high sd<N> making results inconclusive.

Note even though the final answers are the same the results for 8 levels are obviously incorrect since <N> goes down in level 2,5,8 but even here the results are quite good(Also the standard deviations are too large for conclusive results)

3.4 Effect of varying averaging block size

The larger the averaging block size the less density variations it will detect as it passes over the shoal while the smaller the averaging block size the greater the sd<I> will be.Thus a compromise has to be reached.To see the effect of varying the number of points in a block the data from experiment 4 was chosen since it was believed that increasing the averaging block size would decrease sd<I> and thus reduce misallocation.

Band	<N>	IP	<I>	%sd I	sd<I>	sd_{wb}	sd<N>
1	4.305	64	1135	40%	458	29%	4.79
2	5.235	209	2315	39%	907	58%	3.56
3	3.050	159	3877	42%	1063	107%	1.83
4	3.495	90	5738	41%	2399	154%	2.95

$<N_{av}>=4.153+/-1.71$

Table 3.4.1

Note errors in <N> for levels 3 and 4 due to large deviations in sd_{wb}

There are too few independant points per layer.This causes large sd<N> .

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	3.492	42	1143	38%	432	28%	4.316
2	5.035	132	2328	35%	830	53%	4.219
3	2.176	143	3785	42%	1581	102%	1.214
4	4.738	38	4865	36%	1749	113%	7.173

<N_{av}>=3.672+/-1.887

Table 3.4.2 351 points in a block.The sd_{wb} is slightly lower but not enough to prevent misallocation as a result the<N> values are not in increasing order.The effect of increasing block size on <N> is difficult to monitor because the low number of statistically independant points also has a deleterious affect on <N> and because of the resulting uncertainty in the estimate of<N>it is difficult to notice the effect of changing the block size on <N>

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	2.919	22	1136	27%	314	20%	4.605
2	2.445	95	2759	29%	793	51%	1.741
3	2.012	137	3580	30%	1078	69%	1.115
4	6.511	19	5239	24%	1277	82%	16.55

Table 3.4.3 711 points There is a large sd_{wb} for levels 2,3,4 and a large sd<N> for band 1 and 4 making these results inconclusive but the sd_{wb} can be seen to have improved over the results for 351 points. The effect of reduction in misallocation can be seen to have an improvement in the estimates of<N>in that there are not such large irregular decreases in <N> when going to a higher band than for 351 points.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	3.77	273	2040	53%	1092	35%	1.90
2	2.81	250	4550	56%	2576	83%	1.30

Table 3.4.4 2 bands-351 points. In order to reduce the sd<N> due to too few independant points in a band fewer bands were chosen so that the effect of averaging block size could be more clearly seen.

Band	<N>	IP	<I>	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.995	117	2452	30%	739	47%	1.192
2	2.237	156	3785	29%	1110	72%	1.203

Table 3.4.5 2 bands-711 points. Increasing the averaging block size has resulted in more sensible estimates of <N> for each band.

Note this is not a very good example because the high sd<N> masks the improvement in <N> achieved by increasing averaging block size but it can still be seen that increasing the number of independant points in the block size has a beneficial effect on <N> for each level.

3.5 Comparison with echo integration

To compare with echo integration work out $<N_{av}>$+/-sd$<N_{av}>$ for each experiment then divide by cell size to get ρ_{stat}+/-sdρ_{stat} then divide by ρ_{int} to get the ratio of the two

$$\rho_{stat}=(<a^4>/<a^2>^2)(<b^4>/<b^2>^2)/(\pi r_g^2 c\tau\ (<I^2>/<I>^2-2))$$
$$=(<a^4>/<a^2>^2)(<b^4>/<b^2>^2)<N_{av}>/(2\pi r_g^2 c\tau)$$

$<a^4>/<a^2>^2=(\gamma^2+4\gamma+2)/(\gamma+1)^2$ for $\gamma=0$ $<a^4>/<a^2>^2=2$ $<b^4>/<b^2>^2=245$
$=245<N_{av}>/(\pi r_g^2 c\tau)=<N>/(19.2r_g^2\tau)$

for expl τ =.14*10 r_g^2=22.5*24.4=549 therefore ρ =$<N_{av}>$/1.475

Exp	<N>	equiv cell volume	ρ_{stat}	ρ_{int}	ρ_{stat}/ρ_{int}
1	1.25+/-.35	1.475	.847+/-.236	1.56	.543+/-.151
2A	.65+/-.30	6.026	.107+/-.048	.39	.274+/-.123
2B	1.67+/-.52	16.524	.101+/-.031	.39	.259+/-.079
3A	1.03+/-.17	3.672	.281+/-.047	.78	.360+/-.060
3B	2.42+/-1.25	4.644	.522+/-.269	.78	.669+/-.344
4	4.15+/-1.71	5.56	.747+/-.307	1.42	.526+/-.216
5	1.71+/-.46	5.39	.317+/-.085	1.78	.178+/-.047
6A	4.31+/-1.19	9.8	.440+/-.121	na	
6B	7.05+/-2.71	12.39	.569+/-.218	na	

The reason these ratio's of ρ_{stat} / ρ_{int} are not unity is mainly due to misallocation due to too few independant sample points in the averaging block. The error due to misallocation is not represented by the error due to standard deviation(shown here) and is an error over and above that of standard deviation leading to poor estimates of number density. To overcome this smaller pulse lengths and greater depth ranges and higher ping rates have been chosen for the next set of data collected. This will give more independant samples in a block thus reducing misallocation as well as providing more independant samples in a density region. The higher ping rate will also give more detailed and hence accurate data on the shoal density variations along the track that the ship is travelling and should therefore lead to better results.

4 AFRICANA CRUISE DATA(June 1987)

This data is part of a second set of data .The experiments were chosen to have as short a pulse length as possible and as many pings as possible to increase the total number of independant points. The depth range was increased to 5 m to allow more independant points with depth as well.

Hereafter MI refers to midband intensity-the intensity in the middle of the allocated band.This is compared to the average intensity for the band <I> which theoretically should be equal.

4.1 Analysis for TDE2

This experiment was run on a dispersed shoal of Redeye at night at depth 13-18m.

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.28	10057	2461	2870	39%	954	33%	.0733
2	3.13	6467	6532	5741	31%	2031	71%	.298
3	11.28	2536	11019	8611	26%	2898	101%	3.59
4	11.52	1031	19191	11481	26%	5047	176%	5.84

$<N_{av}>$=3.66+/-.55

Table 4.1.1 Ping 1-1112

From the table it can be seen that for band 1 twice the sd<I> is less than the width of the band but in band 2,3,4 it is greater than the width of the band. This means that some of the points will definately be wrongly allocated to other bands. It is true however that band 4 includes all the points above L including those above L therefore <I> would be larger than it should be this would make sd<I> larger than it should be.

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.27	1494	1761	1651	39%	686	21%	.188
2	2.10	1528	4737	4954	34%	1610	49%	.353
3	7.48	803	8048	8257	27%	2173	66%	3.21
4	5.12	1067	15190	11560	29%	4405	133%	1.52

$<N_{av}>$=3.39+/-.641

Table 4.1.2 Ping 1-277

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.93	1237	857	776	43%	368	24%	.143
2	1.26	1626	2143	2329	39%	835	53%	.178
3	1.50	1066	3810	3882	37%	1409	90%	.274
4	1.93	983	6733	5435	34%	2289	147%	.394

$<N_{av}>$=1.36+/-.12

Table 4.1.3 Ping 278-555

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.74	983	2643	1949	35%	925	23%	.344
2	3.00	1783	5484	5848	31%	1700	43%	.535
3	8.13	1288	9762	9747	27%	2635	67%	2.906
4	15.64	858	15794	13647	25%	3948	101%	10.92

$<N_{av}>$=6.3+/-2.06

Table 4.1.4 Ping 556-833

Band	N	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	2.34	1000	1884	1359	33%	621	22%	.506
2	2.39	1938	3820	4078	32%	1222	45%	.374
3	4.09	1158	6785	6798	30%	2008	74%	1.041
4	5.06	853	11083	9518	29%	3169	116%	1.676

$<N_{av}>$=3.24+/-.41

Table 4.1.5 Ping 834-1112

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.05	1929	1197	1087	41%	490	22%	.132
2	1.37	2176	2977	3262	38%	1131	52%	.171
3	2.24	1413	5478	5438	33%	1807	83%	.401
4	3.36	1326	9675	7614	30%	2902	133%	.729

$<N_{av}>$=1.84+/-.176

Table 4.1.6 13-15m for ping 1-1112

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.73	1872	2152	1778	35%	753	21%	.248
2	3.32	2357	5132	5335	30%	1539	43%	.537
3	11.19	1321	8859	8893	26%	2303	64%	4.91
4	6.96	1293	16106	12450	27%	4348	122%	2.22

$<N_{av}>$=5.09+/-1.055

Table 4.1.7 16-18m for ping 1-1112

	ρ_{stat}	ρ_{int}	ρ_{stat}/ρ_{int}
Ping1-1112	3.067+/-.463	2	1.533+/-.231
Ping 1-277	2.84+/-.536	2.28	1.245+/-.235
Ping 278-555	1.139+/-.10	1.07	1.064+/-.09
Ping 556-833	5.279+/-1.73	2.691	1.961+/-.64
Ping 834-1112	2.715+/-.343	1.87	1.451+/-.183
13-15m	1.85+/-.177	1.48	1.24+/-.119
16-18m	3.46+/-.718	2.49	1.38+/-.288

Table 4.1.8 comparison of ρ_{stat} with ρ_{int}

All the values of ρ_{stat}/ρ_{int} fall within the same range when standard deviation is taken into account. The standard deviation tells which result is closest to the correct value and which results are too far away and must be ignored. The $sd\rho$ given in the table above will be an underestimation of the error in ρ because it does not take account of the error due to misallocation (to the wrong band) of the points.Misallocation is due to too few independant points in the averaging block which will be a problem whenever $sd_{wb}>50\%$
This is the case for all but the lowest band in most of the analysis above

4.2 Analysis of TSE 3

Analysis of dispersed shoal of anchovy at night.

For depth of 10-15m

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.94	8336	766	684	43%	329	24%	.056
2	1.59	11843	1955	2053	36%	704	51%	.088
3	2.57	8001	3351	3423	32%	1072	78%	.203
4	2.64	6744	5731	4793	32%	1834	134%	.230

Table 4.2.1 Ping 1-1927

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.88	1865	597	485	35%	209	22%	.276
2	1.28	3225	1360	1455	39%	530	55%	.129
3	1.77	2050	2440	2425	35%	854	88%	.244
4	1.88	908	3463	3394	35%	1212	124%	.396
5	2.39	359	4652	4364	33%	1535	158%	.868

$<N_{av}>$=1.64+/-.114

Table 4.2.2 Ping 1-481

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.02	1975	761	622	41%	312	25%	.126
2	1.71	3148	1752	1867	35%	613	49%	.188
3	2.43	1928	3062	3112	32%	980	78%	.383
4	3.39	963	4388	4357	31%	1360	109%	.866
5	4.01	377	5809	5602	29%	1684	135%	1.773

$<N_{av}>$=2.01+/-.173

Table 4.2.3 Ping 482-962

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	1.67	1051	1290	940	36%	464	24%	.316
2	2.69	3787	2734	2820	32%	875	46%	.315
3	4.40	2595	4570	4701	29%	1325	70%	.776
4	3.89	856	6726	6582	30%	2018	107%	1.125
5	6.34	248	8750	8463	28%	2450	130%	4.427

$<N_{av}>$=3.31+/-.325

Table 4.2.4 Ping 963-1443

Band	<N>	IP	<I>	MI	%sd<I>	sd<I>	sd_{wb}	sd<N>
1	.57	2194	637	701	51%	325	23%	.063
2	1.19	2671	1932	2103	39%	753	53%	.130
3	2.18	1957	3419	3505	33%	1128	80%	.328
4	2.60	1075	4963	4908	32%	1588	113%	.563
5	2.53	450	6876	6310	32%	2200	157%	.84

$<N_{av}>$=1.51+/-.123

Table 4.2.5 Ping 1444-1924

	ρ_{stat}	ρ_{int}	ρ_{stat}/ρ_{int}
Ping 1-1927	2.43+/-.094	1.358	1.79+/-.07
Ping 1-481	2.14+/-.149	.957	2.236+/-.155
Ping 482-962	2.62+/-.173	1.233	2.12+/-.140
Ping 963-1443	4.32+/-.325	1.85	2.33+/-.175
Ping 1444-1924	2+/-.160	1.386	1.44+/-.115

Most ratio's fall within the same range . Except for errors due to standard deviation there is a constant ratio of approximately 2.1

The ratio's are higher than 1 indicating that the estimate of $\gamma \simeq 0$ might have been wrong and that $\gamma \simeq \infty$ might have been more correct since this will halve the ratio.

In the data analysed all but the first band had sd_{wb} which was too high. This may also have led to error.

Choosing less density levels in the analysis for this chapter may have led to more accurate results however the increased number of independant points did decrease the sd<N>

5. CONCLUSION

There are density variations within fish shoals which have to be taken into account when using this model.When the shoals are divided into regions of nearly constant density and the statistical technique is applied to regions of constant density and the resultant density estimates are combined to find the overall average density for the fish shoal it is found to compare well with the estimate of fish number density obtained by the echo integration method.

In the data analysis of chapter 4 the difference between the estimates of fish number density using the two methods was clearly seen to be due to the standard deviation in the statistical method.To bring down this standard deviation it is necessary to have still more independant points in the data set which effectively means shorter pulse lengths and a higher ping rate. Another way of decreasing the standard deviation is to have lower values of <N> by decreasing the resolution cell size by employing narrower beamwidths and shorter pulse lengths. Unfortunately on the equipment available data for shorter pulselengths could not be obtained. Going to a shorter pulse length also means going to a higher frequancy in order to avoid pulse distortion though higher frequancies run into propagation difficulties. Thus there is a limit to the maximum fish number density that can be measured by this method.

Another cause of errors in the estimate of <N> comes about when the shoal is partitioned into regions of different density. If sd<I> for the averaging block is larger than half the difference between the lower and upper level for a density band then the misallocation of data samples to the wrong density band can be quite severe. Thus a low sd<I> has to be attained. This can be done by increasing the number of independant points or reducing the resolution cell size. The number of independant points in the averaging block can be increased by decreasing the pulselength and ping rate or keeping the pulse length and ping rate the same and increasing the size of the block although the block size needs to be kept as small as possible in order to be able to reflect the density variations within the shoal. Misallocation can also be reduced by using less density bands thus increasing the difference between the upper and lower level of the band.

Since any misallocation of data samples to the wrong band will result in errors in the calculation in both the band it should have fallen into as well as the band it did fall into misallocation can be a serious problem. In the data sampled a compromise had to be reached but in future use the problem should be kept to a minimum for the method to work accurately.

