The Quantitative Analysis of Transonic Flows by Holographic Interferometry

Steve Carl Jamieson Parker

A thesis submitted for the degree of:
DOCTOR OF PHILOSOPHY

Submitted to:

THE DEPARTMENT OF ENGINEERING

at

THE UNIVERSITY OF WARWICK

in

October 1993.

British Aerospace,
Sowerby Research Centre,
Bristol.

NUMEROUS ORIGINALS IN COLOUR

For Vera and Cyril Parker

Summary

This thesis explores the feasibility of routine transonic flow analysis by holographic interferometry. Holography is potentially an important quantitative flow diagnostic, because whole-field data is acquired non-intrusively without the use of particle seeding.

Holographic recording geometries are assessed and an image plane specular illumination configuration is shown to reduce speckle noise and maximise the depth-of-field of the reconstructed images. Initially, a NACA 0012 aerofoil is wind tunnel tested to investigate the analysis of two-dimensional flows. A method is developed for extracting whole-field density data from the reconstructed interferograms. Fringe analysis errors are quantified using a combination of experimental and computer generated imagery. The results are compared quantitatively with a laminar boundary layer Navier-Stokes computational fluid dynamics (CFD) prediction. Agreement of the data is excellent, except in the separated wake where the experimental boundary layer has undergone turbulent transition.

A second wind tunnel test, on a cone-cylinder model, demonstrates the feasibility of recording multi-directional interferometric projections using holographic optical elements (HOE's). The prototype system is highly compact and combines the versatility of diffractive elements with the efficiency of refractive components. The processed interferograms are compared to an integrated Euler CFD prediction and it is shown that the experimental shock cone is elliptical due to flow confinement.

Tomographic reconstruction algorithms are reviewed for analysing density projections of a three-dimensional flow. Algebraic reconstruction methods are studied in greater detail, because they produce accurate results when the data is ill-posed. The performance of these algorithms is assessed using CFD input data and it is shown that a reconstruction accuracy of approximately 1% may be obtained when sixteen projections are recorded over a viewing angle of $\pm 58°$. The effect of noise on the data is also quantified and methods are suggested for visualising and reconstructing obstructed flow regions.

Contents

List of Figures

List of Tables

Nomenclature

Quantities are defined in order of appearance. Multiple definitions are only used when there is no risk of confusion.

Chapter 2

$\underline{v}_{flow}$ - flow velocity

$\underline{v}_{seed}$ - velocity of seeding particle

τ - particle response time

t - time

a_r - particle radius

ρ_p - particle density

ρ_f - fluid density

λ - wavelength in vacuum

ν - kinematic viscosity

∇ - gradient operator

$\underline{H}$ - magnetic field intensity

j - $\sqrt{-1}$

k_0 - wavenumber in vacuum

$\mathcal{H}$ - complex refractive index

$\underline{E}$ - electric field intensity

ϵ - electric permittivity

μ - magnetic permeability

σ - specific conductivity

ω - angular frequency

c - speed of light in vacuum

q - dimensionless constant

I_n - normalised intensity

$I_\perp$ - light polarised perpendicular to plane of observation

$I_{||}$ - light polarised parallel to plane of observation

σ_{RS} - Rayleigh scattering cross-section

n_m - refractive index of surrounding medium

n_s - refractive index of scattering particle

ΔE - energy liberated in electronic transition

Δt - lifetime of transition

h - Planck's constant

$\Delta\omega$ - angular line-width (approx.)

I_0 - peak or input intensity

$1/\Gamma$ - combined radiative life-times

ω_0 - mid-band angular frequency

$\Delta\omega_{1/2}$ - full-width-half-maximum angular line-width

Chapter 2 continued

K - Boltzmann's constant

T - absolute temperature

M - mass

ω_0' - Doppler shifted mid-band angular frequency

v_f - velocity of recording medium in a film drive

m - magnification of imaging system

v_{max} - maximum flow velocity

x, y, z - spatial co-ordinates

$\mathcal{F}$ - Fourier transform operator

δ - Dirac delta function

$\otimes$ - convolution operator

$i(x, y)$ - intensity distribution or transmittance function

J_1 - first order Bessel function

ψ_d - angle of diffraction

R - constant radius

θ - constant angle

ρ - radial spatial frequency co-ordinate

d - constant distance

N - integer of summation

ΔT - time between sequential exposures

f - frequency

Δf - Doppler shift

v - velocity component

α, β - constant angle

ω_1, ω_2 - Doppler shifted frequencies

i_s - signal current

i_n - noise current

η - quantum efficiency

I - constant intensity

$\Delta I / I_0$ - fractional attenuation of incident beam

μ_a - absorption coefficient

$p(x, y, z)$ - pressure field

L - constant length

dl - length integration element

ρ_0 - stagnation density

$\rho(x, y)$ - local static density

M - local Mach number

Chapter 2 continued

γ	-	ratio of specific heats	r_0	-	fixed position vector
n_i, n_t	-	refractive indices on either side of an interface	f	-	focal length of imaging lens/system
θ_i, θ_t	-	angles of incident and transmitted rays	w	-	aperture width
			Δl	-	optical path length change
$n(x,y,z)$	-	refractive index distribution	$\Phi(x,y)$	-	phase distribution
K_g	-	Gladstone-Dale constant	$r(x,y)$	-	complex amplitude of reference wave
$\rho(x,y,z)$	-	density field			
$\underline{r}$	-	position vector	$o(x,y)$	-	real amplitude of object wave
s	-	scalar			
$\underline{b}$	-	unit vector parallel to the ray direction	γ	-	degree of coherence
			$a(x,y)$	-	unmodulated background of a fringe pattern
$b(x,y)$	-	fringe contrast			

Chapter 3

$P(x_h, y_h)$	-	holographic recording plane	$t(x, y)$	-	transmittance function
O_D	-	diffraction amplitude	t_0	-	bias transmittance
$.^*$	-	denotes complex conjugation	β	-	gradient of photographic t/e curve
$f\#$	-	f-number of imaging system			
$\underline{\hat{k}}$	-	unit vector along a ray path	T	-	exposure time
d_f	-	fringe spacing	$\mathbf{O_i}$	-	complex amplitude of *ith* object wave
d_s	-	speckle diameter			
x_0, y_0, z_0	-	co-ordinates of a fixed position vector	$\mathbf{R}$	-	complex amplitude of reference wave
Δn	-	change in refractive index produced by compressible flow	w_x	-	width of hologram in the x direction
n_0	-	refractive index of static air	h_y	-	height of hologram in the y direction
x_d, y_d	-	co-ordinates in the plane of a diffuser	z_l	-	z co-ordinate of fringe localisation

Chapter 4

$\phi(x,y)$ - phase distribution

P_{jn} - noise equivalent power of Johnson noise

P_{sn} - noise equivalent power of shot noise

B - temporal bandwidth

R_0 - feedback resistance

R_l - load resistance

e - electronic charge

$\overline{P}$ - average optical power

m_f - fringe modulation

P_o - optical power of object beam

P_r - optical power of reference beam

SNR - signal-to-noise ratio

δ_p - constant phase step

β_f, α_f - Fourier coefficients

δ_x - phase step in x-direction

$i_k(x,y)$ - kth phase stepped interferogram

q_p - quantisation phase error

Q - number of quantised intensity levels

ϵ_1, ϵ_2 - linear and quadratic phase error coefficients

γ_n - error parameter in phase stepped fringe analysis

d_m - displacement of piezo mirror

f_x - carrier frequency in x-direction

$c(x,y)$ - complex valued interferogram with zero order removed

u, v - spatial frequency co-ordinates

u_x - location of first order

$\mathcal{R}$,$\mathcal{I}$ - real and imaginary components

Chapter 5

i_s	-	sampled intensity distribution	PTF	-	phase transfer function
a	-	spacing between pixels or sampled values	MTF	-	modulation transfer function
			i_{ip}	-	input intensity
f_b	-	positive cut-off frequency	i_{op}	-	output intensity
f_{-b}	-	negative cut-off frequency	i_{min}	-	minimum intensity
f_N	-	spatial Nyquist frequency	i_{max}	-	maximum intensity
δf	-	spatial frequency of moiré fringes	b	-	pixel width
			c	-	proportionality constant
$g(x)$	-	synthetic fringe profile	i_0	-	amplitude of sinusoidal intensity distribution
$n(x)$	-	gaussian noise distribution			
$h(x,y)$	-	point spread function	k_x	-	wave number at carrier frequency
α_w, β_w	-	parameters used to describe window function	ψ	-	constant phase difference
			$\Phi(ip)_i$	-	ith input phase value
$\otimes$	-	convolution operator	$\Phi(op)_i$	-	ith output phase value
$\mathcal{L}$	-	non-linear operator	o_1, o_2	-	object wave amplitudes
m,n	-	number of pixels in x and y directions	k_i	-	ith constant defining non-linear response
OTF	-	optical transfer function			

Chapter 6

Re	-	Reynolds number	λ_1,λ_2	-	recording and replay wavelengths
u	-	flow speed	dof	-	depth-of-field
l_c	-	chord length	d_r	-	spatial resolution
μ	-	dynamic viscosity	R	-	ideal gas constant
λ_a	-	geometrical factor	ϵ_ρ	-	percentage density error
t_a	-	aerofoil thickness	$\rho_{EXP}(M)$	-	experimental density field at Mach M
h_w	-	working section height	$\rho_{CFD}(M)$	-	CFD density field at Mach M
d_w	-	thickness of wire required to cause flow separation	u	-	object distance
θ_r	-	angle between reference beams			
$\overline{f}$	-	mean fringe spacing			
v	-	image distance			