Although the method is independant of target strength it still does depend on the statistical nature of the backscatter more precisely on the probability density function of the backscattered amplitude from a single fish. More work will have to be done on fish scatter statistics particularly to work out for various fish species to see how much does vary if at all.(may also vary with fish movement.)

Although beam divergence was not taken into account for work done in this thesis the work could be extended to do that.

This would result in the density being a function of depth and time as opposed to the assumption in this thesis that density was time invariant.

The method may also be used in conjunction with the echo integration method to determine the target strength of the fish although it will take further work to improve the accuracy of the statistical method to such an extent. When the true beamshape was taken into account it made the estimate of number density a little lower than the estimate obtained assuming a conical beam. A further improvement can be obtained by working out the beamshape for a scattering layer which is what a fish shoal is instead of a scattering volume as was done.

6. REFERENCES

Abramowitz and Stegun
"Pocket Book of Mathematical Functions"
Verlag Harri Deutsch

Bodholt,H.(1969)
"Quantitive Measurement of Scattering Layers"
ICES/FAO Acoustic Training Course, Svolvaer, Norway 2-15 March, 1969

Clay,C.S. and Heist,B.G. (1984)
"Acoustic Scattering by Fish - Acoustic Models and a Two Parameter Fit"
J. Acoust. Soc. Am,75(4),1077-1083

Clay,C.S. and Medwin,H (1977)
"Acoustical Oceanography: Principles and Applications"
John Wiley and Sons,Inc.

Davis,P.J. and Rabinowitz,P. (1967)
Numerical Integration
Blaisdell Publishing Company, Waltham Massachusetts
Denbigh,P.N. and Smith,Q. (1987) see last reference

Denbigh,P.N. and Weintroub,J. (1985)
"Fish Stock Assesment by a Statistical Analysis of Echo Sounder Signals"
12th I.C.A Associated Symposium on Underwater Acoustics

Ehrenburg,J.E. (1973)
"Estimation of the Intensity of a Filtered Poisson Process and its Application to Acoustic Assesment of Marine Organisms"
Phd Thesis, WSG 73-2, Division of Marine Resources,University of Washington.

Ehrenburg,J.E. (1974)
"Recursive Algorithm for Estimating the Spatial Density of Acoustic Scatterers"
J. Acoust. Soc. Am.,vol 56, no.2,Aug.,Pg 542-547

Ehrenburg,J.E.
"Calculation of the Constants Needed to Scale the Output of an Echo Integrator"
Biomass Handbook no.7, NOAA/NMFS

Ehrenburg,J.E. and Lytle,D.W. (1972)
"Acoustic Techniques for Estimating Fish Abundance"
IEEE Trans, GE-10,pg 138-145

Jobst,W.J. and Smits,T.I. (1974)
"Mathematical Model for Volume Reverberation: Experiment and Simulation"
J. Acoust. Soc. Am vol55,no. 2,Feb,pg 227-236

Lozow,J.B.
"Transducer Directivity: A Simple Calculation of its Spatial Averages"
The Charles Stark Draper Laboratory,Inc,Cambridge,Massachusetts,U.S.A.
pg495-524

Middleton,D. (1972)
"A Statistical Theory of Reverberation and Similar First-order Scattered Fields- Part IV: Statistical Models"
IEEE Transactions on Information Theory,Vol IT-18,no.1,Jan, Pg 68-90

Papoulis,A. (1965)
"Probability Random Variables, And Stochastic Processes"
McGraw Hill Kogakusha, Ltd

Parzen,E. (1962)
"Stochastic Processes"
San Francisco,Holden-Day

Peterson,M.L.,Clay,C.S. and Brandt,S.B. (1976)
"Acoustic Estimates of Fish Density and Scattering Function"
J. Acoust. Soc. Am.,vol 60,no.3,Sept

Pusey,P.N.,Schaefer,D.W. and Koppel,D.E. (1974)
"Single Interval Statistics of Light Scattered by Identical Independant Scatterers"
J. Phys. A. Math.,Nucl,Gen.,vol 7,no.4,pg530-540

Rice,S.O. (1944,1945)
"Mathematical Analysis of Random Noise"
Bell Sys Tech Journ,July vol XXIII,no.3,pg 282-332 and Jan, vol XXIV,no.1,pg 46-156

Spiegel,M.O. (1968)
"Mathematical Handbook"
McGraw-Hill,Inc.

Weintroub,J. (1986)
"Fish Stock Assesment by a Statistical Analysis of Echo Sounder Signals"
U.C.T.,Oct.(Masters Thesis)

Wilhelmij,P. (1983)
"A Statistical Approach to Determine the Number Density of Random Scatterers from Backscattered Pulses"
U.C.T.,Apr.(Masters Thesis)

Wilhelmij,P. and Denbigh,P. (1984)
"A Statistical Approach to Determining the Number Density of Random Scatterers from Backscattered Pulses"
J. Acoust. Soc. Am 76(6),Dec., pg 1810-1818

Denbigh, P.N. and Smith, Q. (1987)
"Experiments with an Acoustic Technique for Estimating Fish Number Density, not Requiring Fish Target Strength"
International Symposium on Fisheries Acoustics, June 22 - 26, 1987, Seattle, Washington, U.S.A.

7.0 APPENDICES

APPENDIX 1

The purpose of this appendix is to derive the equations for $<I^2>, <I^3>, <I^4>$ from the formula $<I_N^n> = \sum (n!)^2/((n_1!)^2(n_2!)^2 \ldots (n_N!)^2) \prod_{i=1}^{N} <(a^2)^{n_i}> = \sum M_{N,n}^2 <(a^2)^{n_i}>$ where $n_1, n_2 \ldots n_N \in (0,1,2 \ldots n)$ and $\sum_{i=1}^{N} n_i = n$

The values of $<I_N^2>, <I_N^3>$ and $<I_N^4>$ are derived by choosing N=4 and finding expressions for $<I_4^2>, <I_4^3>$ and $<I_4^4>$ and then to try to discover the answer by induction for any N.

To calculate the values of $n_1, n_2 \ldots n_N$ the following algorithm in BASIC was used. Here the values for n are 2,3,4

```
INPUT "n=";n
FOR I=0 TO n
FOR J=0 TO n
FOR K=0 TO n
FOR L=0 TO n
IF I+J+K+L=n PRINT "n1=";I,"n2=";J,"n3=";K,"n4=";L
NEXT L
NEXT K
NEXT J
NEXT I
```

A1.1 Equation for $<I_N^2>$

Using n=2 in the preceding program and the expression $M_{N,n} = n!/(n_1!n_2! \ldots n_N!) = 2!/(n_1!n_2! \ldots n_N!)$ one gets

n_1	n_2	n_3	n_4	$<a^{2n_1}>$	$<a^{2n_2}>$	$<a^{2n_3}>$	$<a^{2n_4}>$	$M_{4,2}^2$	$\prod_{i=1}^{4}<a^{2n_i}>$		
0	0	0	2	1	1	1	$<a^4>$	1	$<a^4>$		
0	0	1	1	1	1	$<a^2>$	$<a^2>$	2^2	$<a^2>^2$	$4<a^2>^2(1)$	
0	0	2	0	1	1	$<a^4>$	1	1	$<a^4>$		
0	1	0	1	1	$<a^2>$	1	$<a^2>$	2^2	$<a^2>^2$	$4<a^2>^2(2)$	
0	1	1	0	1	$<a^2>$	$<a^2>$	1	2^2	$<a^2>^2$		$4<a^2>^2(1+2+3)$
0	2	0	0	1	$<a^4>$	1	1	1	$<a^4>$		
1	0	0	1	$<a^2>$	1	1	$<a^2>$	2^2	$<a^2>^2$		
1	0	1	0	$<a^2>$	1	$<a^2>$	1	2^2	$<a^2>^2$	$4<a^2>^2(3)$	
1	1	0	0	$<a^2>$	$<a^2>$	1	1	2^2	$<a^2>^2$		
2	0	0	0	$<a^4>$	1	1	1	1	$<a^4>$		

Thus $\langle I_4^2\rangle=4\langle a^4\rangle+2^2\langle a^2\rangle^2(1+2+3)$

For any N $\langle I_N^2\rangle=N\langle a^4\rangle \ +2^2\langle a^2\rangle^2(1+2+3.....+(N-1)$

$=N\langle a^4\rangle+2^2\langle a^2\rangle^2N(N-1)$

A1.2 Equation for $\langle I_N^3\rangle$

Here $M_{4,3}=3!/(n_1!n_2!n_3!n_4!)$

n_1	n_2	n_3	n_4	$(3!)^2\prod_{i=1}^{4}\langle a^{2n_i}\rangle/(n_i!)^2$	
0	0	0	3	$\langle a^6\rangle$	
0	0	1	2	$3^2\langle a^2\rangle\langle a^4\rangle$	$3^2\langle a^2\rangle\langle a^4\rangle(2)$
0	0	2	1	$3^2\langle a^2\rangle\langle a^4\rangle$	
0	0	3	0	$\langle a^6\rangle$	
0	1	0	2	$3^2\langle a^2\rangle\langle a^4\rangle$	
0	1	1	1	$6^2\langle a^2\rangle^3$	
0	1	2	0	$3^2\langle a^2\rangle\langle a^4\rangle$	$3^2\langle a^2\rangle\langle a^4\rangle(4)$
0	2	0	1	$3^2\langle a^2\rangle\langle a^4\rangle$	
0	2	1	0	$3^2\langle a^2\rangle\langle a^4\rangle$	
0	3	0	0	$\langle a^6\rangle$	
1	0	0	2	$3^2\langle a^2\rangle\langle a^4\rangle$	
1	0	1	1	$6^2\langle a^2\rangle^3$	$6^2\langle a^2\rangle^3(4)$
1	0	2	0	$3^2\langle a^2\rangle\langle a^4\rangle$	
1	1	0	1	$6^2\langle a^2\rangle^3$	
1	1	1	0	$6^2\langle a^2\rangle^3$	
1	2	0	0	$3^2\langle a^2\rangle\langle a^4\rangle$	
2	0	0	1	$3^2\langle a^2\rangle\langle a^4\rangle$	$3^2\langle a^2\rangle\langle a^4\rangle(6)$
2	0	1	0	$3^2\langle a^2\rangle\langle a^4\rangle$	
2	1	0	0	$3^2\langle a^2\rangle\langle a^4\rangle$	
3	0	0	0	$\langle a^6\rangle$	

$\langle I_4^3\rangle=4\langle a^6\rangle+3^2\langle a^2\rangle\langle a^4\rangle(6+4+2)+6^2\langle a^2\rangle^3(4)$

$=4\langle a^6\rangle+3^2\langle a^2\rangle\langle a^4\rangle(4(4-1))+6\langle a^2\rangle^3(4(4-1)(4-2))$

In general $\langle I_N^3\rangle=N\langle a^6\rangle+3^2\langle a^2\rangle\langle a^4\rangle N(N-1)+6\langle a^2\rangle^3N(N-1)(N-2)$

A1.3 Equation for $<I_N^4>$

n_1	n_2	n_3	n_4	$(4!)^2\prod_{i=1}^{4}<(a^{2n_i})>/(n_i!)^2$
0	0	0	4	$<a^8>$
0	0	1	3	$4^2<a^2><a^6>$
0	0	2	2	$6^2<a^4>^2$
0	0	3	1	$4^2<a^2><a^6>$
0	0	4	0	$<a^8>$
0	1	0	3	$4^2<a^2><a^6>$
0	1	1	2	$12^2<a^2>^2<a^4>$
0	1	2	1	$12^2<a^2>^2<a^4>$
0	1	3	0	$4^2<a^2><a^6>$
0	2	0	2	$6^2<a^4>^2$
0	2	1	1	$12^2<a^2>^2<a^4>$
0	2	2	0	$6^2<a^4>^2$
0	3	0	1	$4^2<a^2><a^6>$
0	3	1	0	$4^2<a^2><a^6>$
0	4	0	0	$<a^8>$
1	0	0	3	$4^2<a^2><a^6>$
1	0	1	2	$12^2<a^2>^2<a^4>$
1	0	2	1	$12^2<a^2>^2<a^4>$
1	0	3	0	$4^2<a^2><a^6>$
1	1	0	2	$12^2<a^2>^2<a^4>$
1	1	1	1	$24^2<a^2>^4$
1	1	2	0	$12^2<a^2>^2<a^4>$
1	2	0	1	$12^2<a^2>^2<a^4>$
1	2	1	0	$12^2<a^2>^2<a^4>$
1	3	0	0	$4^2<a^2><a^6>$
2	0	0	2	$6^2<a^4>^2$
2	0	1	1	$12^2<a^2>^2<a^4>$
2	0	2	0	$6^2<a^4>^2$
2	1	0	1	$12^2<a^2>^2<a^4>$
2	1	1	0	$12^2<a^2>^2<a^4>$
2	2	0	0	$6^2<a^4>^2$
3	0	0	1	$4^2<a^2><a^6>$
3	0	1	0	$4^2<a^2><a^6>$
3	1	0	0	$4^2<a^2><a^6>$
4	0	0	0	$<a^8>$

terms in $<a^8>$ sum to $<a^8>(1+1+1+1)=4<a^8>$

terms in $<a^2><a^6>$ sum to $4^2<a^2><a^6>(6+4+2)=4^2<a^2><a^6>4(4-1)$

terms in $<a^4>^2$ sum to $6^2<a^4>^2(1+2+3)=18<a^4>^2 4(4-1)$

terms in $<a^2>^2<a^4>$ sum to $12^2<a^2>^2<a^4>(3+9)=72<a^2>^2<a^4>4(4-1)(4-2)$

There is only one term in $<a^2>^4$ and that is $24^2<a^2>^4=24<a^2>^4 4(4-1)(4-2)(4-3)$

Thus $<I^4>=4<a^8>+(16<a^2><a^6>+18<a^4>^2)N(N-1)+72<a^2>^2<a^4>N(N-1)(N-2)$
$+24<a^2>^4N(N-1)(N-2)(N-3)$

and $<I^4>=N<a^8>+(16<a^2><a^6>+18<a^4>^2)N(N-1)+72<a^2>^2<a^4>N(N-1)(N-2)$
$+24<a^2>^4N(N-1)(N-2)(N-3)$

APPENDIX 2

The purpose of this appendix is to derive the moments of the Ricean pdf. Rice(1944,1945) gives the pdf for the envelope R of a sine wave Pcospt accompanied by noise with variance Ψ_0 as $P(R)=(R/\Psi_0)\exp(-R^2+P^2)/(2\Psi_0)I_0(RP/\Psi_0)$ and the n'th moment as $\overline{R}^n=(2\Psi_0)^{n/2}\Gamma(n/2+1)\,{}_1F_1(-n/2;1;-P^2/(2\Psi_0))$ where ${}_1F_1$ is the confluent hypergeometric function (pg 189 Abramowitz and Stegun) and Γ is the gamma function.