Chapter 7

M_∞	-	freestream Mach number	n	-	no. of optical channels
p_0	-	stagnation (total) pressure	A	-	exposure area
$r_o(x,y,z)$	-	object source point	s	-	photographic sensitivity
$r_r(x,y,z)$	-	reference source point	ϵ	-	proportion of beam
$r_c(x,y,z)$	-	reconstruction source point			intercepted by HOE
R_o, R_r, R_c	-	curvature of spherical waves	d_e	-	real emulsion thickness
A_0	-	amplitude of wavefront	Λ	-	fringe spacing
S	-	spherical aberration	n	-	bulk refractive index
		coefficient	f	-	HOE focal length
C_x, C_y	-	comatic aberration	t_w	-	window thickness
		coefficients	ν	-	Poisson's ratio
A_x, A_y, A_{xy}	-	astigmatic aberration	$w(x,y)$	-	window deformation
		coefficients	p	-	pressure loading
E	-	optical energy	a, b	-	window dimensions
η	-	diffraction efficiency	E	-	Young's modulus
N	-	no. of superimposed			
		holograms			

Chapter 8

$f(x,y)$	-	perturbation of refractive index field	l_i	-	projection value for ith ray
p	-	projection co-ordinate	w_{ij}	-	weighting factor for ith ray passing through jth pixel
q	-	frequency domain projection co-ordinate	w_{max}	-	maximum weighting factor
θ	-	projection angle	f_j	-	value of reconstructed field at jth pixel
ds	-	line integration element	f_j^k	-	estimated field value at jth pixel
r	-	radial co-ordinate in axisymmetric flow	f_j^{k+1}	-	improved field value at jth pixel after kth+1 iteration
R	-	maximum extent of axisymmetric flow	l_{exp}	-	projection measurement for ith ray
$l(p,\theta)$	-	integrated projection	λ	-	relaxation factor
M_{proj}	-	number of projections	N	-	total number of values in the field
$l'(p,\theta)$	-	filtered projection	e_j	-	error in jth reconstructed pixel value
$h(p)$	-	impulse response of filter function	e_{rms}	-	rms reconstruction error
α	-	dummy variable	$\rho'(x,y)$	-	amended density field
δ	-	sampling interval	$\rho(x,y)$	-	CFD density field phantom
$b(x)$	-	basis function			
Δ_x	-	parameter used in B-sline basis function			
i	-	ray number			
j	-	pixel number			

Chapter 8 continued

ρ_{min}	-	minimum density value in the field	$l_e(p,\theta)$	-	projection of estimated field
$f_e(x,y)$	-	estimate for the unknown density field	$f_c(x,y)$	-	complementary field

Where appropriate, upper case characters denote the spatial Fourier transform of the corresponding lower case symbols.

Acknowledgements

I would like to thank the Sowerby Research Centre, British Aerospace, PLC and the Royal Commission of the 1851 Exhibition, for financially supporting this work programme.

From the Sowerby Research Centre I would like to thank Mrs Carren Holden, for her regular technical assistance on computing issues and for introducing me to aerodynamics. I would also like to thank: Dr Len Cooke for industrial supervision; Dr Philip Salter for his experimental advice on test; Mr Oswald Burgess for his constant technical and moral support; Dr Tim Slack for consultation on image processing; and members of the Aerodynamics Department for use of their computational fluid dynamic (CFD) predictions.

From Warwick University I would like to thank Dr Peter Bryanston-Cross for technical advice and academic support. Also from Warwick University, I would like to show my appreciation to Dr Tom Judge, for providing an elegant solution to the phase unwrapping problem. From Bolton Institute of Higher Education, I am grateful to Dr Landauro and Dr Jackson for their consultancy on algebraic reconstruction techniques.

I would also like to thank the staff of City University, London, and Cranfield Institute of Technology, for use of their wind tunnels and for providing technical support.

Finally, I would like to show my appreciation to Michelle Clark for her patience during the write-up period.

Declaration

The work presented in this thesis was undertaken as part of an applied research programme at the Sowerby Research Centre. All experimental work, including optical system design and wind tunnel testing, was performed by the author. Theoretical studies on optical techniques and flow diagnostics were also a result of the author's own research.

Mrs Carren Holden was a co-worker on computer algorithm development and provided a large software contribution on fringe analysis, computational fluid dynamics and tomographic imaging. Phase unwrapping software was supplied under contract by Dr Tom Judge of Warwick University. Bolton Institute of Higher Education performed a small prelimary study into the use of algebraic reconstruction techniques for reconstructing limited angle projection data; this contribution to the thesis is minor and is clearly referenced in chapter 8.

Chapters 5,6 and 7 of the thesis have formed a basis for the following papers, poster sessions and presentations.

- Invited presentation at the College of Aeronautics, Cranfield Institute of Technology, December 1990.
- Poster session at the *3rd International Conference on Holographic Systems, Components and Applications*, organised by the Institute of Electrical Engineers, Heriot-Watt University, Edinburgh 16-18 September 1991.
- Presentation at the *International Conference on Optical Methods and Data Processing in Heat and Fluid Flow*, organised by the Institute of Mechanical Engineers, City University, London 2-3 April 1992.
- Poster session at the *European Forum on Wind Tunnels and Wind Tunnel Test Techniques*, organised by the Royal Aeronautical Society, Southampton University, 14-17 September 1992.

- Presentation at the *International Conference on Applied Optics and Opto-Electronics*, organised by the Institute of Physics, Leeds University, 14-17 September 1992.

- C.M.E. Holden, S.C.J. Parker and P.J. Bryanston-Cross. *Quantitative 3D Holographic Interferometry for Flow Field Analysis.* Opt. and Lasers in Engineering Vol. 19, pp 285-298 (1993).

- S.C.J. Parker, C.M.E. Holden and P.J. Bryanston-Cross. *The Study of Transonic flows by Quantitative Holographic Interferometry.* J. Aerospace Eng. Proc. Part G. To be published in 1993.

In addition, holograms produced by the author, of transonic flow over a two-dimensional NACA 0012 aerofoil, are on display at the British Science Museum; the Engineering Department of Warwick University; the Sowerby Research Centre, British Aerospace; and Metrolaser, Los Angeles.

1 Introduction

At present, a purely theoretical analysis of most transonic flows is inadequate and computational predictions must be augmented by experimental measurements. A suitable diagnostic must be non-intrusive and whole-field, because transonic flows are perturbation sensitive and extend a significant distance from the aerodynamic body. Holographic interferometry fulfills these requirements and does not suffer from the practical restrictions imposed by particle seeding.

This thesis describes the development of a holographic diagnostic tool for the determination of three-dimensional density information from compressible flows. Initially, a simple two-dimensional flow was studied and then an optical system was designed and constructed, for recording multi-directional interferometric data. Fringe analysis and tomographic reconstruction techniques are presented for analysing the holographic measurements.

Holographic interferometry may be used to visualise the flow in a wind tunnel, by simultaneously reconstructing optical wavefronts which have separately traversed a working section containing static air and a compressible flow. High speed flows modify the density distribution of the gas and distort the transmitted wavefronts. Consequently, when the unperturbed and distorted wavefronts are reconstructed simultaneously, they interfere and produce a fringe pattern. The whole-field density data that can be extracted from interferograms provides diagnostic information and may be used in the validation of computational fluid dynamics (CFD) predictions. If the holograms are recorded by a pulsed laser then transient flow features can be visualised and wind tunnel vibration may be tolerated.

Chapter 2 presents a literature study and technical review of alternative optical diagnostics that utilise elastic and inelastic scattering processes. The relative merits of these techniques are examined and it is shown that they are largely unsuitable for use in high speed wind tunnel testing. In general, artificial particle seeding is required to obtain a strong signal from methods using elastic scattering and diagnostics based on

inelastic scattering operate most successfully when applied to reactive flows containing suitable molecular species. When wind tunnel testing at transonic speeds, the large mass flow rates often preclude particle and molecular seeding; and inconveniently short UV wavelengths must be used to study inelastic interactions.

Holographic interferometry does not utilise scattering processes, but works by visualising density changes in the fluid. Consequently, the technique offers the prospect of quantitative flow analysis, without the inconvenience, or loss of measurement accuracy, incurred by particle or molecular seeding. At low subsonic speeds, air density is uniform and therefore interferometry cannot be used. In supersonic flows, the density changes abruptly through shock waves and fringe densities are often too high for accurate data extraction. At intermediate transonic speeds, however, the density variations are extended in space and ideally suited to examination by holographic interferometry.

Although density is an important parameter for characterising compressible flows, a complete understanding of fluid behaviour around a body requires the measurement of velocity and an additional state variable. Holographic interferometry may be used to calculate the temperature field if simultaneous recordings are made at two different wavelengths. However, velocity is a vector quantity and is most easily measured using seeded flows. In this respect, holographic interferometry should be used in conjunction with at least one other method and treated as a complementary technique. An advantage of holographic interferometry over other diagnostics, is its ability to qualitatively visualise flows. In addition to providing quantitative information, holographic interferograms can be optically post-processed, to yield schlieren, shadowgraph and interferometric images from a single recording.

After a literature review and preliminary discussion of holographic principles, double-exposure, real-time, time-averaged and differential holographic interferometry are explained in the context of wind tunnel testing. A brief comparison is made of the lasers and recording materials that are suitable for implementing holographic interferometry in practical environments. The relative merits of holography over other

diagnostics and its practical limitations are discussed. An experimental study of specular and diffuse illumination recording geometries is then made to determine which configuration optimises fringe contrast and minimises noise. This preliminary investigation is important, because the accuracy of interferometric data is dependent on the quality of the reconstructed images.