The pdf of the echo amplitude $\hat{a}$ due to a fish is
$w_R(a)=(2a/\sigma_d)\exp(-(a^2+\sigma_c)/\sigma_d)I_0(2a\sqrt{\sigma_c}/\sigma_d)$ and $\gamma=\sigma_c/\sigma_d$
A comparison of these two notations yields the following identities.
$\sigma_d/2=\Psi_0$; $a=R$; $\gamma=\sigma_c/\sigma_d=P^2/(2\Psi_0)$
Therefore $\langle a^n\rangle=(\sigma_d)^{n/2}\Gamma(n/2+1)M(-n/2;1;-\gamma)$ where M is the alternative notation for a confluent hypergeometric function.
To work out $\langle a^2\rangle,\langle a^4\rangle,\langle a^6\rangle,\langle a^8\rangle$ only even values of n are needed therefore put $n=2m$ where $m=1,2,3....$
Thus $n/2=m$ and $\langle a^m\rangle=(\sigma_d)^m\Gamma(m+1)M(-m;1;-\gamma)$
$=(\sigma_d)^m m!M(-m;1;-\gamma)$ since $\Gamma(m+1)=m!$ (pg 101 Spiegel)
$M(-n;\alpha+1;x)=n!/(\alpha+1)_n L_n^{\alpha}(x)$ (pg 194, Abromowitz and Stegun)
where $L_n^{\alpha}(x)$ is the Generalised Laguerre Polynomial.
for $\alpha=0$ $M(-n;1;x)=n!/(1)_n L_n^0(x)$
$(a)_n$ is the Pach hamers symbol $(a)_n=\Gamma(a+n)/\Gamma(a)$ and $(1)_n=\Gamma(1+n)/\Gamma(1)=n!/0!=n!$
Therefore $M(-n;1;x)=n!/n!L_n^0(x)=L_n(x)$ where $L_n(x)$ is the Laguerre polynomial.
$L_n(x)=\sum_{m=0}^{n}(-1)^m\binom{n}{n-m}x^m/m!$ (pg 334,Abromowitz and Stegun)
Thus $L_1(x)=1-x$
$L_2(x)=(2-4x+x^2)/2!$
$L_3(x)=(6-18x+9x^2-x^3)/3!$
$L_4(x)=(24-96x+72x^2-16x^3+x^4)/4!$
Putting $x=-\gamma$, $\langle a^{2m}\rangle=(\sigma_d)^m m!L_m(-\gamma)$
Therefore $\langle a^2\rangle=\sigma_d\,(1+\gamma)$
$\langle a^4\rangle=\sigma_d^2\,(2+4\gamma+\gamma^2)$
$\langle a^6\rangle=\sigma_d^4\,(6+18\gamma+9\gamma^2+\gamma^3)$
$\langle a^8\rangle=\sigma_d^8\,(24+96\gamma+72\gamma^2+16\gamma^3+\gamma^4)$

APPENDIX 3

The purpose of this appendix is to calculate the moments $\langle N\rangle,\langle N^2\rangle,\langle N^3\rangle,\langle N^4\rangle$ for various distribution pdf's.

A3.1 Poisson pdf moments

The moments of the Poisson pdf are best calculated from it's characteristic function $f(w)=\exp(a(\exp(jw)-1))$ using the moment thereom which states that $(d^n f/dw^n)_{jw=0}=j^n E(x^n)$

Letting $t=jw$ gives $f(t)=\exp(a(e^t-1))$. Now $df/dw=df/dt.dt/dw=jdf/dt$

$d^2f/dw^2=jd/dw(df/dt)=j(dt/dw)(d^2f/dt^2)=j^2d^2f/dt^2$. In general $d^nf/dw^n=j^nd^nf/dt^n$

Therefore $E(x^n)=(d^nf/dt^n)_{t=0}$

$<N>=E(x)=(df/dt)_{t=0}=(\exp(a(e^t-1))ae^t)_{t=0}=(fae^t)_{t=0}=af(0)=a$

$<N^2>=E(x^2)=(d^2f/dt^2)_{t=0}=((df/dt)ae^t+fae^t)_{t=0}=(ae^t(df/dt+f))_{t=0}=a(a+1)=a^2+a$

$=<N>^2+<N>$

Similarly $<N^3>=E(x^3)=(d^3f/dt^3)_{t=0}=a^3+3a^2+a=<N>^3+3<N>^2+<N>$

and $<N^4>=<N>^4+6<N>^3+7<N>^2+<N>$

A3.2 Binomial pdf moments

For a binomial pdf $f(w)=(p.\exp(jw)+q)^n$ putting $t=jw$ $f(t)=(pe^t+q)^n$

Let $z=pe^t+q$ therefore $f(z)=z^n$ and $(d^nz/dt^n)_{t=0}=(pe^t)_{t=0}=p$ for all n.

Also $z(0)=p+q=1$

$<N>=(df/dt)_{t=0}=(nz^{n-1}dz/dt)_{t=0}=np$

$<N^2>=(d^2f/dt^2)_{t=0}=(n(n-1)z^{n-2}(dz/dt)^2+nz^{n-1}d^2x/dt^2)_{t=0}=n^2p^2-(np)p+np$

$=<N>^2-<N>(<N>/n)+<N>$

A3.3 Uniform pdf moments

For a uniform pdf $P(N=i)=1/(n+1)$ where $i=0,1,2.....n$ $\quad \Sigma\equiv\sum_{i=0}^{n}$

$<N^k>=\Sigma i^kP(N=i)=P(N=i)\Sigma i^k=(1/(n+1))\Sigma i^k$

$<N>\ =(1/(n+1))\Sigma i=(1/(n+1))n(n+1)/2=n/2$

$<N^2>=(1/(n+1))\Sigma i\ =(1/(n+1))n(n+1)(2n+1)/6=n(2n+1)/6=<N>(4<N>+1)/3=4<N>^2/3+<N>/3$

APPENDIX 4

The purpose of this appendix is to calculate the values of $<b^4>/<b^2>$ for different transducers using numeric integration.

From Ehrenburg(1973,APPB) $E(b^n(\theta,\phi))=\int_0^{2\pi}\int_0^{\pi/2} b^n(\theta,\phi)f(\theta,\phi)d\theta d\phi$

When scattering density is constant throughout the volume this expression reduces to $E(b^n(\theta,\phi))=<b^n(\theta,\phi)>=(1/(2\pi))\int_0^{\pi/2}\sin\theta\int_0^{2\pi}b^n(\theta,\phi)d\phi d\theta$

A4.1 Circular piston transducer

For a circular piston transducer $b(\theta,\phi)=D^2(\theta,\phi)=(2J_1(\pi(d/\lambda)\sin\theta)/(\pi(d/\lambda)\sin\theta))^2$

(Clay and Medwin pg 144)

Thus the $<b^n(\theta,\phi)>=1/(2\pi)\int_0^{\pi/2}\sin\theta\bullet b^n(\theta)\bullet\int_0^{2\pi}d\phi\bullet d\theta=\int_0^{\pi/2}b\ (\theta)\sin\theta d\theta$

This integral had to be evaluated numerically using the following program in BASIC.

```
1: INPUT"XO=";D,"X2P=";E,"P=";F:A=0:X=0:Y=0:GOSUB5:D=(E-D)/2/F
2: GOSUB 5:Y=4Y:GOSUB 5:A=Y+A:F=F-1:IF F  0 GOTO 2
3: BEEP 3:PRINT "ANS.",AD/3:END
5: A=Y+A:X=X+D:Z=(d/λ)*π*SIN X:IF Z=0 LET Y=0:RETURN
6: Z=Z/2:J=Z:K=0:Y=0
7: K=K+1:Y=Y+J:J=-JZZ/K/(K+1):IF ABS(J) 1E-7 GOTO 7
8: Z=2Z:y=Z/(d/λ)/π*(4*Y*Y/Z/Z)^n:RETURN
```

This program is basically a simpsons rule integration program(P.J.Davis 1967). XO is the start limit of the integration and X2P is the end limit of the integration. P is the number of steps in the integration.

Subroutine 5 includes a function definition. Line 7 calculates J_1 the Bessel function of the first kind of order 1 and line 8 calculates the function $b^n(\theta)\sin\theta$. The effective diameter of the transducer can be calculated from the experimental value of beamwidth. The 3dB beamwidth=8.8 degrees.

Thus the angle at the 3dB point is 8.8/2=4.4 degrees.

At the 3dB point of the transducer $k.a.\sin\theta_{3dB}=1.6=(d/\lambda)\sin\theta_{3dB}$.(pg 146,Clay and Medwin).Thus $d/\lambda=6.64$

This result can be checked by calculation using the angle to the first null which was 11.5 from experimental data. $d/\lambda=3.83/(\pi\sin 11.5)=6.115$ (pg 146,Clay and Medwin)

Choosing P=40 and $d/\lambda=6.6$ and n=2 and 4 yields

$\langle b^4\rangle/\langle b^2\rangle^2=1.11041\text{x}10^{-3}/(2.12824\text{x}10^{-3})^2=245$

A4.2 Rectangular Transducer

An approximate expression for $\langle b^n\rangle$ for a rectangular transducer is given by Lozow .

$\langle b^n\rangle=(3/(2n))(\lambda/(\pi\sqrt{(ab)}))^2(1-\exp(-(4n/3)(\pi\sqrt{(ab)}/\lambda)^2\sin(\phi_m/2)))/(1-\cos\phi_m)$

with $\phi=\pi/2$ and $a/\lambda=b/\lambda=6$, $(\lambda/(\pi\sqrt{(ab)})=1/(\pi^2 36)$

and $\langle b^4\rangle/\langle b^2\rangle^2=(2/3)(\pi^2 36).1,0002=237$

The calculations for this appendix are based on the theoretical beampattern . It would be even better to measure the actual beam patterns of the transducers used and do a numerical integration on the measured beam pattern since this may differ from the theoretical ones.

The results above can still be refined further because they have been calculated with the assumption that the fish occur within the whole volume of the water below the transducer. In fact fish occur and the data is collected from a scattering layer. The results above can be modified for a scattering layer(Ehrenburg 1973,pg 56). For a scattering layer one has again a non-homogenous Filtered Poisson Process since the density will vary with depth.Above the scattering layer it will be zero and in the scattering layer it will be constant.(Non-homogenous filtered Poisson Processes are beyond the scope of this thesis so the scattering layer modification will not be calculated.

While in the calculation above $f(\theta,\phi)=\sin\theta/(2\pi)$ with a scattering layer at d_0

$$f(\theta,\phi,t)=\left\{\begin{array}{ll} 0 & \text{if } \theta< \cos{}^{1}(2d_0/(ct)) \\ \\ \sin\theta/(2\pi(1-(2d_0/(ct))) & \text{if } \theta\geqq \cos{}^{1}(2d_0/(ct)) \end{array}\right\}$$

Note that this calculatipon involves beam divergence and hence is non-stationary and involves a non-homogenous filtered Poisson Processes which is beyond the scope of this thesis so the scattering layer modification will not be calculated.

APPENDIX 5

The purpose of this appendix is to determine the errors in the estimates of normalised moments caused by using a finite number of measurements in evaluating the averages of I^n and of I more accurately than in Wilhelmij and Denbigh(1984)

A5.1 Determining the standard deviation in second normalised moment of intensity

Consider the variance in the n'th normalised intensity moment.

Write $<I^n>/<I>^n=xy^{-n}$ where $x=<I^n>$ $y=<I>$

If $z=xy^{-n}$ then $\sigma_z^2=(\partial z/\partial x)^2\sigma_x^2 +(\partial z/\partial y)^2\sigma_y^2 +2(\partial z/\partial x)(\partial z/\partial y)\mu_{11}$ (Papoulis pg212)

σ_z^2 is the variance in z

σ_x^2 is the variance in x

σ_y^2 is the variance in y

$$\sigma_z^2=(y^{-n})^2\sigma_x^2 +(-nxy^{-n-1})^2\sigma_y^2+(-2nxy^{-2n-1})\mu_{11}$$

$$=\sigma_x^2/(y^n)^2+n^2(x/y^{n+1})^2\sigma_y^2-2nx/y^{2n+1}\mu_{11}$$

$$=\sigma_x^2/\langle I\rangle^{2n}+n^2(\langle I^n\rangle/\langle I\rangle^{n+1})^2\sigma_y^2-2n(\langle I^n\rangle/\langle I\rangle^{2n+1})\mu_{11}$$

σ_x^2=variance in $\langle I^n\rangle$=<variance in I^n>=(variance in I^n)/M

$=(\langle I^{2n}\rangle-\langle I^n\rangle^2)/M$ since $\mathrm{var}(x)=E(x^2)-E^2(x)$ (Papoulis pg 144)

and $\sigma_y^2=(\langle I^2\rangle-\langle I\rangle^2)/M$

$$\mu_{11}=\mathrm{cov}(\langle I^n\rangle,\langle I\rangle)=E(\langle I^n\rangle\langle I\rangle)-E(\langle I^n\rangle)E(\langle I\rangle)$$

$$=(ME(I^{n+1})+M(M-1)E(I^n)E(I))-\langle I^n\rangle\langle I\rangle$$

$$=(M\langle I^{n+1}\rangle+M^2\langle I^n\rangle\langle I\rangle-M\langle I^n\rangle\langle I\rangle-M^2\langle I^n\rangle\langle I\rangle)/M^2$$

$$=(M\langle I^{n+1}\rangle-M\langle I^n\rangle\langle I\rangle)/M^2$$

$$=(\langle I^{n+1}\rangle-\langle I^n\rangle\langle I\rangle)/M$$

Therefore

$$\sigma_z^2=(\langle I^{2n}\rangle-\langle I^n\rangle^2)/(\langle I\rangle^{2n} M)+n^2(\langle I^n\rangle/\langle I\rangle^{n+1})^2(\langle I^2\rangle-\langle I\rangle^2)/M$$

$$-2n(\langle I^n\rangle/\langle I\rangle^{2n+1})(\langle I^{n+1}\rangle-\langle I^n\rangle\langle I\rangle)/M$$

$$=(\langle I^{2n}\rangle/\langle I\rangle^{2n}-(\langle I^n\rangle/\langle I\rangle^n)^2)/M+n^2(\langle I^n\rangle/\langle I\rangle^n)^2(\langle I^2\rangle/\langle I\rangle^2-1)/M$$

$$-2n(\langle I^n\rangle/\langle I\rangle^n)(\langle I^{n+1}\rangle/\langle I\rangle^{n+1} -\langle I^n\rangle/\langle I\rangle^n)/M$$

$$=(\langle I^{2n}\rangle/\langle I\rangle^{2n}-(\langle I^n\rangle/\langle I\rangle^n)^2+n^2(\langle I^n\rangle/\langle I\rangle^n)^2(\langle I^2\rangle/\langle I\rangle^2)-n^2(\langle I^n\rangle/\langle I\rangle^n)^2$$

$$-2n(\langle I^n\rangle/\langle I\rangle^n)(\langle I^{n+1}\rangle/\langle I\rangle^{n+1})+2n(\langle I^n\rangle/\langle I\rangle^n)^2)/M$$

$$=(\langle I^{2n}\rangle/\langle I\rangle^{2n}-2n(\langle I^{n+1}\rangle/\langle I\rangle^{n+1})(\langle I^{n}\rangle/\langle I\rangle^{n})+(\langle I^{n}\rangle/\langle I\rangle^{n})(n^2(\langle I^2\rangle/\langle I\rangle^2)-n^2 +2n-1))/M$$

$$=(\langle I^{2n}\rangle/\langle I\rangle^{2n}-2n(\langle I^{n+1}\rangle/\langle I\rangle^{n+1})(\langle I^{n}\rangle/\langle I\rangle^{n})+(\langle I^{n}\rangle/\langle I\rangle^{n})^2(n^2(\langle I^2\rangle/\langle I\rangle^2) -(n-1)^2))/M$$