The accuracy of a holographic flow diagnostic system is highly dependent on the methods of data acquisition and processing. Consequently, chapters 4 and 5 are dedicated to a study of electro-optic reconstruction systems and fringe analysis. Phase shifting and Fourier transform techniques are examined in detail, as these are the most practical and successful fringe analysis strategies. Phase measurement accuracy is quantified for the Fourier transform method, using a combination of computer generated synthetic data and wind tunnel imagery. Error analysis is particularly important when the interferometric phase measurements are reconstructed tomographically to provide three-dimensional flow information. If the holographic projections are processed inaccurately, the tomographic reconstruction algorithm may not converge to a stable solution.

Chapters 6 and 7, describe wind tunnel tests performed at Cranfield Institute of Technology and City University, London. During the first test, conducted in the Cranfield 9" by 7.5" transonic wind tunnel, the noise characteristics of specular and diffuse illumination holograms were compared. Dual reference beam holograms were subsequently recorded with the specular interferometer to determine the density field around a NACA 0012 two-dimensional aerofoil. Both narrow and wide angle dual reference beam configurations were implemented, to facilitate phase sensitive processing of the interferometric data. The narrow angle system, with approximately 1° between the two reference beams, was used to investigate data extraction by Fourier transform fringe analysis; whereas the wide angle system, with 28° between the beams, was used to study phase stepped fringe analysis.

Prior to running the Cranfield tunnel, the resolution and stability of the optical system were assessed. The wind tunnel tests were conducted at Mach 0.80 and 0.90

to avoid periodic flow and provide data which is relevant to the civil aircraft industry. Absolute and differential interferograms are presented, and the flow features are described qualitatively. The use of holographic interferometry as a quantitative diagnostic is then illustrated by processing the interferograms by both phase stepped and Fourier transform methods. The accuracy of these two strategies is considered, with an emphasis on using the techniques in an industrial environment. After fringe analysis, the phase distributions of the interferograms are converted into two-dimensional density fields by path-optimised phase unwrapping software. These results are compared quantitatively to a Navier-Stokes CFD prediction and the discrepancies are shown pictorially as a pseudo-colour difference map.

The wind tunnel test at City University investigated the potential of holographic optical elements (HOE's) for generating multi-directional interferometric data. An angular field-of-view of 116° was obtained by suspending the HOE's in front of the test section windows. The images are shown to have a resolution of 100μm, which exceeds that obtained using conventional schlieren optics. A cone-cylinder body was chosen to provide an intermediate test case between two-dimensional and fully three-dimensional flow. Experimental shock angles are compared to predictions from an Euler CFD code and cone theory, for freestream Mach numbers in the range 1.0 to 1.2. The holographic images show that the shock cone over the model is elliptical: a feature which would have been difficult to detect using other diagnostics.

Finite fringe interferograms are recorded of a cone-cylinder model at Mach 1.20 and viewing angles of 0° and ±58°. These results are processed by Fourier transform fringe analysis and compared with line integrated predictions from an Euler CFD code. During wind tunnel testing it was discovered that the windows bowed inwards, due to a pressure drop in the working section. The distortion of the glass introduced a set of parabolic error fringes, which appeared superimposed on the interferometric data. The surface distortion was determined by performing a series of calibration runs with an empty tunnel and subtracted from the phase maps. These phase maps, with the error surface removed by digital post-processing, are shown to be in good

agreement with computational predictions, although the asymmetry of the shock cone and viscous effects at the cone-cylinder junction, produce some discrepancies.

Chapter 8 describes the implementation of computed aided tomography (CAT) for the reconstruction of limited angle multi-directional interferograms. Initially, the central slice theorem is used to show that, in principle, a reconstruction accuracy exceeding 5% may be obtained with a viewing angle of only $\pm 45°$. The performance of ART algorithms is then investigated for geometries where the field-of-view is restricted, or the number of projections is small. A plane of density values, computed from an Euler CFD code modelling flow around a cone-cylinder body, is used to generate the input projection data. The accuracy of the ART algorithms is evaluated by differencing the reconstructed fields from the original phantom. Also, reconstruction errors are quantified when the views are corrupted by computer generated noise, to simulate the use of holographic imagery instead of 'perfect' data. Finally, Fourier domain low-pass filtering is investigated for suppressing aliasing artefacts.

To conclude this thesis, two optical configurations are presented which could be used to record multi-directional interferometric data in a production wind tunnel. The first design is based on the spatially multiplexed system used at City University. An attraction of this geometry is that all of the views are recorded simultaneously. A principal disadvantage, however, is that the flow must be located in the area of beam intersection. Consequently, the field-of-view of each optical channel must be large, or the flow region must be highly confined.

An alternative system is proposed, whereby the views are acquired sequentially by rotating the model between successive holographic recordings. A disadvantage of this approach is that data capture is prolonged and flow unsteadiness may generate inconsistencies between views. To overcome these limitations, a real-time recording geometry is described that can acquire interferometric data at TV frame-rates, without intermediate photographic development of the holograms. The advantages of this design are that the optical configuration is simplified and each of the views has a common optical path.

2 Review of Optical Flow Diagnostics

2.1 The Interaction of Photons with Matter

2.1.1 Introduction

Experimental aerodynamics involves the measurement of flow temperature, pressure, density and velocity. Conventional diagnostics normally require the insertion of a probe into the flow. Instruments in common use include surface mounted transducers for measuring static pressure; the pitot-static tube for calculating Mach number from simultaneous measurements of stagnation and static pressure; and the hot wire anemometer for determining Mach number from changes in electrical resistivity.

The accuracy of flow diagnostics is dependent on both Mach number and Reynolds number. For example, a review by Steinback [1] concludes that hot wire anemometry is unreliable at compressible speeds, because the heat transfer from the wire is a complicated function of velocity and state variables. A more fundamental limitation of intrusive diagnostics at transonic and supersonic speeds is that probes perturb the flow, have a finite response time, and are subject to large mechanical forces. In reactive high enthalpy flows there is also the risk of destroying the measuring device.

Probes record data at only a few discrete points and therefore they provide an incomplete picture of the flow. If the sensor is scanned to sample the entire field then the flow must be in a steady state to provide spatially consistent data. It is difficult to study unsteady processes intrusively, because turbulence and mixing produce features which vary rapidly in both space and time. In practice, unsteady flow measurements can be temporally averaged, although this inevitably results in a loss of physical understanding. The deficiencies of time-averaged measurements are demonstrated by Bearman *et al* in a study of three-dimensional wake flows using particle image velocimetry (PIV) [2]. In their investigation they show that the time-averaged vortex structure behind motor vehicles is markedly different from the instantaneous flow. At perturbation sensitive transonic speeds, these differences are likely to be just as

important.

In addition to physical limitations, point measurement techniques have economic penalties associated with the lengthy data acquisition times, and elaborate instrumentation of the model and wind tunnel. Nevertheless, intrusive methods are reliable and provide data which requires negligible post-processing. Also, these devices are simple to interface to analogue and digital electronics, thereby enabling rapid data storage and analysis. Consequently, novel flow diagnostics must be highly automated and have proven accuracy if they are to gain acceptance alongside established intrusive technologies.

An indirect qualitative understanding of complex unsteady flows may be gained by visualising the flow optically. These preliminary investigations can be subsequently augmented by conventional intrusive measurements to provide quantitative aerodynamic data in critical regions of the flow. The expense and complexity of using both optical and intrusive diagnostics is sometimes prohibitive. Recently, however, advances in laser and computer technology suggest that optical diagnostics may be used for both the initial qualitative investigation and the subsequent quantitative study.

Optical diagnostics offer significant advantages over conventional intrusive techniques. A principal benefit is that the photons in a beam of light impart negligible momentum to the fluid through which they propagate and therefore optical interrogation may be regarded as wholly non-intrusive. In general, however, optical measurements are difficult to relate to fluid properties and the abundance of data generated by these techniques cannot be accessed quickly due to bottlenecks in data acquisition, processing and storage. These disadvantages have been relaxed by recent advances in charge coupled devices (CCD's), high power pulsed lasers, high speed computing and digital image processing. This has led to the refinement of a host of complementary optical techniques, which may be broadly classified into two main categories:

1. Methods which measure the properties of the fluid directly.

2. Methods which examine the dynamics of entrained foreign matter.

The first of these categories is primarily concerned with the calculation of fluid density from measurements made on the associated refractive index field. Physical processes which alter the speed of light through a fluid, and therefore modify the refractive index, include compressible flow, combustion, convective heat transfer, fluid mixing and stratification. Schlieren, shadowgraph and interferometry are the principal techniques for visualising these changes in the refractive index field. Historically, these methods were used qualitatively, with interferometry used occasionally in quantitative studies for flows with a simple symmetry (two-dimensional or axisymmetric) [3].

Recent progress in image processing hardware and software suggests that the automatic analysis of interferometric data will become increasingly common. An inherent problem, however, when measuring an extended density field from modifications of the refractive index field, is that the results are integrated along the line of observation. Inversion of the data is complicated, except for idealised two-dimensional and axisymmetric density fields, and tomographic reconstruction techniques are needed to analyse three-dimensional flows. In practice, optical access is usually restricted and therefore the projection data is unevenly sampled, resulting in aliasing errors in the experimental density data.

An alternative to measuring variations in the refractive index field, is to examine the intensity and spectral properties of light scattered from a localised point. The scattered signal may be produced by an elastic or inelastic interaction with molecules in the flow; or it may be enhanced by scattering light from artificially seeded, or naturally occurring, microscopic particles. When a molecular interaction is used, the scattering process is inefficient and the input light beam must be tightly focussed to localise the measurement and provide sufficient optical power density. The illumination must be scanned to cover the entire field and the flow information is derived spectroscopically. Although data capture and spectral analysis is complicated, improvements in pulsed laser technology at short wavelengths (frequency doubled Nd:YAG and UV excimer lasers), have led to an increase in the popularity of these techniques [4, 5].