For second normalised moments $n=2$ and variance in $\langle I^2\rangle/\langle I\rangle^2$ equals

$$(\langle I^4\rangle/\langle I\rangle^4-4(\langle I^3\rangle/\langle I\rangle^3)(\langle I^2\rangle/\langle I\rangle^2)+(\langle I^2\rangle/\langle I\rangle^2)^2(4(\langle I^2\rangle/\langle I\rangle^2)-1))/M$$

$$=(\langle I^4\rangle/\langle I\rangle^4-4(\langle I^3\rangle/\langle I\rangle^3)(\langle I^2\rangle/\langle I\rangle^2)+4(\langle I^2\rangle/\langle I\rangle^2)^3-(\langle I^2\rangle/\langle I\rangle^2)^2)/M$$

A5.2 Derivation for Variance When Dealing with Constant Amplitude Scatterers and a Poisson Distribution

For constant amplitude scatterers and a Poisson distribution

$\langle I^2\rangle/\langle I\rangle^2=2+1/\langle N\rangle$

$\langle I^3\rangle/\langle I\rangle^3=6+9/\langle N\rangle+1/\langle N\rangle^2$

$\langle I^4\rangle/\langle I\rangle^4=24+72/\langle N\rangle+34/\langle N\rangle^2+1/\langle N\rangle^3$

$$\sigma^2=(\langle I^4\rangle/\langle I\rangle^4-4(\langle I^3\rangle/\langle I\rangle^3)(\langle I^2\rangle/\langle I\rangle^2)+4(\langle I^2\rangle/\langle I\rangle^2)^3-(\langle I^2\rangle/\langle I\rangle^2)^2)/M$$

$$=((24+72/\langle N\rangle+34/\langle N\rangle^2+1/\langle N\rangle^3)-4(6+9/\langle N\rangle+1/\langle N\rangle^2)(2+1/\langle N\rangle)+4(2+1/\langle N\rangle)^3 -(2+1/\langle N\rangle)^2)/M$$

$$=(4+20/\langle N\rangle+13/\langle N\rangle^2+1/\langle N\rangle^3)/M$$

Therefore standard deviation in the second normalised moment of intensity

$=\mathrm{sd}\langle I^2\rangle/\langle I\rangle^2= ((4+20/\langle N\rangle+13/\langle N\rangle^2+ 1/\langle N\rangle^3)/M)$

as $\langle N\rangle\to\infty$ $\mathrm{sd}\langle I^2\rangle/\langle I\rangle^2\to 2/\surd M$ which was derived by Wilhelmij and Denbigh(1984)

A5.3 Derivation for Variance When Dealing With Rayleigh Amplitude Statistics and Poisson Distribution

$\langle I^2\rangle/\langle I\rangle^2=2!(1/\langle N\rangle+1)$

$\langle I^3\rangle/\langle I\rangle^3=3!(1/\langle N\rangle^2+3/\langle N\rangle+1)$

$\langle I^4\rangle/\langle I\rangle^4=4!(1/\langle N\rangle^3+7/\langle N\rangle^2+6/\langle N\rangle+1)$

$$\sigma^2=(4!(1/\langle N\rangle^3+7/\langle N\rangle^2+6/\langle N\rangle+1)-4.3!(1/\langle N\rangle^2+3/\langle N\rangle+1)2!(1/\langle N\rangle+1) +4.(2!)^3(1/\langle N\rangle+1)^3-(2!)^2(1/\langle N\rangle+1)^2)/M$$

$$=4(2/\langle N\rangle^3+17/\langle N\rangle^2+10/\langle N\rangle+1)/M$$

Therefore $\mathrm{sd}\langle I^2\rangle/\langle I\rangle^2= (4(2/\langle N\rangle^3+17/\langle N\rangle^2+10/\langle N\rangle+1)/M)$

as $\langle N\rangle\to\infty$ $\mathrm{sd}\langle I^2\rangle/\langle I\rangle^2\to 2/\surd M$ as for constant amplitude scatterers

A5.4 Determining the standard deviation in $\langle N\rangle$

To obtain the standard deviation in $\langle N\rangle$ use the fact that $\langle N\rangle=(\langle a^4\rangle/\langle a^2\rangle^2)/(\langle I^2\rangle/\langle I\rangle^2-2)$ to get the $sd\langle N\rangle$ in terms of $sd\langle I^2\rangle/\langle I\rangle^2$

Let $g=\langle N\rangle$, $k=\langle a^4\rangle/\langle a^2\rangle^2$, $z=\langle I^2\rangle/\langle I\rangle^2$ then $g=k/(z-2)$ where k is assumed to be a constant.

Therefore $\sigma_g^2=(dg/dz)^2\sigma_z^2$ (pg 152,Papoulis)

Therefore $\sigma_g=|dg/dz|\sigma_z$ where $\sigma_z=sd\ \langle I^2\rangle/\langle I\rangle^2$

$\sigma_g=sd\ \langle N\rangle$

$(dg/dz)=k(-1)(z-2)^{-2}$ therefore $\sigma_g=(k/(z-2)^2)\sigma_z=(k/(z-2))^2\sigma_z/k=g^2\sigma_z/k$

ie $sd\langle N\rangle=\langle N\rangle^2(sd\langle I^2\rangle/\langle I\rangle^2)/(\langle a^4\rangle/\langle a^2\rangle^2)$

Thus for constant amplitude scatter the standard deviation in $\langle N\rangle$ equals $\langle N\rangle^2\sqrt{((4+20/\langle N\rangle+13/\langle N\rangle^2+1/\langle N\rangle^3)/M)}$ since $\langle a^4\rangle/\langle a^2\rangle^2=1$

and $sd\langle N\rangle=\langle N\rangle^2\sqrt{((2/\langle N\rangle^3+17/\langle N\rangle^2+10/\langle N\rangle+1)/M)}$ for Rayleigh amplitude statistics since $\langle a^4\rangle/\langle a^2\rangle^2=2$

A5.5 Determining the Standard Deviation in $\langle I\rangle$ for the purpose of block averaging

When doing block averaging there is a need to determine how many independant points M are needed in the calculation of $\langle I\rangle$.

variance in $\langle I\rangle=\sigma_y^2=(\langle I^2\rangle-\langle I\rangle^2)$

Therefore $sd\langle I\rangle=\sqrt{((\langle I^2\rangle-\langle I\rangle^2)/M)}$

sd as a fraction of $\langle I\rangle=(sd\langle I\rangle)/\langle I\rangle=\sqrt{((\langle I^2\rangle/\langle I\rangle^2-1)/M)}$

$=\sqrt{(((\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+2-1)/M)}$

$=\sqrt{(((\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+1)/M)}$

Thus $sd\langle I\rangle=\langle I\rangle\sqrt{(((\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+1)/M)}$

As $\langle N\rangle\to\infty$ the sd as a fraction of $\langle I\rangle=1/\sqrt{M}$

The $sd\langle I\rangle$ as a percentage of $\langle I\rangle=\%sd\langle I\rangle=100\sqrt{((2/\langle N\rangle+1)/M)}$

<N>	%sd <I>
.1	45%
.5	22%
1.0	17%
2.0	14%
4.0	12%

Table A5.5.1

M	%sd<I>
50	20%
100	14%
200	10%
300	8%
500	6%
1000	4%

Table A5.5.2

It can be seen from the tables and the expressions that the %sd<I>decreases with increasing density and increasing number of statistically independant points. It is important to note that although the %sd<I> decreases with increasing density the sd<I>may increase with increasing density.

A5.6 Determining the Standard Deviation in <N> For a Non-homogenous Density Shoal

When working with a shoal in which density varies ie one that does not have a constant mean number density the shoal is divided into regions of constant mean number density and then the average density is found for the whole shoal using $\langle N_{av}\rangle=(\Sigma m_i\langle N_i\rangle)/(\Sigma m_i)$ from section 2.9.3

The standard deviation in $\langle N_{av}\rangle=\mathrm{sd}\langle N_{av}\rangle=\sigma_{\langle N_{av}\rangle}$

The variance in $\langle N_{av}\rangle=\sigma^2_{\langle N_{av}\rangle}=\mathrm{var}(\Sigma(m_i/\Sigma m_i)\langle N_i\rangle)$

From pg 211,241 Papoulis if $x_1,x_2,x_3 \ldots\ldots\ldots x_n$ are uncorrelated.

$\sigma^2{}_{x_1+x_2+x_3+..x_n}=\sigma^2_{x_1}+\sigma^2_{x_2}+\sigma^2_{x_3}+\ldots\ldots\ldots+\sigma^2_{x_n}$

For $x=at$ $\quad \sigma^2_x=E(x^2)-E^2(x)=E(a^2t^2)-E^2(at)=a^2E(t^2)-a^2E^2(t)=a^2(E(t^2)-E^2(t))=a^2\sigma^2_t$

Therefore $\mathrm{var}(\Sigma(m_i/\Sigma m_i)\langle N_i\rangle)=\Sigma\mathrm{var}((m_i/\Sigma m_i)\langle N_i\rangle)=\Sigma(m_i/\Sigma m_i)^2\mathrm{var}\langle N_i\rangle$

$=(\Sigma m_i^2\mathrm{var}\langle N_i\rangle)/(\Sigma m_i)^2$

Therefore $\mathrm{sd}\langle N\ \rangle=\sigma_{\langle N_{av}\rangle}=\surd(\Sigma m_i^2(\mathrm{sd}\ \langle N_i\rangle)^2)/(\Sigma m_i)$

A5.7 Determining the Standard Deviations Which Include the Effects of Beamshape

The purpose of this appendix is to include the effect of a non-ideal beamshape in the expressions for standard deviation derived in the preceding sections of the appendices. This can be done by taking the generalised expression for the variance in the second normalised moment from the end of appendix A5.1.

$$\mathrm{var}(\langle I^2\rangle/\langle I\rangle^2)=(\langle I^4\rangle/\langle I\rangle^4-4(\langle I^3\rangle/\langle I\rangle^3)(\langle I^2\rangle/\langle I\rangle^2)$$
$$+4(\langle I^2\rangle/\langle I\rangle^2)^3-(\langle I^2\rangle/\langle I\rangle^2)^2)/M$$

Note that this expression is completely general and holds for any distribution and any scatterer statistics.
For a Poisson distribution and arbitrary scatterer type (from pg 2-12 of thesis)

$$\langle I^2\rangle/\langle I\rangle^2=(\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+2$$
$$\langle I^3\rangle/\langle I\rangle^3=(\langle a^6\rangle/\langle a^2\rangle^3)/\langle N\rangle^2+9(\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+6$$
$$\langle I^4\rangle/\langle I\rangle^4=(\langle a^8\rangle/\langle a^2\rangle^4)/\langle N\rangle^3+(16(\langle a^6\rangle/\langle a^2\rangle^3)$$
$$+18(\langle a^4\rangle^2/\langle a^2\rangle^4))/\langle N\rangle^2+72(\langle a^4\rangle/\langle a^2\rangle^2)/\langle N\rangle+24$$

replacing $\langle a^n\rangle$ by $\langle a^n\rangle\langle b^n\rangle$ as in section 2.4 pg 2-17

for convenience let $(\langle b^4\rangle/\langle b^2\rangle^2)(\langle a^4\rangle/\langle a^2\rangle^2)=x$

$(\langle b^6\rangle/\langle b^2\rangle^3)(\langle a^6\rangle/\langle a^2\rangle^3)=y$

$(\langle b^8\rangle/\langle b^2\rangle^4)(\langle a^8\rangle/\langle a^2\rangle^4)=z$

and $\langle N\rangle=n$

Thus $\langle I^2\rangle/\langle I\rangle^2=x/n+2$

$\langle I^3\rangle/\langle I\rangle^3=y/n+9x/n+6$

$\langle I^4\rangle/\langle I\rangle^4=z/n^3+((16y+18x^2)/n^2)+72x/n+24$

Therefore $\mathrm{var}\{\langle I^2\rangle/\langle I\rangle^2)=[(z-4xy+4x^3)/n^3+(5x^2+8y)/n^2+20x/n+4]/M$

$$\left.\begin{array}{l}\text{For constant amplitude scatterers } x=y=z=1\\ \text{For Rayleigh scatterers } x=2;y=6;z=24\end{array}\right\}\text{for an ideal beamshape}$$

Using the same program for the same circular transducer from APPENDIX4.1 and choosing n=6 yields $\langle b^6 \rangle = 7.59123\text{x}10^{-4}$

choosing n=8 yields $\langle b^8 \rangle = 5.77504\text{x}10^{-4}$

This gives $\langle b^6 \rangle / \langle b^2 \rangle^3 = 78750$ and $\langle b^8 \rangle / \langle b^2 \rangle^4 = 28149623$

Using the approximate expression for $\langle b^n \rangle$ for a rectangular transducer in APPENDIX4.2 one gets the following expression

$$\langle b^{2m} \rangle / \langle b^2 \rangle^m = (3/(4m))(1/(\pi^2\ 36)/(3/(2.2)^m\ (1/(\pi^2\ 36)^m$$

$$= (48\pi^2)^{m-1}/m$$

Putting m=3 gives $\langle b^6 \rangle / \langle b^2 \rangle^3 = (48\pi^2)^2/3 = 74810$

m=4 gives $\langle b^8 \rangle / \langle b^2 \rangle^4 = (48\pi^2)^3/4 = 26580488$

Thus x,y,z can be worked out for different scatterer types.

For constant amplitude scatterers $\langle a^{2m} \rangle / \langle a^2 \rangle^m = 1$

Thus $x = \langle b^4 \rangle / \langle b^2 \rangle^2$, $y = \langle b^6 \rangle / \langle b^2 \rangle^3$, $z = \langle b^8 \rangle / \langle b^2 \rangle^4$

For Rayleigh scatterers $x = 2.\langle b^4 \rangle / \langle b^2 \rangle^2$; $y = 6.\langle b^6 \rangle / \langle b^2 \rangle^3$; $z = 24.\langle b^8 \rangle / \langle b^2 \rangle^4$

Thus the constants x,y,z can be evaluated for each scatterer type and transducer and substituted into the expression for $\text{var}(\langle I^2 \rangle / \langle I \rangle^2)$.

Thus $\text{sd}\langle N_e \rangle = \langle N_e \rangle \sqrt{(\text{var}(\langle I^2 \rangle / \langle I \rangle^2))/x}$ where $\langle N_e \rangle$ is the equivalent number of fish in a hemisphere.(see pg 2-17)

Thus sd $\rho_{stat} = \text{sd}\langle N_e \rangle / (\pi r_g^2 c\tau)$ and $\rho_{stat} = \langle N_e \rangle / (\pi r_g^2 c\tau)$

and fish number density $= \rho_{stat}$ +/- $\text{sd}\rho_{stat}$

APPENDIX 6

The purpose of this appendix is to give a listing and explanation of the program used to analyse data.
In the program the following variables are used:
SAMPLE(X,Y) is the array in which the intensity data is stored with X being the ping number and Y the sample number in the ping.
L(K) holds the upper level of the K'th intensity band.
SUMISQ(K) is the sum of intensity data points squared for the K'th intensity band.
SUMI(K) is the sum of the intensity data points for the K'th intensity band.
MS(K) is the total number of intensity points in the K'th band.
NOLEV is the number of intensity bands or regions in the shoal.
NAV(K) is the average number of fish per cell in the K'th intensity band.
EXT enables an extension of the number of bands above the bands specified by NOLEV without changing the seperation between the levels seperating the bands.
CELVOL is the equivalent cell volume in cubic metres.

A description of the algorithm follows:
Lines 142 to 155 read in the stored values of Amplitude into SAMPLE(X,Y) and square them to convert them into intensities and calculates the average intensity IAV for the whole shoal.
Lines 170 to 210 sets the position of the levels which seperate the bands or regions.
Lines 220 and 230 process each ping sample by sample for data containing 80 pings and 200 samples per ping.

Lines 250 to 290 calculates the average values of intensity around SAMPLE(X,Y) and stores the results in BLOKAV. Lines 250 and 260 process all samples in an averaging block of 9 pings by 19 samples and hence determines the size of the averaging block which in this case contains 19x9=171 points. Lines 300 to 365 determine in which band the average value of intensity surrounding (X,Y) falls and once it has found the band sums it with SUMISQ( ), SUMI( ) and MS( ) for that band. Lines 370 to 520 calculates the average value of number of fish per cell for each band and then for the whole shoal.

```
      REAL SAMPLE(100,200),SUMISQ(10),SUMI(10),MS(10)
      INTEGER L(10)
      INTEGER Y,X,I,J,K,NOLEV
      REAL NAV(10),N
      TOTAL=0.
      SUMNMS=0
      SUMMS=0
      DO 1 X=1,80
      DO 2 Y=1,200
      READ(11,*) I,J,SAMPLE(X,Y)
      SAMPLE(X,Y)=SAMPLE(X,Y)*SAMPLE(X,Y)
      TOTAL=TOTAL+SAMPLE(X,Y)
2     CONTINUE
1     CONTINUE
      IAV=TOTAL/16000
       EXT=2
      NOLEV=4
       KNOLEV=NOLEV+EXT
      L(1)=0
      DO 8 I=1,KNOLEV
      L(I+1)=2*I*IAV/NOLEV
8     CONTINUE
       L(KNOLEV+1)=65025
      DO 5 Y=10,191

      DO 5 X=5,76
      BLOKAV=0
      DO 3 J=-9,9
      DO 3 I=-4,4
      BLOKAV=BLOKAV+SAMPLE(X+I,Y+J)
3     CONTINUE
      BLOKAV=BLOKAV/171
      DO 5 K=1,KNOLEV
      IF((L(K).LE.BLOKAV).AND.(BLOKAV.LT.L(K+1)))THEN
          SUMISQ(K)=SUMISQ(K)+SAMPLE(X,Y)*SAMPLE(X,Y)
          SUMI(K)=SUMI(K)+SAMPLE(X,Y)
          MS(K)=MS(K)+1
      ELSE
      ENDIF
5     CONTINUE
      DO 7 K=1,KNOLEV
      IF(SUMI(K).NE.0)THEN
         NAV(K)=2/(MS(K)*SUMISQ(K)/SUMI(K)/SUMI(K)-2)
         ELSE
         NAV(K)=0
      ENDIF
      SUMNMS=SUMNMS+NAV(K)*MS(K)
      SUMMS=SUMMS+MS(K)
      PRINT *,NAV(K),MS(K),SUMISQ(K),SUMI(K)
7     CONTINUE
      CELVOL=4.86
      PRINT *,SUMNMS,SUMMS,CELVOL
      P=SUMNMS/SUMMS/CELVOL
      N=SUMNMS/SUMMS
      PRINT *,P,N
      PRINT *,SAMPLE(1,1),SAMPLE(80,200),SAMPLE(77,192)
      END
```

APPENDIX 7

International Symposium on Fisheries Acoustics
June 22-26, 1987 Seattle, Washington, U.S.A.

Experiments with an Acoustic Technique for Estimating Fish Number Density, not Requiring Fish Target Strength.

by

P.N. Denbigh and Q. Smith

Central Acoustics Laboratory
Department of Electrical and Electronic Engineering
University of Cape Town, Rondebosch 7700, South Africa.

Abstract

The average number of fish in the resolution cell of an echo sounder determines not only the strength of the return signal but also its statistical character. A suitable measure of the statistics can lead to an estimate of fish number density which does not require any prior knowledge of the fish target strength. The appropriate statistical measure is obtained from the experimental variations of the intensity of the echo sounder signal. Previous experiments have shown that the technique has been successful under test tank conditions but considerably in error with real fish shoals. This paper examines the causes of discrepancies in the latter case and shows how the data may be handled differently to produce more accurate and meaningful estimates of fish number density.

Resume

Le nombre moyen de poissons dans le volume échantillonné par un echo-sondeur détermine non seulement l'intensité de l'écho mais aussi ses propriétés statistiques. En utilisant une mesure statistique appropriée il est possible d'estimer la densité en poissons sans savoir à priori l'intensité de la cible. Cette mesure statistique est derivée à partir de variations d'intensité du signal d'écho. Des experiences précédentes ont demontré que cette méthode donne des résultats satisfaisants dans les conditions idéales d'un reservoir de laboratoire, mais non dans de réels bancs de poissons. Cette article étudie les causes de désaccord dans ce dernier cas et montre comment les donneés expérimentales peuvent être utilisees différemment afin d'obtenir une estimation plus précise de la densité en poissons.

List of Symbols (excluding Appendix)

a	amplitude (envelope) of the received echo from an individual fish
b	insonification amplitude taking account of transmit and receive
D	amplitude directivity factor of transducer
e	amplitude of fish echo assuming unity insonification
$E(\ )$	expected value
I	intensity (=square of envelope)
I_{av}	average of all the intensity measurements
$J_1(\ )$	Bessel function of 1st kind and order one
k_i	number of measurements of level i
M	total number of independent measurements
n	parameter of Binomial distribution
N	number of fish in resolution cell

p	parameter of Binomial distribution
$P(\ )$	probability of an outcome
q	parameter of Binomial distribution
r	range
r_g	geometric mean of minimum and maximum ranges
α	phase of echo
γ	property of fish scattering $= \sigma_c / \sigma_d$
θ	angle from boresight of transducer
λ	wavelength
ρ	fish number density
σ	standard deviation of a Gaussian distribution
σ_c	backscattering cross-section of concentrated scattering component of fish
σ_d	backscattering cross-section of distributed scattering component of fish
τ	pulse duration
ϕ	angle from boresight of a transducer
ω	radian frequency of transmit pulse
Ω	solid angle of transducer beam corresponding to 3dB transmit beamwidth
$<\ >$	average value

1. Introduction and summary of previous work

Two main techniques of acoustic fish stock assessment exist. Echo counting can be used when the number of fish in the resolution cell is usually less than one but unfortunately this condition is rarely met. Echo integration is the technique used when the individual fish echoes overlap. The drawback is that an interpretation of the integrated acoustic energy requires a knowledge of fish target strength. A new technique of fish stock assessment has been described by Wilhelmij and Denbigh (1984). It is aimed at producing estimates of the number density of fish in a shoal and can tolerate overlapping echoes while not requiring any knowledge of target strength. It works on the principle that the statistics of the return envelope are determined by the average number of echoes overlapping in the return signal. For example, the central-limit theorem predicts that a very large number of overlapping echoes will result in an instantaneous amplitude which has Gaussian statistics (i.e. Rayleigh envelope statistics). These same statistics cannot be expected if the number of overlapping echoes is small. A suitable measure of the statistical properties is the second normalized moment of intensity (where intensity is defined as the square of the envelope), as this has a very simple mathematical dependence on the mean number <N> of scatterers causing the return signal. The relationship is

$$\langle N\rangle = 1/(\langle I^2\rangle/\langle I\rangle^2 - 2) \qquad (1)$$

where $\langle I^2\rangle/\langle I\rangle^2$ is the second normalized moment of intensity. This formula was verified experimentally by recording and analyzing the return signals when a random matrix of polystyrene spheres in a tank was insonified with high frequency tone bursts. Here <N> referred to the mean number of scatterers in a resolution cell, as this was the average number of scatterers contributing to the return signal at any instant. The results confirmed that the method was suitable for producing sensible estimates of number density.

The above formula used by Wilhelmij and Denbigh arose from a publication by Pusey et al (1974) and involved various assumptions which were appropriate to the model but were questionable for real fish shoals. One of the assumptions was that the echoes from all scatterers were of equal amplitude. Even for fish of constant size this is known to be incorrect, and Denbigh and Weintroub (1986) derived a modified formula for a Rayleigh distribution of echo amplitudes. This differed by a factor of two and was

$$\langle N\rangle = 2/(\langle I^2\rangle/\langle I\rangle^2 - 2) \qquad (2)$$

To determine the number density consider the resolution cell to be $c\tau/2$ in range by $r^2\Omega$ in width, where c is the sound speed, τ is the pulse duration, r is the range and Ω is the solid angle corresponding to the 3dB beamwidth of the sonar in its transmit mode. This gives $\langle N\rangle = \rho r^2\Omega c\tau/2$

$$\therefore \quad \rho = \frac{2}{(\langle I^2\rangle/\langle I\rangle^2 - 2)(0.5c\tau r^2\Omega)} \qquad (3)$$

One of the difficulties of applying this formula can be the variation of the size of resolution cell due to the divergence of the beam. Based upon a derivation by Wilhelmij and Denbigh, which took account of beam divergence, the above formula can be extended to produce the following estimate of number density.

$$\rho = \frac{2}{(\langle I^2\rangle/\langle I\rangle^2_{AV} - 2)(0.5\tau c\, r_g^2\Omega)} \qquad (4)$$

where $\langle I^2\rangle/\langle I\rangle^2_{av}$ is obtained by determining $\langle I^2\rangle/\langle I\rangle^2$ for each and every range and then averaging these values from the different ranges. Compared to Eq. 3 the change to the second term in the denominator is equivalent to changing the resolution cell to $0.5c\tau r_g^2\Omega$ where r_g is the geometric mean of the minimum and maximum ranges.

In 1985 the method was tested at sea against estimates of density obtained by echo-integration (Denbigh and Weintroub, 1986). The targets were night-time layers of anchovy, which were less dense than day-time shoals, and therefore more suited to the statistical method, which requires that the average number of shoals in the resolution cell is fairly small. The integration estimates were obtained from a 38-kHz Simrad EK-5 sounder interfaced to a multi-channel digital integrator. Calibration was by standard sphere and the target strength relationship used an expression for herring obtained by Halldorsson and Reynisson (1983). Further details of the integration methodology are given in Hampton (1987). Data for the statistical method were collected from a 120-kHz Simrad EK-5 sounder synchronised to the 38-kHz sounder. Target identification and sizing was achieved by aimed midwater trawling.

In the best-controlled of the five experiments conducted, 80 of the 120-kHz echo returns (from a total of 3089), were analyzed by the statistical method. The resultant number density estimate was compared with that obtained by integrating all 3089 corresponding 38-kHz returns between approximately the same range limits.

The outcome of this experiment was that the statistical estimate was 3.8 times lower than that obtained by integration. (A similar trend was found in the other experiments, the factors ranging between 3.1 and 5.1.) It is unlikely that the integration method could have been positively biased to this extent as, if integrator survey date were corrected accordingly, the resultant estimate of the anchovy population would be lower than the actual commercial catch made in the two-month period following the survey (I. Hampton, Sea Fisheries Research Institute, pers. comm.). The statistical estimates would therefore appear to have been in error.

The purpose of this present paper is to examine more closely all the theoretical assumptions and, where necessary, to produce more applicable and rigorous formulae. These are then applied to the same experimental data with the intention of producing better agreement with the echo integrator estimates. Although at this stage, the usefulness or otherwise of the technique is being judged by a comparison with results from the echo integrator, evidence is produced to suggest that, subject to confirmation from further experimentation, it might be able to stand alone and be capable of producing reliable estimates of number density, while not requiring a knowledge of target strength.

2. Theoretical Assumptions

Eqns 2 and 3 form the basis of the estimates and their derivations have been given by Denbigh and Weintroub. The following assumptions were made:

(1) the amplitudes of echoes from individual fish have Rayleigh statistics and the echo sounder beam is, in effect, a conical beam whose beamwidth equals the 3dB beamwidth of the transducer in its transmit mode.

(2) there is negligible multiple scattering

(3) there are enough independent measurements of intensity to produce an estimate of the second normalized moment of intensity which is accurate enough to produce a reliable estimate of number density.

(4) the fish within the shoal follow a Poisson volume distribution.

(5) the fish shoal is homogeneous such that the mean amplitudes and mean number densities apply to each and every resolution cell.

These assumptions will be examined. Firstly, however, a new more generalized theory is presented which is suited to predicting the relationship between number density and second normalized moment of intensity when some of these assumptions are modified.