Light interacts weakly with molecules in a flow and therefore image intensifiers and

photon counters must be used to detect the scatter. The signal may be enhanced by particle seeding, so that conventional CCD cameras may be used and data recording is not confined to a single point. Natural particle concentrations in a gas are usually insufficient and therefore the flow must be seeded with foreign matter to increase the scattered signal. Ahlborne [6] and Prandtl [7] are accredited with the first use of particle seeding for the measurement of flow velocity from the streaked images recorded on a time-averaged photograph. In these early experiments, the data was extracted manually and therefore the method was impractical for routine use. Particle tracking is now experiencing renewed interest, due to the availability of inexpensive CCD cameras, digital framestores and high speed computers.

Seeding particles must follow the flow precisely if the velocity field is to be calculated accurately. The trajectory of a particle in a flow field is governed by its size, shape and density with respect to the fluid. In practice, this necessitates the use of small, neutrally buoyant particles, and Melling [8] has shown that at supersonic speeds, the particle diameter should not exceed 0.3μm for regions experiencing high acceleration (e.g. across shocks). An additional condition, which often precludes the use of many substances, is that the scattered signal must be strong enough to be detected amongst noise. In summary, a suitable material for flow seeding must satisfy the following criteria:

1. The particle and fluid densities must be similar.
2. The maximum permissible particle mass is limited by inertia.
3. The minimum permissible particle mass is fixed by Brownian motion.
4. Particle seeding should be monodisperse (single sized) to avoid ambiguities in the measurements.
5. The seeding density should be optimised to provide adequate data sampling while preventing flow disruption.
6. The scattered signal should be large.

7. The seeding agent should be convenient to use (low toxicity, stable, easily deployed, inexpensive and non-corrosive).

The velocity of an entrained seeding particle may be related to the flow velocity by a first order approximation:

$$\underline{v}_{flow} = \underline{v}_{seed} + \tau \frac{d\underline{v}_{seed}}{dt} \tag{1}$$

Where,

$\underline{v}_{flow}$ - flow velocity

$\underline{v}_{seed}$ - velocity of the seeding particle

τ - particle response time

and underlined characters denote vector quantities.

Merzkirch [9] has shown that if the particle and fluid densities are similar, the response time to a discontinuous velocity change is:

$$\tau = \frac{2a_r^2}{9\nu} \tag{2}$$

Where,

a_r - particle radius

ν - kinematic viscosity (1.461×10^{-5} m^2 s^{-1} at STP)

Equation 1 illustrates that errors are most acute in regions of high acceleration because the particles lag behind the flow. Elder [10] has shown that across a shock wave, a velocity change of $300ms^{-1}$ occurs in typically $2\mu s$ for a $0.5\mu m$ particle. The velocity error for neutrally buoyant particles is therefore approximately $0.15ms^{-1}$, which is insignificant at transonic speeds. However, errors become more acute as the particle size increases, and lift forces may occur in regions with a strong velocity gradient. Further inaccuracies arise when there is a mismatch in the density of the particle ρ_p and the fluid ρ_f. Agui and Jimenez [11] have shown that for small density differences, the error increases in proportion to the local Mach number and $(\rho_p/\rho_f -$

1). This brief discussion illustrates that care should be exercised when interpreting results from seeded flows and the velocity mismatch between a fluid and the entrained particles should be accurately quantified.

2.1.2 Elastic Scattering

Photons interact with matter by either elastic or inelastic scattering. Elastic scattering occurs when there is no exchange of energy between the incident photon and the internal energy levels of the atoms. Elastic interactions are further subdivided into Rayleigh, Tyndall or Mie scattering, according to the size of the particles in relation to the illumination wavelength λ. The categories are loosely defined as follows:

Rayleigh scattering: $a_r \ll \lambda$
Tyndall scattering: $a_r \leq \lambda$
Mie scattering: $a_r > \lambda$

To study the scattering process theoretically, it is necessary to solve Maxwell's equations for the diffraction of a plane linearly polarised monochromatic wavefront, by a non-magnetic sphere embedded in a non-conducting homogeneous medium. The time independent Maxwell's equations for the electric and magnetic fields are:

$$\nabla \wedge \underline{H} = -jk_0\mathcal{H}^2\underline{E} \tag{3}$$

$$\nabla \wedge \underline{E} = jk_0\underline{H} \tag{4}$$

Where,

$\underline{H}$ - magnetic field intensity
$\underline{E}$ - electric field intensity
j - $\sqrt{-1}$
k_0 - wavenumber in vacuum (propagation constant)
$\mathcal{H}$ - complex refractive index

The wavenumber and complex refractive index are related to the properties of the material and the angular frequency ω of the electromagnetic field by:

$$\mathcal{H} = \epsilon\mu + j\frac{4\pi\sigma\mu}{\omega} \tag{5}$$

$$k_0 = \frac{\omega}{c} \tag{6}$$

Where,

- μ - magnetic permeability
- σ - specific conductivity
- ϵ - permittivity (dielectric constant)
- c - speed of light in a vacuum

The solution of Maxwell's equations for a sphere in air was first given by Mie in 1908 [12]. This result is summarised in table 1 as a tabulation of normalised intensity values I_0, versus angle, for different scattering regimes [13]. The dimensionless parameter q, defined by equation 7, quantifies the similarity between the particle radius and the laser wavelength. The scattered light is polarised and therefore the intensity values I_n are a normalised summation of these components (equation 8).

$$q = \frac{2\pi a_r}{\lambda} \tag{7}$$

$$I_n = q^2(I_\perp + I_{||}) \tag{8}$$

Where,

- $I_\perp$ - light polarised perpendicular to the plane of observation
- $I_{||}$ - light polarised parallel to the plane of observation

From table 1 it is seen that when the particle diameter exceeds the illumination wavelength and $q > \pi$ (Mie scattering), the scattering cross-section is large and predominantly in the forward direction ($\theta = 0°$). Conversely, in the Rayleigh regime with $q = 0.1$, the cross-section is a factor of $10^8 - 10^{11}$ smaller and the light is scattered almost uniformly.

Mie scatter is produced from large particles and therefore this interaction is used to investigate 'dirty flows'; for example in the study of combustion processes that liberate large quantities of soot. In general, these particles are too massive to follow the flow accurately and therefore Mie scatter is used primarily to examine combustion products rather than the flow that entrains them. In contrast, Rayleigh scatter occurs when the wavelength of light exceeds the particle diameter and therefore the flow may be tracked accurately. The Rayleigh scattering cross-section σ_{RS}, for spherical particles of refractive index n_s embedded in a medium of index n_m, is derived by Morel [14] as:

$$\sigma_{RS} = \frac{16\pi}{3}\left(\frac{2\pi}{\lambda}\right)^4 \frac{n_m^4(n_s^2 - n_m^2)^2}{(n_s^2 + 2n_m^2)^2} a_r^6 \tag{9}$$

Each molecular species has a constant scattering cross-section, and therefore the total scatter is proportional to the number density of the gas. Unfortunately, the Rayleigh scattering cross-section is small and inversely proportional to the fourth power of the illumination wavelength. Consequently, Rayleigh scatter can only be used as a flow diagnostic if a high power narrow band source can be found, which operates at short visible wavelengths or in the UV. Recently, this technique has become increasingly popular due to the development of frequency doubled Nd:YAG and UV excimer lasers. For instance, Barat *et al* used a pulsed Nd:YAG laser to obtain instantaneous flame temperatures in the presence of intense background black body radiation and laser induced glare [15].

One of the main attractions of using Rayleigh scatter as a flow diagnostic is that, in principle, it is possible to determine velocity and all of the state variables [16]. The

θ	$q = 0.01$	$q = 0.1$	$q = 0.5$	$q = 1$	$q = 2$	$q = 5$
0	5.0×10^{-14}	5.0×10^{-8}	1.2×10^{-3}	2.3×10^{-1}	4.3	9.8×10^{2}
90°	2.5×10^{-14}	2.5×10^{-8}	5.0×10^{-4}	3.6×10^{-2}	2.5×10^{-1}	2.7
180°	5.0×10^{-14}	4.9×10^{-8}	7.8×10^{-4}	1.9×10^{-3}	2.0×10^{-2}	1.3

Table 1: Normalised scattered intensity versus angle θ and q, for dielectric spheres of refractive index 1.25

scattering cross-section is proportional to the number density of the gas and therefore the intensity of the scattered signal may be used to calculate density. A limitation of this method is that the measurements are intensity based and therefore reduced in accuracy by absorption and extraneous scatter. The temperature and velocity of a flow may also be determined from Rayleigh scatter, by measuring the principal wavelength and spectral profile. Errors are introduced into these measurements because the linewidth of a laser beam is broadened by various intrinsic and extrinsic processes. Intrinsic natural line broadening, also known as radiative broadening, is governed by Heisenberg's uncertainty principle. A consequence of this quantum mechanical law is that it is impossible to calculate precisely, both the lifetime Δt and the energy ΔE liberated during an electronic transition. Mathematically this may be expressed by an inequality incorporating Planck's constant h:

$$\Delta E \Delta t \geq \frac{h}{4\pi} \tag{10}$$

Hence, as the uncertainty in the photon energy ΔE is equal to $h\Delta\omega/2\pi$, the line-width $\Delta\omega$ is given approximately by:

$$\Delta\omega \sim \frac{1}{2\Delta t} \tag{11}$$

The precise profile of the broadened line is Lorentzian and may be determined by quantum mechanics, as shown by Weisskopf and Wigner in Corney's comprehensive review of spectroscopy [17]:

$$I(\omega) = I_0 \frac{\Gamma/2\pi}{(\omega - \omega_0)^2 + \Gamma^2/4} \tag{12}$$

Where,

I_0 - peak intensity of the profile

$1/\Gamma$ - sum of the radiative life-times of the initial and final electronic states

ω_0 - central frequency of the profile

Natural line broadening is usually negligible when compared to other broadening mechanisms, but it imposes a theoretical resolution limit on the spectroscopic measurement of flow properties.