3. Theory

Consider that there are N overlapping echoes, the kth one given by $a_k \cos(\omega t + \alpha_k)$ where a_k and α_k are random variables. Writing each in terms of its in-phase and quadrature components, I and Q, the resultant intensity, $I^2 + Q^2$, can be written as

$$I_N = \left(\sum_{k=1}^{N} a_k \cos \alpha_k\right)^2 + \left(\sum_{k=1}^{N} a_k \sin \alpha_k\right)^2$$

$$= \sum_{k=1}^{N}\sum_{\ell=1}^{N} a_k a_\ell \cos\alpha_k \cos\alpha_\ell + \sum_{k=1}^{N}\sum_{\ell=1}^{N} a_k a_\ell \sin\alpha_k \sin\alpha_\ell$$

$$= \sum_{k=1}^{N} a_k^2 \cos^2\alpha_k + \sum_{\substack{k=1\\k\neq\ell}}^{N}\sum_{\ell=1}^{N} a_k a_\ell \cos\alpha_k \cos\alpha_\ell$$

$$+ \sum_{k=1}^{N} a_k^2 \sin^2\alpha_k + \sum_{\substack{k=1\\k\neq\ell}}^{N}\sum_{\ell=1}^{N} a_k a_\ell \sin\alpha_k \sin\alpha_\ell$$

Making use of the relationships $\cos^2 x + \sin^2 x = 1$, and $\cos x.\cos y + \sin x.\sin y = \cos(x-y)$, this becomes

$$I_N = \sum_{k=1}^{N} a_k^2 + \sum_{\substack{k=1\\k\neq\ell}}^{N} \sum_{\ell=1}^{N} a_k a_\ell \cos(\alpha_k-\alpha_\ell)$$

$$\therefore \langle I_N \rangle = \sum_{k=1}^{N} a_k^2 = N\langle a^2 \rangle \qquad (5)$$

where $\langle a^2 \rangle$ is the mean value of a^2 in the resolution cell. Similarly

$$\langle I_N^2 \rangle = \langle \{ \sum_{k=1}^{N} a_k^2 + \sum_{\substack{k=1\\k\neq\ell}}^{N} \sum_{\ell=1}^{N} a_k a_\ell \cos(\alpha_k-\alpha_\ell) \}^2 \rangle$$

The mean of each crossproduct term between the first and second of these summations equals zero. Therefore $\langle I_N^2 \rangle$ reduces to

$$\langle I_N^2 \rangle = \langle (\sum_{k=1}^{N} a_k^2)^2 \rangle + \langle (\sum_{\substack{k=1\\k\neq\ell}}^{N} \sum_{\ell=1}^{N} a_k a_\ell \cos(\alpha_k-\alpha_\ell))^2 \rangle$$

$$= \langle \sum_{k=1}^{N} a_k^4 \rangle + \langle \sum_{\substack{k=1\\k\neq\ell}}^{N} \sum_{\ell=1}^{N} a_k^2 a_\ell^2 \rangle + \langle \sum_{\substack{k=1\\k\neq\ell}}^{N} \sum_{\ell=1}^{N} 2a_k^2 a_\ell^2 \cos^2(\alpha_k-\alpha_\ell) \rangle$$

The reduction in the last double summation is best understood by considering a simple example of N=3.

$$\begin{aligned}\langle (\sum_{k=1}^{3} \sum_{\ell=1}^{3} a_k a_\ell \cos(\alpha_k-\alpha_\ell))^2 \rangle &= \langle \{ a_1 a_2 \cos(\alpha_1-\alpha_2) + a_1 a_3 \cos(\alpha_1-\alpha_3) \\ &\quad + a_2 a_1 \cos(\alpha_2-\alpha_1) + a_2 a_3 \cos(\alpha_2-\alpha_3) \\ &\quad + a_3 a_1 \cos(\alpha_3-\alpha_1) + a_3 a_2 \cos(\alpha_3-\alpha_2) \}^2 \rangle \\ &= \langle \{ 2a_1 a_2 \cos(\alpha_1-\alpha_2) + 2a_1 a_3 \cos(\alpha_1-\alpha_3) + 2a_2 a_3 \cos(\alpha_2-\alpha_3) \}^2 \rangle \\ &= \langle \sum_{\substack{k=1\\k\neq\ell}}^{3} \sum_{\ell=1}^{3} 2a_k^2 a_\ell^2 \cos^2(\alpha_k-\alpha_\ell) \rangle\end{aligned}$$

Continuing the evaluation of $<I_N>$ and noting that $<a_k^2\, a_\ell^2> \;=\; <a^2>^2$ if the amplitudes are independent random variables; also that each double summation has N(N-1) terms

$$\therefore <I_N^2> \;=\; N<a^4> + N(N-1)<a^2>^2 + N(N-1)<a^2>^2$$

$$<I_N^2> \;=\; N<a^4> + 2N(N-1)<a^2>^2 \tag{6}$$

where $<a^2>$ and $<a>$ are the mean values of a^2 and a in the resolution cell.

I_N is the intensity when there are N echoes overlapping, i.e.when there are N scatterers in the resolution cell causing the return signal. If the mean number of scatterers in the resolution cell is <N>, let $P_{<N>}(N)$ denote the probability that there are N scatterers in a resolution cell. Considering the mean intensity and the mean square intensity when all resolution cells are considered, we obtain, so long as all cells containing N scatterers have the same $<I_N>$

$$<I> \;=\; \sum_{N=1}^{\infty} <I_N> P_{<N>}(N) \tag{7}$$

and $$<I^2> \;=\; \sum_{N=1}^{\infty} <I_N^2> P_{<N>}(N) \tag{8}$$

Combining these with Eqs. 5 & 6

$$<I> \;=\; <a^2> \sum_{N=1}^{\infty} N\, P_{<N>}(N) \;=\; <a^2>.<N>$$

$$I^2 = \langle a^4 \rangle \sum_{N=1}^{\infty} N \; P_{\langle N \rangle}(N) + \langle a^2 \rangle^2 \sum_{N=1}^{\infty} 2N(N-1)P_{\langle N \rangle}(N)$$

$$= \langle a^4 \rangle \, . \langle N \rangle + 2\langle a^2 \rangle^2 \langle N^2 \rangle - 2\langle a^2 \rangle^2 \langle N \rangle$$

This leads to

$$\frac{\langle I^2 \rangle}{\langle I \rangle^2} = \frac{\langle N \rangle \left[\langle a^4 \rangle - 2\langle a^2 \rangle^2 \right] + 2\langle a^2 \rangle^2 \langle N^2 \rangle}{\langle a^2 \rangle^2 . \langle N \rangle^2}$$

$$= \frac{\frac{\langle a^4 \rangle}{\langle a^2 \rangle^2} - 2}{\langle N \rangle} + 2 \frac{\langle N^2 \rangle}{\langle N \rangle^2} \qquad (9)$$

It is interesting to note that assumptions concerning the distribution of echo amplitude effect the first term only, whereas those concerning the volume distribution effect the second term only.

4. Application of Theory to a Poisson Volume Distribution and Rayleigh Amplitude Statistics

If the mean number of scatterers in the resolution cell is <N>, the Poisson distribution predicts that the probability that there are N scatterers in the cell is given by

$$P_{\langle N \rangle}(N) = \frac{\langle N \rangle^N e^{-\langle N \rangle}}{N!}$$

Papoulis Sec. 5.4 (1965) gives the moments of the distribution

$$E(N^2) = E(N) + (E(N))^2$$

This present paper neglects the difference between expected and average values

$$\therefore \quad \langle N^2 \rangle = \langle N \rangle + \langle N \rangle^2 \tag{10}$$

Consider next a Rayleigh distribution for the amplitudes of individual scatterers i.e.

$$p(a) = \frac{a}{\sigma^2} \exp\left[-a^2/2\sigma^2 \right]$$

It is readily shown that

$$E(a^2) = 2\sigma^2 \simeq \langle a^2 \rangle \; ; \; E(a^4) = 8\sigma^4 \simeq \langle a^4 \rangle$$

Substituting (10) and (11) into (9) we obtain

$$\frac{\langle I^2 \rangle}{\langle I \rangle^2} = 0 + 2 \frac{\langle N \rangle + \langle N \rangle^2}{\langle N \rangle^2} = 2 + \frac{2}{\langle N \rangle}$$

$$\text{or } \langle N \rangle = \frac{2}{\langle I^2 \rangle / \langle I \rangle^2 - 2}$$

This is the same result as in Eq. 2. However, the extended theory of the previous section now allows an investigation of how any changes to the assumptions will modify this formula.

5. Effects of the Beam Pattern and of the Distribution of Backscattering Strengths.

The first assumption of Section 2 is that the amplitudes of the overlapping echoes have Rayleigh statistics and that the echo sounder beam is effectively a perfect conical beam whose beamwidth equals the 3dB beamwidth of the transducer in its transmit mode. This is a very arbitrary assumption. It would seem just as legitimate for example, to choose the equivalent beamwidth to be the 1.5 dB beamwidth as this would correspond to the 3dB beamwidth of the transducer when considered as a combined transmitter and receiver. The purpose of this section is to examine the validity of the assumption and to produce any modifications to Eqs. 2 and 3 that are needed for more accurate estimates of number density.

We define the insonification b as

$$b = D^2(\theta,\phi)$$

where the same transducer is used for transmission and reception. For a circular transducer of diameter d

$$D = \frac{2J_1\left(\frac{\pi d}{\lambda}\sin\phi\right)}{\frac{\pi d}{\lambda}\sin\phi} \qquad (12)$$

If we chose we could altenatively describe the insonification in terms of the probability density function of a fish, located randomly within the hemisphere below the transducer, encountering the insonification b.

The amplitude a of the received echo from an individual fish is the product of the fish echo amplitude e assuming unity insonification and the insonification b taking into account transmission and reception. i.e $a = be$. We wish to evaluate $<a^4>/<a^2>^2$ for insertion into Eq. 9. The fish echo amplitude e can also be described by a pdf and because b and e are independent we have $E(a^n) = E(b^n).E(e^n)$

$$\therefore \quad \frac{E(a^4)}{E(a)^2)^2} = \frac{E(b^4)}{E(b^2)^2} \cdot \frac{E(e^4)}{E(e^2)^2} \qquad (13)$$

Ehrenburg (1973, Fig.B1) gives plots of $E(b^4)$ and $E(b^2)$ for a circular transducer for different values of d/λ . As an example, $d/\lambda = 6$ gives $E(b^4)/E(b^2)^2 = 220$.

Consider next the quotient $E(e^4)/E(e^2)^2$. Clay and Heist (1984) show that scattering from fish can be considered to consist of a concentrated scattering component from the swim bladder described by a backscattering cross-section σ_c,and a distributed scattering component from the skeletal structure described by a backscattering crosssection σ_d. An important parameter is the ratio $\gamma = \sigma_c/\sigma_d$. They find that the situation is then analogous to considering the envelope of a constant amplitude sinusoid plus noise, which is the classic case treated by Rice(1944). Thus Clay and Heist find that the amplitude of the scattering from individual fish is described by a Rician distribution. If the concentrated scattering component is very small, as commonly happens if the fish is long in wavelengths, the distribution reduces to a Rayleigh distribution. However this is a special case situation. It is interesting to note that the theoretical model of Clay and Heist is based upon the changing interference from the scattering centres on the fish due to flexure and movement, and does not consider variations in fish size. However their experimental measurements of individual fish on Lake Michigan include the effects of size variation and do verify the Rician distribution. A derivation for evaluating $E(e^4)/E(e^2)^2$ for a Rician distribution of scattering strengths described by γ is given in the Appendix. Combining this with $E(b^4)/E(b^2)^2 = 220$ for $d/\lambda = 6$ gives

$$\frac{E(a^4)}{E(a^2)^2} = 220 \; \frac{\gamma^2+4\gamma+2}{(\gamma+1)^2}$$

Therefore, equating this result to $\langle a^4\rangle/\langle a^2\rangle^2$ we can substitute into Eq.9. Assuming a Poisson distribution as in Section 4;

$$\frac{\langle I^2\rangle}{\langle I\rangle^2} = \frac{220}{\langle N\rangle} \cdot \frac{\gamma^2+4\gamma+2}{(\gamma+1)^2} - \frac{2}{\langle N\rangle} + 2 + \frac{2}{\langle N\rangle}$$

$$\therefore \quad \langle N\rangle = 220. \frac{\gamma^2+4\gamma+2}{(\gamma+1)^2} \cdot \frac{1}{\langle I^2\rangle/\langle I\rangle^2-2}$$

Considering γ values between zero and infinity it may seem that this equation shows little correspondence to Eq. 2. The reason is that the two equations relate to resolution cells of very different size. The insonification pdf which produced the numerical value of 220 for $E(b^4)/E(b^2)^2$ was based upon the possibility of a fish being anywhere within a hemisphere. On this basis the size of the resolution cell would thus be $2\pi r^2(c\tau/2) = \pi r^2 c\tau$

$$\therefore \quad \rho = \frac{220}{\pi r^2 c\tau} \cdot \frac{\gamma^2+4\gamma+2}{(\gamma+1)^2} \cdot \frac{1}{\langle I^2\rangle/\langle I\rangle^2-2}$$

An equation suitable for comparison with Eq. 3 can be obtained by expressing this last result in terms of the solid angle Ω corresponding to the 3dB beamwidth. From Eq.12 we find D^2 is reduced to 0.5 when $\pi d/\lambda.\sin\phi = 1.617$. Using the relationship $\Omega=2\pi(1-\cos\phi)$ we obtain $\Omega=0.02316$, or $0.02316/\Omega = 1$.

Therefore, multiplying the above equation by $0.02316/\Omega$ for this case of $d/\lambda = 6$

$$\rho = \frac{1}{(\langle I^2\rangle/\langle I\rangle^2-2)(0.62c\tau r_g^2\,\Omega)} \cdot \frac{\gamma^2+4\gamma+2}{(\gamma+1)^2}$$

A repeat of the preceding derivation shows that other values of d/λ give an almost identical formula so long as Ω is modified to be the new beamwidth. In general therefore, including the use of the geometric mean of the minimum and maximum ranges in order to account for beam divergence, we obtain

$$\rho = \frac{\langle N\rangle}{0.62\, c\tau r_g^2\Omega} \qquad (14)$$

where

$$\langle N\rangle = \frac{1}{\langle I^2\rangle/\langle I\rangle^2-2} \cdot \frac{\gamma^2+4\gamma+2}{(\gamma+1)^2} \qquad (15)$$

and Ω is the solid angle corresponding to the 3dB beamwidth of the transducer as a transmitter.

Considering the case of $\gamma = 0$ for a Rayleigh distribution it is seen that Eq. 14 produces estimates which are less by a factor of 1.24 compared with Eq. 3.

The above derivation on its own fails to explain the discrepancies reported in

Section 2 between the estimates using this new statistical technique and those using echo integration. Indeed, by using Eq. 14, the discrepancy is increased, and particularly if $\gamma > 0$ for a non-Rayleigh amplitude distribution.

6. Multiple Scattering

In most of the experimental work previously used for number density estimates the fish shoal occupied a fairly narrow range of depths, and the echo integrator estimates of area number density were typically as low as 6 fish per square metre. The anchovy were about 12cm long so that the fish occupy only a very small fraction of the area. Multiple scattering seems improbable. Experimentally, this is verified because, to a good approximation the statistics of intensity were found to be stationary. It seems unlikely that multiple scattering could cause large errors in the estimates.