Pressure or collisional broadening occurs if the laser beam propagates through a region of high gas pressure because of the strong interaction of neighbouring atoms. This process also results in a Lorentzian profile, which is slightly broader than that produced by natural line broadening. Throughout the visible and UV region of the spectrum, the dominant broadening mechanism in flows is due to random thermal motion. This results in a Doppler shift of the radiation and the broadened line has a Gaussian profile which is typically two orders of magnitude wider than the natural and pressure broadened line width. The Doppler broadened profile has a full-width-half-maximum (FWHM) of:

$$\Delta\omega_{1/2} = \frac{2\omega_0}{c}\sqrt{\left(\frac{2KT}{M}\ln 2\right)} \tag{13}$$

Where,

$\Delta\omega_{1/2}$ - FWHM signal bandwidth
K - Boltzmann's constant
T - absolute temperature of the gas
M - mass of the radiating particle

Hence, if an air flow is interrogated by a red laser with an angular frequency of 3×10^{15}Hz, the Doppler broadened linewidth is approximately 10^{10}Hz at 298K, which is two orders of magnitude larger than a typical natural linewidth. Equation 13 may be used to calculate flow temperature from the width of the broadened spectrum if the mass of the particular molecular species is known and the line width of the interrogating beam is negligible. In addition to line broadening by random molecular motion, the scattered light is also Doppler shifted by the bulk movement of the gas. The central frequency ω_0' of the broadened spectrum may be used to determine flow velocity v by the Doppler relation:

$$|v| = c(1 - \omega_0'/\omega) \tag{14}$$

In conclusion, Rayleigh scattering is a simple flow diagnostic to apply because detailed spectroscopic analysis is unnecessary. Unfortunately this inherent simplicity also means that the separate species concentrations cannot be determined. Another drawback is that Rayleigh scatter is of low intensity and therefore image intensifiers are required unless high power UV lasers are used. The determination of flow velocity and temperature is restricted by the measurement accuracy of the Doppler broadened and shifted spectra. In spite of these limitations, modern UV excimer lasers and photodetector arrays have been used to obtain high quality Rayleigh scatter images of supersonic boundary layers and regions of shock/boundary layer interaction [18]. It is likely that this technology will attract growing interest as excimer laser and CCD technologies mature.

2.1.3 Inelastic Scattering

In addition to elastic scattering, photons may interact inelastically with gas molecules and these processes provide an additional range of flow diagnostics. Inelastic scattering occurs when energy is transferred to the quantised energy states of the molecules and the interactions are categorised as: absorption, fluorescence and Raman shifting. Raman shifting is further subdivided according to which energy states are involved, whether the interaction is linear, and by the relative coherence between the incident and emitted photons.

The following list summarises the specific inelastic scattering modes that occur and figure 1 illustrates which electronic transitions are involved.

1. **Absorption**

   The incident photon is absorbed, promoting the molecule into a higher electronic, rotational or vibrational energy state (or a combination of these). The

excited molecule is then subject to spontaneous emission of a photon within a typical life-time of 10^{-8}s. Fluorescence occurs when the state is longer lived (sometimes hours).

2. **Stimulated Emission**

   Unlike spontaneous emission this interaction is triggered by an incident photon. The emergent photon is in-phase with the incident photon and of identical frequency. Consequently, the resulting beam is coherent and may be amplified in a resonant cavity to produce a highly intense, collimated beam. This process forms the basis of laser operation.

3. **Raman-Stokes Scattering**

   With Raman-Stokes scattering part of the incident photon energy excites the molecule, with the remainder radiated as a photon of lower energy and longer wavelength (incoherent).

4. **Raman-Anti-Stokes Scattering**

   With Raman-Anti-Stokes scattering a pre-excited molecule is reduced to a lower energy state by the incident photon. The process is incoherent, unlike spontaneous emission, because the emitted photon is of shorter wavelength to the incident photon and uncorrelated in phase.

5. **Stimulated Raman Scattering**

   Stimulated Raman scattering forces an excited molecule to release a photon of identical frequency to the incident photon. Unlike spontaneous emission, this can only occur while other Raman scattering processes are present.

In summary, numerous flow diagnostics have been devised that exploit both elastic and inelastic scattering. This thesis is primarily concerned with holographic interferometry, which measures the retardation of unscattered wavefronts by a flow induced refractive index field. A major advantage of this technique over methods based on scattering, is that particle seeding is not required. A drawback of not using scattering

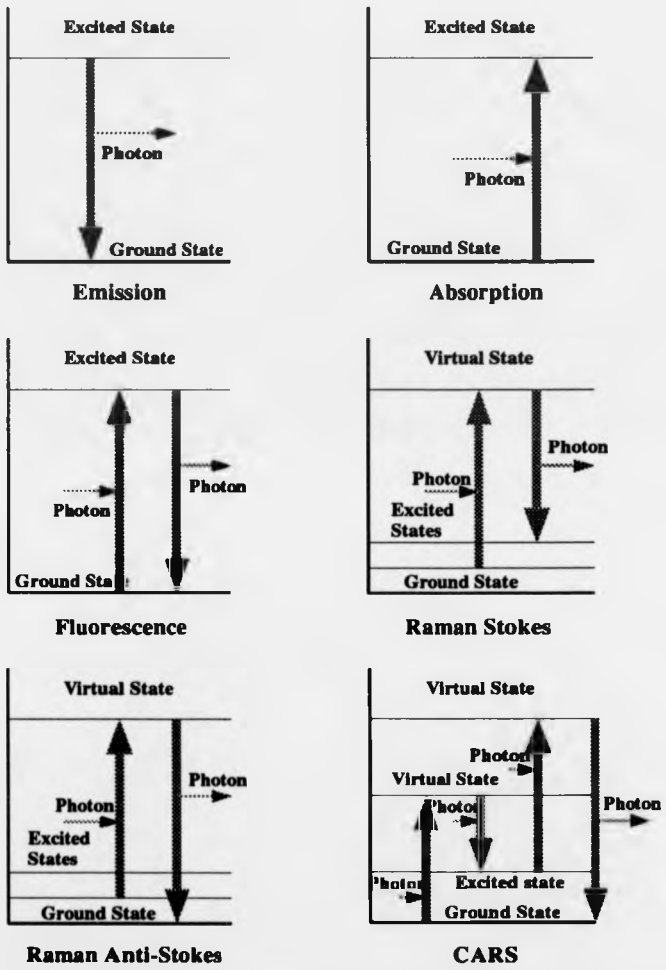


Figure 1: Energy level diagrams for inelastic scattering processes

is that the measurements are delocalised by integration through the flow. The remainder of this chapter presents a literature study of the principal diagnostics that exploit scattering processes and the techniques are compared objectively to holographic interferometry.

2.2 Flow Diagnostics which utilise Elastic Scattering

2.2.1 Particle Image Velocimetry (PIV)

Particle image velocimetry (PIV) originated as a natural extension to using particle seeding for qualitatively visualising flow fields. The technique was pioneered by Dudderar and Simpkins in 1977 [19] and has subsequently been used to study phenomena as diverse as water waves [20] and convective flows [21]. PIV has been applied to mostly low speed flows, but recently Bryanston-Cross *et al* have extended the working range to transonic Mach numbers [22, 23]. At these speeds, the performance of PIV is dominated by how the flow is seeded.

Various mechanisms may be used to introduce the seeding agent into a flow, and the techniques are broadly classified as: atomisation (e.g. by a compressed air nebuliser [24]), vaporisation, condensation [25], sublimisation, chemical activity [26], extraction from a fluid bed [27] and combustion. Numerous seeding agents have been used, and the most popular substances are:

- styrene
- pyrotechnic or tobacco smoke
- silicon oil
- latex particles
- paraffin, oil, kerosene, asphalt or grease
- incense
- helium bubbles
- silicon dioxide
- talcum
- titanium tetrachloride (plus moisture)
- Di-octyl phthalate
- ammonium chloride

- Drager pellets
- wax

A seeded flow may be visualised and recorded photographically by illuminating the entrained particles with a one-dimensional light sheet. Light is scattered elastically from the particles, with the cross-section related to the wavelength by an inverse fourth power law in the Rayleigh limit (equation 9). Consequently, a frequency doubled pulsed Nd:YAG laser is often used as the illumination source, because it produces high peak powers at short visible wavelengths (532nm). System alignment is also simplified using this laser, because its repetition frequency (>50Hz) is high enough for the beam to appear continuous.

Flow velocity can only be determined accurately if the particles follow the stream lines faithfully, respond quickly to forces within the fluid, and do not disturb its properties. At compressible speeds this necessitates the use of microscopic seeding particles ($< 1\mu$m) so that inertial and gravitational effects are minimised and strong flow accelerations are followed (equation 1). Unfortunately, the intensity of Rayleigh scatter is proportional to the sixth power of the particle size (equation 9). Hence, a compromise must be made between particles that are easily visualised and those which map the flow field accurately.

Fluid velocity is determined using PIV by recording a double-exposure photograph of a seeded flow, with an appropriate time delay between the two exposures. Figure 2 shows a PIV image recorded of transonic flow around the trailing edge of a turbine blade [28]. A pulse separation of 0.5μs was used and the heavily exposed areas at the left and bottom of the picture are due to extraneous scatter. Four regions of the flow, depicted in the square boxes of figure 2, were chosen for subsequent image processing.

The velocity components in the plane of the picture may be measured by one of two methods. The direct approach is to image each of the particle pairs and calculate the displacement vector. This method provides a high level of confidence in the measurements, although a knowledge based system must be used to prevent

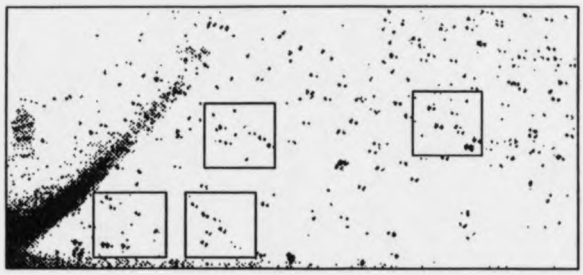


Figure 2: PIV image of transonic flow around the trailing edge of a turbine blade [28]

incorrect pairing of particles. Large errors may also arise if the difference between the respective position vectors cannot be determined accurately, because the image resolution is low, or the particles are closely spaced. The major obstacles to obtaining quantitative measurements from PIV may therefore be summarised as follows:

- Difficulty in photographing micron sized particles due to the limited speed and resolution of the recording material.
- Difficulty in resolving micron sized particles from the photographic record.
- Difficulties in obtaining sufficient seeding density and even flow entrainment.
- Pair identification.
- Limited accuracy in determining the displacement vector.
- Sign ambiguity of the velocity vector.
- Movement of particles out of the light sheet.
- Determination of the z-component of velocity.

Many of these problems have been overcome by using high resolution films in conjunction with fast developers. Bryanston-Cross *et al* have obtained accurate particle images at large stand-off distances, by using diffraction limited optics with the seeding agent dispersed by nebulisers [28]. The directional ambiguity of the velocity vector

may be determined from a *priori* knowledge of the flow physics. Alternatively, the film may be moved between exposures to add a carrier velocity to the particle pairs [29], or the illumination may be scanned by use of birefringent optics [30]. Lauterborn *et al* used a drum camera to transport the film [31] and they show that the directional ambiguity may be removed if the following condition is satisfied:

$$v_f > m v_{max} \tag{15}$$

Where,

v_f - speed of the film
m - magnification of the recording system
v_{max} - maximum flow velocity opposing the film movement

Pair identification has been aided by increasing the number of exposures [21] and by encoding the separate images at two different wavelengths. Alternatively, the photographic film may be replaced by a high speed CCD camera to enable real-time tracking and detection of the particles [32, 33]. Separate encoding of the particles within each pair also enables the velocity direction to be determined without the need for carrier techniques. Strategies for determining the out-of-plane velocity component have involved the use of thicker light sheets, holographic recording in conjunction with dual reference beam methods for separate particle labelling [34, 35, 36], and stereo photography [37]. Bryanston-Cross *et al* have shown that an additional advantage of using holographic recording when a short time separation is used, is that Young's fringes appear superimposed on the reconstructed defocussed image of the particle pair [38]. These fringes may then be used to directly calculate the velocity vector at that point in the flow.

An indirect method for measuring the velocity field recorded by the particle pairs is to use the diffractive properties of microscopic point images to produce macroscopic interference fringes. If a photographic negative of a particle pair is interrogated by a plane wave from an unexpanded laser beam, the Fraunhöfer diffraction pattern $I(u, v)$ is given by the power spectrum of the aperture function:

$$I(u,v) = |\mathcal{F}[\{\delta(x-x_1, y-y_1) + \delta(x-x_2, y-y_2)\} \otimes i(x,y)]|^2 \qquad (16)$$

Where,

$\mathcal{F}$ -	Fourier transform operator
$(x_1, y_2); (x_2, y_2)$ -	co-ordinates of the particle pair
δ -	Dirac delta function
$\otimes$ -	convolution operator
$i(x,y)$ -	transmittance function of a single particle

To a first approximation the particle image is a circular hole of radius R. Therefore the diffraction pattern from a single particle image is the classical Airy pattern:

$$I(kR\sin\psi_d) = I_0 \left[\frac{2J_1(kR\sin\psi_d)}{kR\sin\psi_d}\right]^2 \qquad (17)$$

Where,

J_1 -	first order Bessel function
k -	propagation number $(2\pi/\lambda)$
ψ_d -	angle through which the incident beam is diffracted
I_0 -	peak intensity of the diffraction pattern

Hence, the diffraction pattern from the pair of apertures is given by the convolution theorem as:

$$I(\rho) = I_0' \cos^2 \left\{\frac{\pi}{\lambda}[(x_2 - x_1)\rho\cos\theta + (y_2 - y_1)\rho\sin\theta]\right\} J_1^2\left(\frac{\rho kR}{d}\right) \qquad (18)$$

Where,

ρ, θ -	polar co-ordinates in the plane of the diffraction pattern
I_0' -	peak intensity of the diffraction pattern
d -	distance to the screen from the apertures

Equation 18 describes a sinusoidal fringe pattern, which is analogous to that produced by Young's original double-slit experiment. Flow velocity may be determined by measuring the mean spacing and orientation of these fringes. The method is not

particularly accurate, however, because the fringes are modulated by the Airy diffraction pattern generated by the individual particle images. When the photographic negative is interrogated, it must be ensured that the beam width exceeds the mean particle pair spacing so that fringes are produced. If more than one particle pair falls within the interrogation area, then cross interference occurs between particles from different pairs. This process generates a noise 'halo' in the diffraction pattern and complicates the determination of fringe spacing and orientation. If the flow velocity is approximately constant within the interrogation area, then the particle pairs lie predominantly in the same direction and have similar spacings. In this instance, the Young's fringes can be analysed accurately and the measurement errors can be used to quantify velocity deviations in the area.

An alternative to optical interrogation of the particle images is to generate the Fourier transform digitally using a fast Fourier transform (FFT) algorithm [39]. The boxed regions of figure 2 were used to form four 256×256 binary raster images (figure 3), which were subsequently Fourier transformed to generate Young's fringe patterns (figure 4). The 2D FFT performs approximately $2N^2\log_2 N$ multiplications for N data points, opposed to N^4 multiplications required when computational redundancy is not invoked. Although the FFT is an efficient algorithm, transforming images of this size digitally has historically required many minutes CPU time. In addition, sampling errors are introduced during digitisation, spatial resolution is lost, and the dynamic range is compressed. Consequently, analogue optical processing methods have been favoured, although this trend is currently reversing due to sustained improvements in computer hardware. At present, a 256×256 FFT requires 1s CPU time on a SPARC 10 workstation and real-time processing is possible if dedicated hardware is used.

Although the velocity vector may be determined directly from Young's fringes, a more accurate method, which is conducive to automatic processing, involves forming the autocorrelation function of the particle pair images. In practice this is achieved by computing the inverse Fourier transform of the fringe pattern. Figure 5 shows the autocorrelation functions of the four images extracted from figure 2.

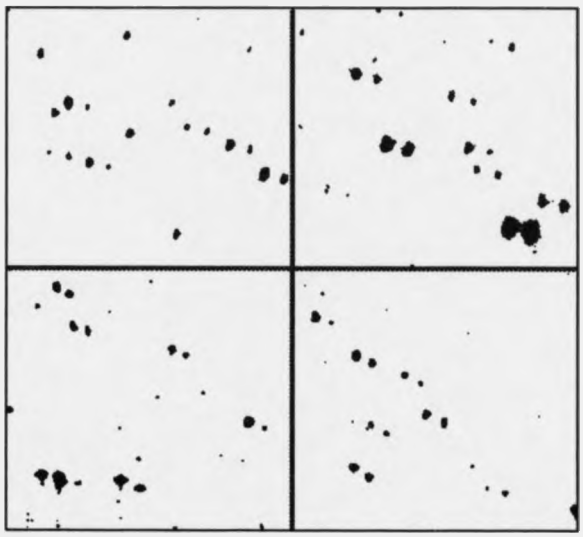


Figure 3: Flow regions selected for processing by Fourier analysis

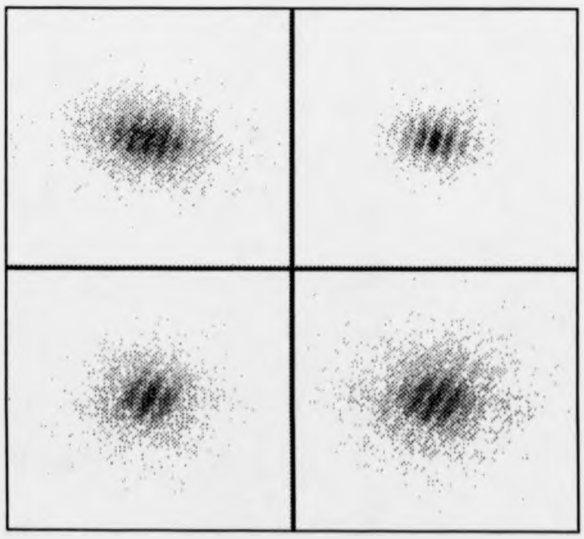


Figure 4: Computer generated Young's fringe patterns

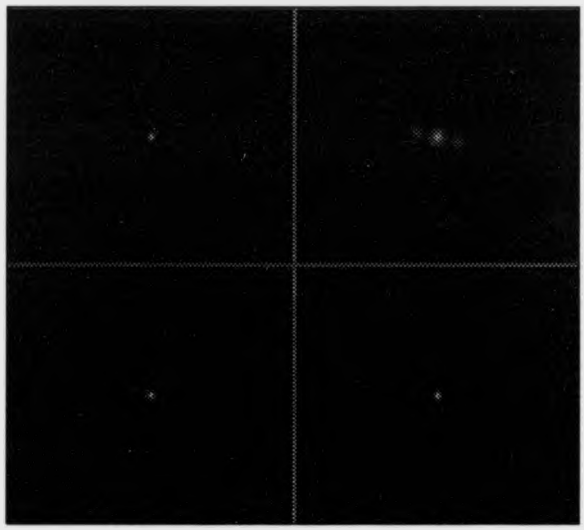


Figure 5: Autocorrelation functions of the particle pairs

The autocorrelations consist of an on-axis zero order, with symmetrically disposed positive and negative first orders. The velocity vector lies parallel to the vector joining the first orders and has a magnitude $|v|$ given by:

$$|v| = \frac{d}{2\Delta t m} \tag{19}$$

Where,

d - diagonal distance between the first orders

ΔT - time between the two exposures

m - magnification of the imaging system

If the pulse separation and magnification of the imaging system are known, the speed of the particles can be determined from the pixel displacement. In the examples shown, the direction and number of pixels travelled across the 2048×1024 image in the 0.5μs pulse duration, are as follows:

Top left image	(18.7±0.2) pixels/s at an angle of 15.5° ± 0.5° to the horizontal
Top right image	(19.7±0.2) pixels/s at an angle of 14.7° ± 0.5° to the horizontal
Bottom left image	(15.4±0.2) pixels/s at an angle of 15.4° ± 0.5° to the horizontal
Bottom right image	(17.5±0.2) pixels/s at an angle of 17.5° ± 0.5° to the horizontal

The accuracy of the velocity measurements is reduced if the particles are not monodisperse, so that a particular pair dominates the field. The errors are particularly severe if there is a single large particle in the image that cross-correlates strongly with all of the individual particles. The accuracy may be improved by pre-processing the image with digital erosion and dilation algorithms to normalise the area of each of the particles.

A further limitation of Fourier transform analysis is that the identity of the particles within each pair is lost and therefore the sign of the velocity vector cannot be determined from the autocorrelation function. This restriction may be overcome if the twin exposures are recorded separately and then cross-correlated to retain directional information. Other methods involve encoding the separate exposures with different phase values [40] or wavelengths [41].

An alternative to computing the correlation function digitally is to use an analogue coherent optical processing system. In practice this is difficult to achieve in a single step using coherent light, because the intensity distribution rather than the complex amplitude must be inverse transformed. Two step methods, which generate an intermediate intensity distribution prior to forming the inverse Fourier transform, involve expensive specialised devices and materials (e.g. spatial light modulators, photorefractives or photochromics). Consequently, most future developments will probably involve digital image processing.

In summary, PIV is a powerful technique that can provide direct measurements of the velocity field. Before the method can be used extensively, however, it is necessary to design an optical system for determining the z-component of velocity. Also, a

fully automated optical and digital processing system will have to be developed, to facilitate rapid extraction and logging of velocity information. A comparison of PIV and other diagnostics should be made to assess the accuracy of particle seeding and data extraction at compressible speeds.

2.2.2 Laser Doppler Anemometry (LDA)

In 1842 an Austrian physicist, Christian Johann Doppler, discovered that when acoustic or electromagnetic radiation is reflected from a moving body, the scattered signal is frequency shifted. The Doppler shift was first exploited in flow analysis by Yeh and Cummins in 1964 [42], when they calculated the velocity of liquids. Similar results were demonstrated in gases by Foreman *et al* in 1965 [43].

When monochromatic light of frequency f is scattered from a seeded flow and examined spectroscopically, a Doppler shifted spectral component at $(f + \Delta f)$ is observed. The Doppler shift of visible light is typically only 600MHz at transonic speeds and therefore accurate calculation of the flow velocity by direct spectral measurement is impractical. A better technique for detecting small frequency differences is to heterodyne the Doppler shifted light with an unscattered reference beam. One particular experimental configuration for performing reference beam LDA is shown in figure 6. Optimum performance of the system requires that the recombined beams are mutually coherent, of similar intensity, and are matched in polarisation. A detailed theoretical and practical discussion of this design, and explanation of the associated signal processing, is given by Drain [44]. Many other similar geometries have been proposed, which differ in the way that the reference beam is generated and recombined with the signal beam.

The signal detected using the reference beam technique is modulated at the beat frequency Δf, and may be analysed electronically by frequency demodulation. Subsequently, the flow velocity may be calculated from the demodulated signal by use of the following equation:

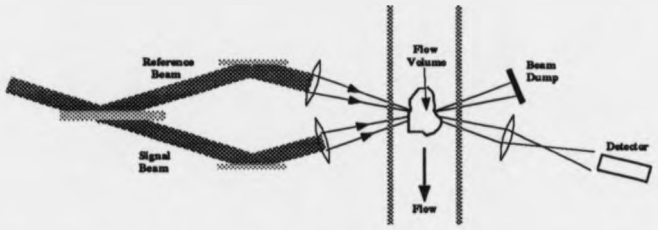


Figure 6: Typical design of a reference beam LDA system

$$\Delta f = \frac{2v}{\lambda}\sin\left(\frac{\alpha}{2}\right) \tag{20}$$

Where,

- v - velocity component orthogonal to the beam bisector
- α - angle between the incident reference and signal beams

Alternatively, two signal beams of similar intensity may be shone through the flow so that they intersect at the measurement point. A self-aligning differential LDA system, originally proposed by Durst and Whitelaw [45], is shown in figure 7.

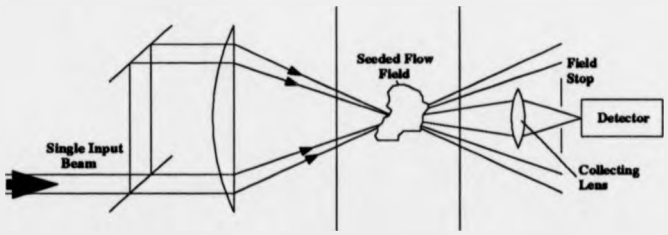


Figure 7: Typical design of a differential LDA system [44]

As the two laser beams are not co-linear, the Doppler shift of the elastically scattered light is different for the two component beams. The temporally modulated signal recorded by a photomultiplier tube can be explained by considering the interference of these two beams:

$$\begin{aligned} I(t) &= |I_0\sin(\omega_1 t) + I_0\sin(\omega_2 t)|^2 \\ &= I_0'\sin^2\left[(\omega_1+\omega_2)\,t/2\right]\left[1+\cos(\omega_1-\omega_2)t\right] \end{aligned} \tag{21}$$

Where,

I_0 - intensity of the two beam (made equal to simplify the analysis)

ω_1, ω_2 - Doppler shifted frequencies of the respective signal beams

I_0' - peak signal intensity detected by the photomultiplier

Only the difference frequency $(\omega_1 - \omega_2)$ of equation 21 is temporally resolved by a detector. The modulated signal is produced when light is scattered from a particle, as it traverses the interference pattern formed by the intersecting beams. If the beam geometry is known, this signal may be used to determine the flow velocity orthogonal to the direction of the fringes. If the beams are separated by an angle β, the relationship between the Doppler shift and particle velocity is:

$$\Delta f = \frac{2v}{\lambda} \sin\left(\frac{\beta}{2}\right) \tag{22}$$

An advantage of differential LDA over the reference beam method is that the frequency of the scattered signal is independent of the angle of reception. Consequently, the entire cone of rays received by the collecting lens may be processed. In comparison, the Doppler shift with the reference beam method is a function of the collecting angle and therefore the aperture must be stopped down to obtain accurate results. This reduces fluence and therefore the differential method is preferred in practical systems where the Doppler signal is weak. A low f-number system is particularly useful when the seeding densities are low and the particles are small.

The highest signal-to-noise power ratio that can be achieved with a heterodyned LDA system is limited by shot noise at the detector, which is governed by:

$$\left(\frac{i_s}{i_n}\right)^2 = \frac{\eta I}{hf\Delta f} \tag{23}$$

i_s - signal current

i_n - noise current

η - quantum efficiency

I - intensity at the detector

In practice, the SNR is reduced by speckle noise, because the scattered signals are of random phase and are detected over a finite collecting area. An additional source of measurement inaccuracy arises due to transit time broadening of the Doppler shift. This occurs because the volume encompassed by the intersecting beams is of finite size, so that when an individual particle passes through the region, the scattered signal lies within a gaussian envelope. Consequently, the signal at the Doppler frequency is broadened and a velocity error arises. Other mechanisms which result in spectral broadening include:

- Variations in the scattering angle.
- Presence of velocity gradients within the illumination region.
- Natural line broadening processes (see section 2.1.2).

The accepted accuracy of LDA is 1-2.5%, or 0.1% when the data is averaged. However, this refers to only a single velocity component and errors are compounded when multiple channel systems are considered.

The directional ambiguity of the velocity measurement may be resolved by frequency shifting one of the beams using an electro-optic Pockel cell, or an acousto-optic Bragg cell. These devices encode the Doppler signal onto a carrier wave by scanning the interference fringes at megahertz modulation frequencies. Aliasing errors may be prevented by ensuring that the associated carrier velocity exceeds all negative particle velocities.

Three beam pairs are required to measure the three velocity components. Measurement accuracy is maximised if orthogonal beam pairs are used, although in practice the beam geometry is usually dictated by optical access. The three channels have to be uniquely encoded so that the Doppler signals may be separated at the detectors without crosstalk. A review of methods used to encode the signals is given by Bell *et al* [46]. The most popular method is to use the three principal lines of the argon-ion laser, although alignment errors occur if the beam paths are long. Owen *et al*

have developed a system [47] which alleviates this problem by using only the 488nm and 514nm argon-ion lines, with the 514nm component split into two orthogonally polarised beams. This system, installed in the NASA Ames 6'×6' supersonic wind tunnel, employed a 15W laser modulated at 40MHz by a quartz acousto-optic Bragg cell. The results from their study show that three component velocity measurements may be made on-line in a production wind tunnel, although frequent alignment and maintenance of the system is required.