7. Number of Measurements

Wilhelmij and Denbigh (1984) included an analysis giving the fractional error $\Delta\langle N\rangle/\langle N\rangle$ in the estimates of the number of scatterers in the resolution cell, as a function of the number M of independent measurements of intensity. The formula obtained for identical scatterers was

$$\left|\frac{\Delta\langle N\rangle}{\langle N\rangle}\right| = \frac{2}{M^{\frac{1}{2}}} . \langle N\rangle$$

For Rayleigh scatterers the error would be reduced by two.

However, certain simplifying assumptions were made which rendered the derivation accurate only if the envelope statistics of the return waveform were closely Rayleigh, as would be the case only for large <N>. This condition is not applicable in the experiments reported and a much more satisfactory examination of the problem has been reported by Weintroub (1986). He does a computer simulation of a fish shoal for various values of <N> , assuming that individual fish have Rayleigh amplitude statistics and a Poisson volume distribution. He calculates the intensities as a function of range and of transducer position and then applies the statistical technique to estimate <N>. These estimates are then compared with the true values of <N>. An example of his simulation is shown in Table 1 for Rayleigh scatterers. 1000 independent measurements were used to perform the estimates and 100 independent simulations were undertaken to obtain good estimates of the mean and standard deviation of the estimates of <N>. It is interesting to note that a bias exists in the estimates. This arises because, considering the denominator term ($<I^2>/<I>^2$ - 2),too low an experimental value of $<I^2>/<I>^2$ can bring the value of $<I^2>/<I>^2$ very close to two, with the denominator term then approaching zero. This has much more impact on the estimate than an equal but positively biassed error in $<I^2>/<I>^2$.

The echo integrator estimates in all of the night time experiments corresponded to a mean number of fish in the resolution cell less than four. The statistical estimates were mostly based upon about 1000 independent measurements and it may be concluded from Table 1 therefore that the discrepancy reported by Denbigh and Weintroub between estimates by the echo integrator and by the statistical technique is too large to be explained by an insufficient number of measurements.

8. Volume Distribution of Fish

An important assumption in the theory leading to Eqns. 2 and 3 is that the fish are Poisson distributed in volume. The purpose of this section is to examine the sensitivity of the estimate to the validity of this assumption. It is very convenient to restrict the investigation to a Rayleigh distribution as then $\langle a^4\rangle/\langle a^2\rangle^2 = 2$ and the first term of Eq.9 equals zero. The remaining term depends solely on the volume distribution. Consider the reduced form of Eq.9 which is applicable for fish with Rayleigh amplitude statistics.

$$\frac{\langle I^2\rangle}{\langle I\rangle^2} = 2\,\frac{\langle N^2\rangle}{\langle N\rangle^2} \tag{16}$$

For a given value of <N> we wish to see how much the second normalized moment depends on the volume distribution assumed. Conversely, for a given experimental value of $\langle I^2\rangle/\langle I\rangle^2$, we wish to find how the estimate of <N> depends on the volume distribution assumed.

The way this will be done is by an example. Three distributions are considered, each having a true value of <N> = 3. In each case, the second normalized moment is evaluated using Eq.16. The estimate of <N> is made from this second normalized moment on the assumption of a Poisson volume distribution. Clearly the case of a Poisson distribution should be considered first.

Case 1 Poisson Distribution

This case was treated in Section 4 and led to the formula

$$\frac{\langle I^2\rangle}{\langle I\rangle^2} = 2 + \frac{2}{\langle N\rangle}$$

The example of $\langle N\rangle = 3$ gives $\langle I^2\rangle/\langle I\rangle^2 = 2.67$. Conversely, a measured value of $\langle I^2\rangle/\langle I\rangle^2 = 2.67$ would predict $\langle N\rangle = 3$ since the correct volume distribution is assumed. The distribution $P_{\langle N\rangle}(N)$ for $\langle N\rangle = 3$ is shown in Fig. 1a.

Case 2 Binomial Distribution

The expression for a binomial distribution is

$$P(N) = \frac{n!}{(n-N)!\ N!} \cdot p^n q^{n-N}$$

where $q = 1-p$.

From Papoulis, the first moment is np and the second moment is $n(n-1)p^2 + np$.

Consider a Binomial distribution with $n = 6$ and $p = 1/2$, such that the first moment $\langle N\rangle = 3$, as before.

$$\langle N^2\rangle = 6(6-1)(1/2)^2 + 3 = 10.5.$$

Therefore Eq. 16 becomes

$$\langle I^2\rangle/\langle I\rangle^2 = 2(10.5)/3 = 2.333.$$

Suppose that this is the experimentally determined value of the second normalized moment. If we are operating on the assumption of a Poisson distribution we apply the formula

$$\langle N \rangle = 2/(\langle I^2 \rangle / \langle I \rangle^2 - 2)$$

and obtain as our estimate that $\langle N \rangle = 6$. The error is considerable compared with the true value of 3. The Binomial distribution is plotted in Fig.1b and it is seen that it does not differ over-greatly from the Poisson distribution. The Poisson distribution has a better theoretical justification but, in spite of this, the sensitivity of this fish stock estimation technique to the assumptions of volume distribution is clearly revealed. It will be noted that the Binomial distribution causes too high an estimate. This is opposite to the error of the experiments. A substitution for other Binomial distributions shows that the more peaky the probability distribution, the greater is the error, though again always of the opposite polarity to the error in the experimental results.

Case 3 Uniform Distribution

Consider a uniform distribution with a mean value of 3 as before. As shown in Fig. 1c this signifies a probability of 1/7 that there are 0, 1, 2, 3, 4, 5 or 6 fish in any resolution cell. The second moment is given by

$$\langle N^2 \rangle = 1/7\ (0^2 + 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = 13$$

Therefore Eq. 16 becomes

$$\langle I^2\rangle/\langle I\rangle^2 \quad = \quad 2(13)/3^2 \quad = \quad 2.89.$$

Going backwards to predict $\langle N\rangle$ on the assumption of a Poisson distribution, we apply the formula $\langle N\rangle = 2/(\langle I^2\rangle/\langle I\rangle^2 - 2)$, and obtain $\langle N\rangle = 2.25$. This time our estimate is too low. It is 2.25, whereas it should be 3.

It is possible to generalize the two preceding cases to demonstrate that a volume distribution which is more peaky than the Poisson distribution will give too high an estimate, and that the error can be considerable, and that a volume distribution which is more uniform than the Poisson distribution will give too low an estimate, but that the error is relatively small. On the basis of these findings, it does not appear that deviations from the Poisson distribution can explain the discrepancies between experimental estimates and what are believed to be the true densities as determined by echo integration.

9. Inhomogoneities in the Fish Shoal

An inspection of the theoretical derivations of Section 3 shows that the quantities $\langle a^2\rangle$ and $\langle a^4\rangle$ refer to the mean values of a^2 and a^4 within a resolution cell. It is assumed that these mean values are equally applicable for all resolution cells. The quantities $\langle N\rangle$ and $\langle N^2\rangle$ refer to the mean values of $N^2$and N^4 within a resolution cell and it is assumed that they are also equally applicable for all resolution cells.

A close examination of the experimental data shows that, apart from the fine intensity structure that is to be expected from statistical variations, there is also a coarse structure. The most likely explanation is that there is a variation in fish number density throughout the shoal. It follows that the estimates previously reported by Denbigh and Weintroub were made under a false premise - that of homogoneity within the shoal. In fact it is inhomogeneous. The problem may be greatly reduced by dividing the shoal into regions of different intensity and applying the statistical technique of number density estimation to each separately. The procedure adopted has been as follows.

a) the average intensity I_{av} is determined using all data points from the shoal. This value is then used to derive a large set of intensity bands, 0 to $I_{av}/2$, $I_{av}/2$ to I_{av}, I_{av} to $3I_{av}/2$, and $3I_{av}/2$ to $2I_{av}$, $2I_{av}$ to $5I_{av}/2$, $5I_{av}/2$ to $3I_{av}$, etc.etc.. These bands are designated Band 1, Band 2, Band 3 etc.

b) at each point of the shoal the average of the intensities from the surrounding points is determined. Too many surrounding points should not be used in this averaging process as the spatial resolution within the fish shoal is then too poor. The use of too few points is even worse as the estimated intensity may then be determined by the statistical variations of intensity from point to point, and not by true regional variations. The use of about twenty independent points has been found to provide a very satisfactory compromise.

c) each data point is assigned to that band within which its averaged intensity falls.

d) the estimate of <N> is determined for each band by the statistical technique, using only the data points associated with that band.

e) the values of <N> for the different bands are averaged appropriately, taking into account the number of data points within each band, to find an overall average value of $\langle N\rangle$ for the whole shoal. If the estimates in any bands are clearly inaccurate,as will be seen to occur when there are too few data points in that band, the estimate in that band is replaced by an interpolated value.

Concerning the accuracy of the estimates it should be noted that the higher bands are prone to error, giving estimates of $\langle N\rangle$ which are too low. One problem is that more measurements are needed when $\langle N\rangle$ is large (see Sec. 7). In general there are less. The second problem concerns imperfections in the averaging procedure and is best understood by considering the simplified example of a shoal of perfectly uniform density. Because of the statistical variation of intensities between resolution cells some of the measurements of intensity can be expected to be very large, in spite of the uniformity of the shoal. Consider averaging a block of twenty intensities which happens to contain such an exceptionally large value. As the shoal is uniform the answer should fall within the band corresponding to the average for the whole shoal. However the one large value may well cause it to fall into a higher band. The statistical distribution of the intensities in this higher band will be different from the other band and will be incorrect. It will tend to contain an excess of large values because it has arisen from blocks containing one very large value among the twenty. Consider the effect of this on the value of $\langle I^2\rangle/\langle I\rangle^2$ in the higher band. We have $\langle I^2\rangle/\langle I\rangle^2$ equals $(1/20 \sum I^2)/(1/20 \sum I)^2$ or $20(\sum I^2)/(\sum I)^2$. If there is one very large dominant value of I among the twenty, $\sum I^2 \simeq (\sum I)^2$ and $\langle I^2\rangle/\langle I\rangle^2$ tends towards the value 20. This causes a meaningless and, with reference to Eq. 15, a very low estimate of $\langle N\rangle$ for that band. The likelihood of such situation is reduced by averaging more than twenty points in the first place but then the penalty is a degraded spatial resolution within the fish shoal. As stated previously the use of twenty independent points has been found to provide a good compromise for our experiments. Fewer than twenty has produced anomalous results, probably because the statistics in each band are upset by data points being allocated to the wrong band.

Experimental results are summarised in Table 2. The first column gives the number of the intensity band. The last entry in this column applies when all data points are used together i.e. when there is no quantization of the intensities into bands, which is equivalent to the situation reported previously by Denbigh and Weintroub. The second and third columns give, respectively, the corresponding estimates of <N>, using Eq. 15, and the number of data points used. The anchovy were mostly between nine and ten wavelengths long and, following the findings of Clay and Heist for fish long in wavelengths, a γ value of zero was used in deriving <N>. The fourth and fifth columns are similar except using only the even pings. The sixth and seventh columns use only the odd pings. It should be noted that, due to the time between samples being an eleventh of the pulse duration, only one data point in eleven is statistically independent.

For the data points corresponding to a particular band the mean intensity is at the centre of that band and, using the values from column 2 of Table 2, Fig. 2 shows a plot of the estimate of <N> for the first four bands versus the mid-point of each band. Its linearity attests to the validity of the technique. The estimates for higher bands are not shown in Fig. 2 but are not consistent with the first four; this deviation is attributed to too few data points and to the problem discussed above. It should be noted that all the higher bands together constitute only 14% if the total data. Further corroboration to the validity of the statistical technique comes from columns 4 and 6 of Table 2. These give estimates of <N> which are totally independent of one another as they arise from alternate pings. Nevertheless they agree reasonably well with one another for the first four bands and also with the estimates of column 2 using all the pings.

The estimate of overall number density is based upon the straight line of Fig. 2. It passes through the first point and it is considered therefore that there are 5059 data points predicting $\langle N\rangle = 0.39$. Based upon the straight line the second point appears to be slightly in error. Using a corrected value we have 1935 data point predicting $\langle N\rangle = 3 \times 0.39$. Continuing this procedure there are 2190 data point predicting $\langle N\rangle = 5 \times 0.39$, 2089 points predicting $\langle N\rangle = 7 \times 0.39$, 913 predicting $\langle N\rangle = 9 \times 0.39$, etc. Therefore, for the whole shoal

$$\langle N\rangle_{AV} = \frac{0.39(5059 + 3\times1935 + 5\times2190 + 7\times2089 + \ldots)}{5059 + 1935 + 2190 + 2089 + \ldots}$$

$$= 1.686$$

It will be noted that this estimate is very much higher than that of the last entry in Table 2 which corresponds to the case reported previously by Denbigh and Weintroub, where inhomogeneities were neglected.

Using Eq. 14, which includes beam pattern effects, $\rho = \langle N\rangle_{av}/(0.62c\,\tau\, r_g^2\, \Omega)$. The circular transducer used had a 10° beamwidth giving $\Omega = 2\pi(1 - \cos 5^\circ) = 0.0239$ steradians. The minimum and maximum ranges were 22.5m and 24.4m, giving as the geometric mean $r_g = 23.4$m. The pulse had sharp leading and trailing edges and a duration of 0.14ms.

$$\therefore \rho = \frac{1.686}{0.62 \times 1500 \times 0.00014 \times 23.4^2 \times 0.0239}$$

$$= 0.984 \text{ fish/m}^3$$

The echo integrator result from the Sea Fisheries Institute was 1.65 fish/m³. Although a discrepancy still exists it is greatly reduced compared with previously reported results. Bearing in mind that echo integration results can have errors, arising particularly from incorrect assumptions of fish target strength, it appears that, by dividing the shoal into regions of different average intensity, the statistical technique may now be producing useful estimates.

10. Conclusions

A technique has been developed which produces estimates of fish number density without requiring a knowledge of fish target strength. The theory described in earlier publications has been extended to take account of the transducer beam pattern and of Rician scattering from the individual fish. It has been shown that major errors can occur in the estimates if the fish shoal is inhomogeneous, but that, in the example discussed, estimates showing reasonable agreement with echo integrator results can be obtained if the fish shoal is first divided into regions of constant scattering strength. Much more experimentation is needed before any firm conclusions can be drawn but the results so far are very encouraging. The greatest limitation to the technique is that the shoal should be of sufficiently low density that only a few fish fall into the resolution cell. The use of short pulses and a narrow beam is clearly very beneficial. Short pulses have the added advantage that they generate more independent intensity measurements for the required statistical analysis. If further confirmation of the technique should be forthcoming it is possible that it could provide an alternative to echo integration for fish stock assessment. Alternatively it might be used in conjunction with an echo integrator, providing an independent estimate of number density in a low density region of the shoal and thereby enabling the target strength to be deduced for use by the echo integrator elsewhere in the shoal.

Acknowledgements

We would like to express our thanks to the Sea Fisheries Institute for the opportunity to participate in the cruise on R.S. Africana. In particular we would like to thank Mr Ian Hampton of for providing the echo integration results and for much encouragement and advice. We would also like to thank Mr Jonathan Weintroub, previously of this laboratory and an author of a previous paper on this topic, for producing the experimental data on which much of this paper is based.

Appendix

Clay and Heist show that the p.d.f. of the echo amplitude e due to a fish is

$$W_R(e) = \frac{2e}{\sigma_d} \exp\left[- \frac{e^2+\sigma_c}{\sigma_d} \right] I_o \left[\frac{2e\,\sigma_c^{\frac{1}{2}}}{\sigma_d} \right] \qquad \text{(A1)}$$

where σ_c is backscattering cross section of the concentrated scattering component and σ_d that of the distributed components. They also introduce a term γ where $\gamma = \sigma_c / \sigma_d$.

From Eq. 13 the requirement is to evaluate $<e^4>/<e^2>^2$

The solution to evaluating the second and fourth moments is found in Rice (1945). He considered the envelope R of a sine wave Pcos pt accompanied by noise. The pdf is

$$p(R) = \frac{R}{\psi_o} \exp\left[- \frac{R^2+P^2}{2\psi_o} \right] I_o\left[\frac{RP}{\psi_o}\right]$$

where Ψ_0 is the variance of the noise.

This equation becomes the same as Eq. (A1) after the following identities:

$$\frac{\sigma_d}{2} \equiv \psi_o \quad ; \quad e \equiv R \quad ; \quad \gamma \equiv \frac{\sigma_c}{\sigma_d} = \frac{P^2}{2\psi_o}$$

Rice shows (Eq. 3.10-12) that

$$\overline{R^n} = (2\psi_o)^{n/2} \Gamma(\frac{n}{2} + 1)\, {}_1F_1 (- \frac{n}{2} \; ; 1 \; ; - \frac{P^2}{2\psi_o})$$

where Γ represents the gamma function and ${}_1F_1$ a confluent hypergeometric function. After making use of the identities:

$$\overline{e^n} = (\sigma_d)^{n/2} \Gamma(\frac{n}{2} + 1)\, {}_1F_1 (- \frac{n}{2} \; ; 1 \; ; - \gamma)$$

Abramowitz and Stegun (1965) give an alternative notation for the confluent hypergeometric function, namely ${}_1F_1(a;b;z) = M(a,b,z)$. Applying this notation and using $\Gamma(m+1) = m!$

$$\overline{e^2} = \sigma_d \Gamma (2)M(-1, 1, -\gamma) = \sigma_d M(-1, 1, -\gamma)$$

and

$$\overline{e^4} = \sigma_d^2 \Gamma (3)M(-2, 1, -\gamma) = 2\sigma_d^2M(-2, 1, -\gamma)$$

But from Sec. 13.1.2 of Abramowitz and Stegun

$$M(a,b,z) = 1 + \frac{az}{b} + \frac{(a)_2}{(b)_2} \cdot \frac{z^2}{2!} + \;\ldots\; \frac{(a)_n}{(b)_n} \cdot \frac{z^n}{n!} + \ldots$$

where $(a)_n = a(a+1)(a+2) \ldots (a+n-1); \; (a)_0 = 1$

$$\therefore \frac{\overline{e^4}}{(\overline{e^2})^2} = \frac{2\sigma_d^2\left[1+(-2)(-\gamma)+\frac{(-2)(-1)}{(1)(2)} \cdot \frac{(-\gamma)^2}{2} + 0 \;\ldots\; +0\right]}{\sigma_d^2\left[1+\frac{(-1)}{(1)}(-\gamma) + 0 \;\ldots\ldots\ldots\; +0\right]^2}$$

$$\therefore \frac{\overline{e^4}}{(\overline{e^2})^2} = \frac{\gamma^2+4\gamma+2}{(\gamma+1)^2}$$

References

Abramowitz, M., and I.A. Stegun. 1965. Handbook of mathematical functions. Dover Publications Inc., New York.

Clay, C.S., and B.G. Heist. 1984. Acoustic scattering by fish - acoustic models and a two-parameter fit. J. Acoust. Soc. Am. 75(4):1077-1083.

Denbigh, P.N., and J. Weintroub. 1986. Fish stock assessment by a statistical analysis of echo sounder signals. 12th I.C.A. Congress Symposium on Underwater Acoustics, Halifax.

Ehrenburg, J.E. 1973. Estimation of the intensity of a filtered Poisson process and its application to acoustic assessment of marine organisms. Doctoral dissertation. Published by Division of Marine Resources, Univ. of Washington.

Halldorsson, O. and P. Reynisson. 1983. Target strength measurements of herring and capelin in situ at Iceland. In Notken, O. and S.C. Venema (eds) Symposium on fisheries acoustics. Selected papers of the ICES/FAO Symposium on fisheries acoustics. Bergen, Norway 21-24 June 1982. FAO Fish Rep. 300: 78-84.

Hampton, I. 1987 - Acoustic study on the abundance and distribution of anchovy spawners and recruits. In The Benguela and Comparable Ecosystems. Payne, A.I.L., Gulland, J.A. and K.H. Brink (eds). S. Afr. J. Mar. Sci. 5.

Papoulis, A. 1965. Probability, random variables and stochastic processes. McGraw-Hill.

Pusey, P.N., Schaefer, D.W., and D.E. Koppel. 1974. Single-interval statistics of light scattered by identical independent scatterers. J. Phys. A Math. Nucl. Gen. 7: 530-540.

Rice, S.O. 1944. Mathematical analysis of random noise. Bell System Tech. J. 23: 282-333.

Rice, S.O. 1945. Mathematical analysis of random noise. Bell System Tech. J. 24: 96-157

Weintroub, J. 1986. Fish stock assessment by a statistical analysis of echo sounder signals. M.Sc. Thesis. Univ. of Cape Town, Rondebosch 7700, R.S.A.

Wilhelmij, P. and P.N. Denbigh. 1984. A statistical approach to determining the number density of random scatterers from backscattered pulses. J. Acoust. Soc. Am. 76(6): 1810-1818.

<u>Table 1</u> Simulation results for Rayleigh scatterers; 100 trials were performed at each density using 1000 samples per trial.

True $< N >$	Mean $< I^2 > / < I >^2$	Mean est. of $< N >$	s.d. in est. of $< N >$
0.10	21.43	0.10	0.011
0.50	5.952	0.51	0.068
1.00	3.948	1.05	0.150
2.00	2.975	2.11	0.369
4.00	2.507	4.23	1.14

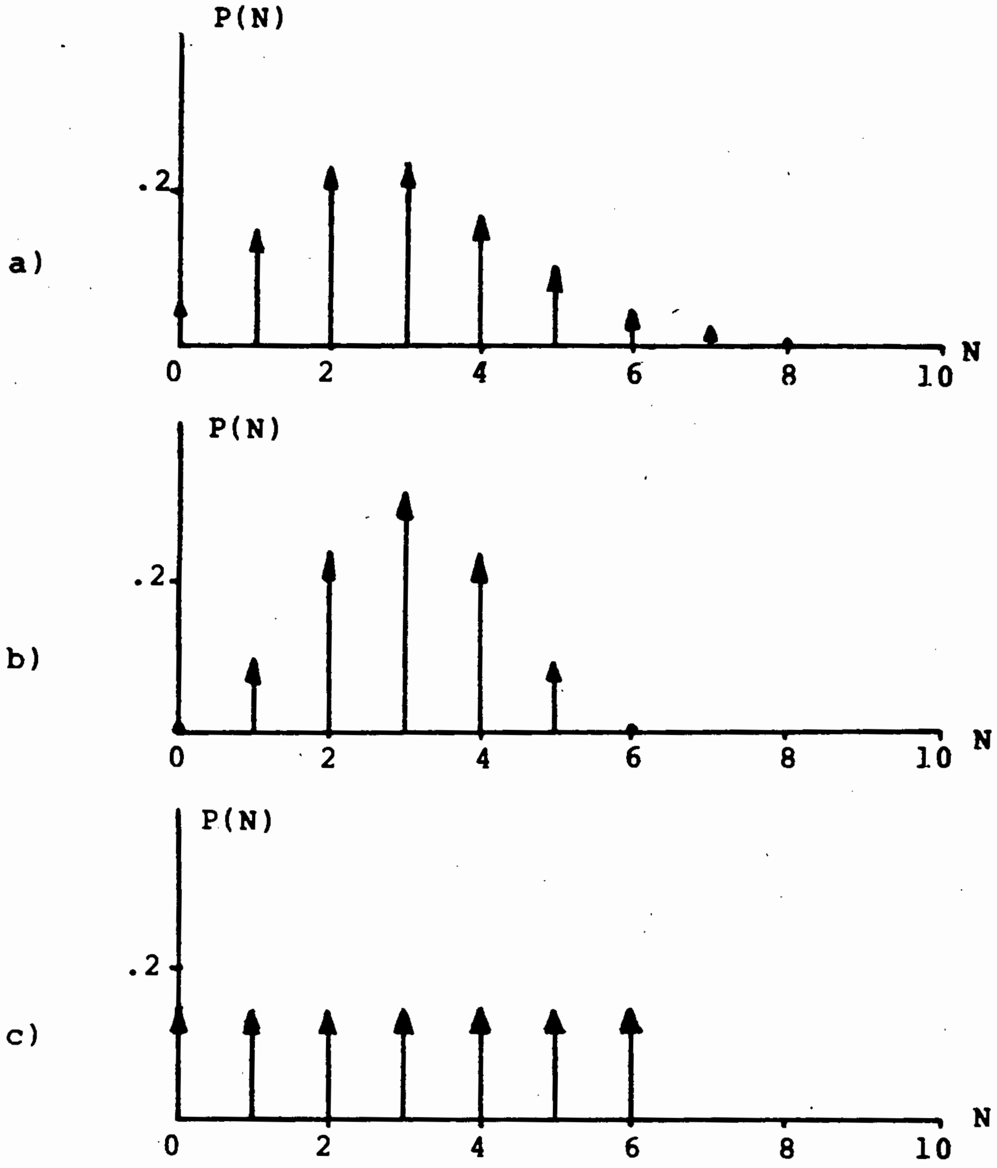

Fig.1 Probability distibutions having <N>=3

a) Poisson b) Binomial c) Uniform

Table 2 Estimates of <N> for different intensity regions of fish shoals.

Intensity Band	All pings		Even pings		Odd pings	
	<N>	No. of points	<N>	No. of points	<N>	No. of points
1	.39	5059	.39	2533	.40	2526
2	.93	1935	.98	985	1.05	950
3	1.93	2190	2.20	1097	1.73	1093
4	2.91	2089	2.18	1026	4.04	1063
5	2.25	913	1.97	434	3.05	479
6	1.69	454	2.17	232	2.03	222
7	1.21	211	1.96	121	1.85	90
8	3.01	107	5.12	45	2.54	62
9	6.00	63	4.06	30	-10.6	33
10	20.78	83	3.98	49	-6.6	34
ALL BANDS	.53	13104	.45	6552	.71	6552

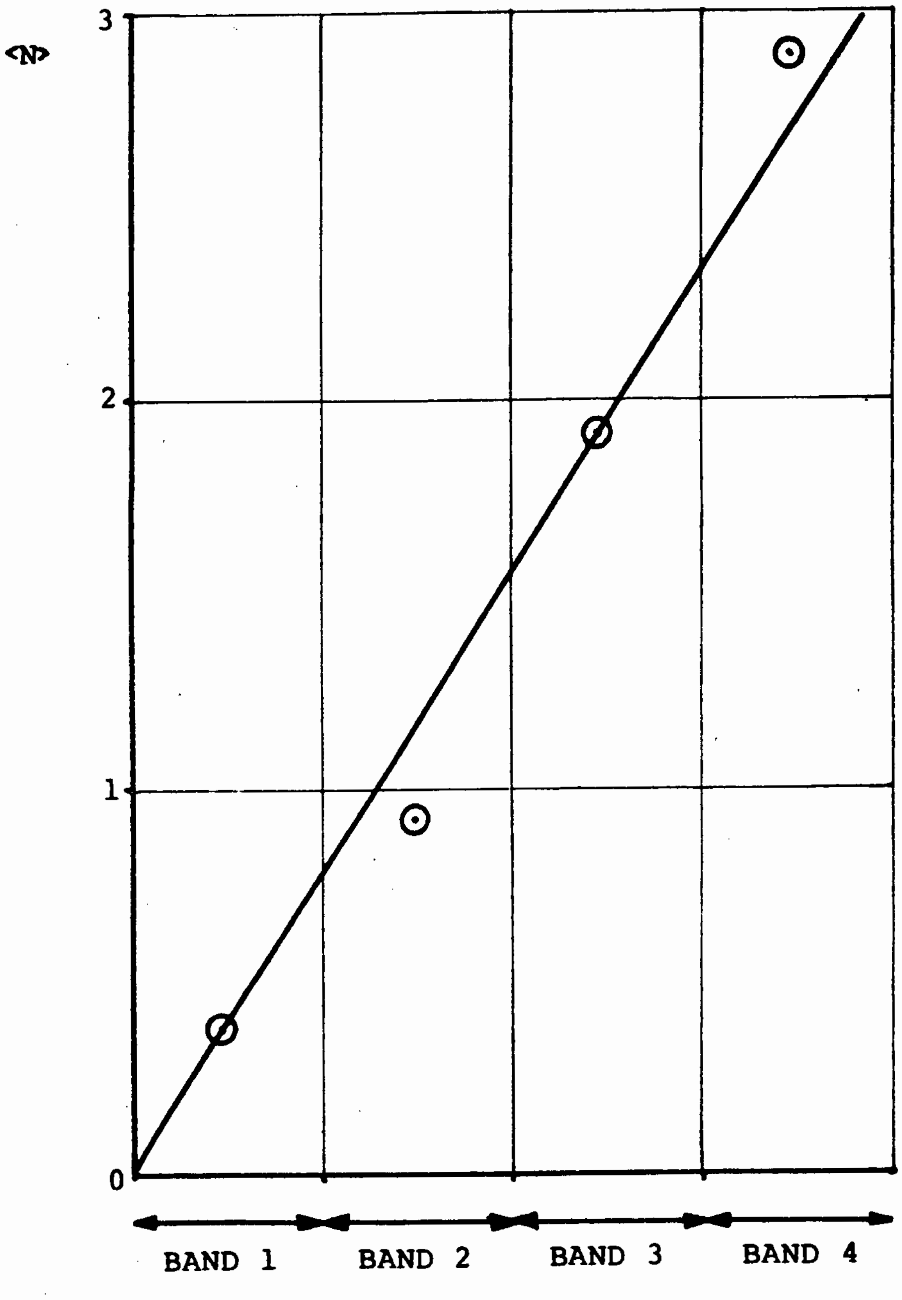

Fig.2 Estimates of <N> for different intensity regions of fish shoal