System complexity is further increased if whole-field measurements are required and usually it is impractical to obtain the three velocity components at more than one point in the flow. In high speed wind tunnels, problems are experienced in obtaining sufficient seeding densities. Consequently, the Doppler shifted signal is usually weak and photon counting, or correlation, is required to improve the signal-to-noise ratio. Nevertheless, commercial systems are now widely available (e.g. LDA 2000 marketed by Oriel) and LDA has been applied successfully to the study of complex flow regions, such as supersonic boundary layers [48].

In summary, LDA is a proven technique for obtaining velocity information at a few discrete points in a flow. The principal merits of LDA over other diagnostics are that the measurements are non-intrusive and the spatial resolution is very high (20-100μm). In addition, the dynamic range is large, with the technique capable of measuring flow velocities ranging from 1mms^{-1} to supersonic speeds. Although processing times may be very long, the signals are received effectively instantaneously and therefore turbulent phenomena may be studied. In spite of these capabilities, LDA is used infrequently because it is time consuming, expensive in terms of wind tunnel run-time, requires seeded flows, and can only provide spatial information if the results are temporally averaged.

2.2.3 Laser Two Focus (L2F) Velocimetry

In 1968, shortly after development of the laser Doppler system, D. Thompson invented the tracer particle fluid velocity meter [49]. This device was the forerunner to

the laser two focus (L2F) velocimeter (also known as the laser transit anemometer). L2F operates by measuring the time-of-flight of particles as they travel between two focussed laser beams. When the technique was first used to measure flow velocity, the resulting signals were confused by even mild turbulence levels. This situation improved with the introduction of photon correlators [50, 51, 52] and multi-channel analysers [53, 54, 55], which enabled statistical processing of the signals.

L2F has achieved a status similar to LDA and a range of commercial systems are now available. The technique is most suited to analysis of high speed flows which have a low deviation in the local rms velocity (turbulent intensities <10%). L2F can operate in confined regions and therefore the technique has become established in the study of turbomachinery [56, 57, 58]. The method has been used to a lesser extent in wind tunnel testing [59, 60] and in the study of jets [61].

Data extraction from L2F measurements is accomplished by statistical methods and therefore the accuracy is governed by the number of signal pairs sampled. A detailed analysis of the mathematics involved in processing L2F signals is given by Schodl [62]. When turbulent intensities approach 10%, the variation in the velocity measurements is large and prohibitively long data acquisition times are required. At higher turbulence levels, the mean velocity tends to zero and meaningful results cannot be obtained. The practical difficulties encountered when applying L2F to production testing are largely similar to those of LDA.

2.3 Flow Diagnostics based on Inelastic Scattering

2.3.1 Absorption/Emission Spectroscopy

Emission and absorption spectroscopy may be used to analyse electromagnetic radiation after it has inelastically scattered from molecules in a flow. Spectroscopic measurements are made by tuning the wavelength of a dye laser to a resonant absorption peak. The absorption or emission spectra may then be analysed to determine temperature, pressure, species concentration and flow velocity. Optical absorption is the least complicated spectroscopic diagnostic to implement and combustion processes are ideally suited to examination by this method. In particular, species concentrations may be calculated for OH, NH, NH_2, CO and O_2.

The absorption of a quasi-monochromatic beam by a gas is defined by the Beer-Lambert law:

$$\Delta I/I_0 = 1 - \int_0^L \exp[-\mu_a p(x,y,z)]dl \tag{24}$$

Where,

$\Delta I/I_0$ - fractional attenuation of the incident laser beam
μ_a - absorption coefficient
$p(x,y,z)$ - partial pressure of the absorbing species
L - integration length through the flow

The absorption coefficient is a function of temperature, oscillation strength, line shape, and Boltzmann fraction of the resonant transition. Flow information may be extracted by looking at transitions which are only strongly dependent on one of these variables.

An advantage of optical absorption over other spectroscopic techniques is that low power lasers may be used, with optical fibres used to deliver the beams. A flexible delivery system is useful because absorption measurements are integrated through the flow and therefore beams must be passed through the test section at many angles to enable tomographic reconstruction of the three-dimensional data. Tomographic analysis

may introduce large errors into the data, because the attenuation of the transmitted beam is small and difficult to measure accurately. The absorption coefficient may be determined more precisely if amplitude modulation is used.

Chang *et al* avoided the data extraction problem of path integrated measurements, by studying a two-dimensional flow in a shock tube [63]. Nitric oxide (NO) was created in the heated flow and a continuous-wave ring dye laser was tuned to the characteristic 225nm absorption band. Resonances of N_2 and O_2 were not chosen because their major electronic absorption bands lie deeper in the UV. This paper includes a detailed description of the data reduction process for calculating velocity, temperature, pressure, density and mass flux. At present, absorption methods are not used routinely because of their limited accuracy in three-dimensional flows and the requirement for an expensive tunable laser.

2.3.2 Laser Induced Fluorescence (LIF)

Laser induced fluorescence (LIF) is similar to absorption spectroscopy in that it involves resonantly exciting a particular molecular electronic transition in a flow. Unlike absorption measurements, however, the resulting signal is the result of a spontaneous decay of the excited state. If a Stokes transition is involved, the excited atom does not return to its initial state after emitting a photon, but instead it drops to an interim state with the emission of a photon of lower energy. The excess energy promotes the molecule into a higher vibrational, rotational or electronic state, which may result in 'quenching' of the signal if there is non-radiative collisional energy transfer between molecules. Quenching complicates data analysis, although it is possible to choose transitions that are unaffected. The anti-Stokes process is a similar interaction, which occurs when the molecule is pre-excited so that the emitted photon is of higher energy than the one absorbed. If the transition happens rapidly ($\sim 10^{-7}$s) then the process is known as fluorescence, whereas phosphorescence occurs when the transition is delayed (sometimes many minutes).

LIF is most effective when applied to a clearly defined electronic transition, so

that the emitted radiation can be identified from noise. Consequently, LIF is popular for studying reactive flows [64] where there are an abundance of combustion products with suitable spectroscopic characteristics (e.g. the hydroxyl radical). In non-reactive flows, fewer transitions are available and artificial seeding is often necessary. Nitric oxide (NO) and iodine are two strong candidates for seeding nitrogen and air flows. When NO is chosen, the ratio of the fluorescent intensities at 225.857nm and 225.875nm may be used to calculate temperature. Unfortunately NO is toxic and the fluorescent signal is quenched by oxygen, thereby reducing the signal-to-noise ratio. Many other good seeding agents for LIF are also highly toxic and unsafe to use.

The recent development of excimer lasers that operate in the far UV, has enabled flow analysis using the fluorescence of intrinsic oxygen molecules. Molecular oxygen has a ro-vibronic resonant transition at 193.055nm in the Schumann-Runge band [65]. This absorption line may be excited by using a pulsed ArF laser to pump a tunable dye laser, with the output beam passed through a frequency doubling non-linear crystal to obtain wavelengths in the far UV. In addition to the convenience of operating in unseeded flows, the transition is largely unaffected by collisional quenching. Unfortunately the detected signal is a function of temperature and density, and therefore independent measurements must be made to separate these variables. Technical difficulties are compounded by the complexity and expense of the multiple laser system.

To localise LIF measurements and provide sufficient power densities, the incident beam must be focussed into the flow volume. Recently, planar laser induced fluorescence (PLIF) has been used to obtain a plane of measurements by expanding the beam into a light sheet [66]. With PLIF, the fluorescent intensity is weak and therefore an image intensifier is required to amplify and gate the signal detected by the CCD camera. A single velocity component may be ascertained from the Doppler shift of the scattered light, although the spectrum of the incident laser beam must be ascertained to calculate the shift accurately. Alternatively, counter propagating beams may be heterodyned so that the detected signal is modulated at twice the Doppler frequency. To measure all three velocity components requires the use of three orthogonal light

sheets, with separate optics and detectors.

The number density of a particular molecular species may be determined from the fluorescent intensity, by measuring the population of the excited state. Similarly, the temperature of the gas may be calculated from the Boltzmann distribution, if the relative populations of two states close to the ground level are measured. The pressure field may be calculated from temperature and number density, by using the equation of state for an ideal gas. In practice, the extraction of data from LIF measurements is time consuming (typically two minutes per point with little sign of future improvement) and is complicated by several factors:

- Spectroscopic broadening processes.
- Radiation trapping.
- Quenching.
- Rotational redistribution.
- Low signal strength.
- Amplitude noise due to imperfections and dirt on the optics.
- Calibration difficulties.
- Scarcity of electronic transitions which separate the variables of velocity, temperature and pressure.

In general these problems can be overcome by making additional measurements with a complementary diagnostic, although this inevitably increases the cost and complexity of experimental flow analysis. When a light sheet is used, the intensity distribution of the two-dimensional image is usually a result of more than one variable. Hence, the spatial information is gained at the expense of spectroscopic data and the technique becomes largely qualitative. If suitable electronic transitions are available, LIF can measure velocity to an accuracy of approximately $5ms^{-1}$. Although this figure

is acceptable at compressible flow speeds, LIF is primarily confined to combustion research. However, the technique is likely to gain in popularity for transonic wind tunnel testing, because of the increased availability of high power UV lasers.

Recently, a similar diagnostic to LIF has been developed which exploits fluorescence/phosphorescence, of photochromic/photoluminescent chemicals seeded into the flow. The technique, known as laser induced photochemical anemometry (LIPA), entails writing a grid of laser lines into the seeded flow with a pulsed laser. This pattern persists with a lifetime of 1μs-10s depending on the chemical used and distorts as it propagates with the flow. Images of the grid can subsequently be used to calculate flow velocity, vorticity and strain rate [67]. The principal limitation of LIPA is the inconvenience caused by seeding the flow with exotic chemicals.

2.3.3 Raman Scattering

The spontaneous Raman effect was first predicted in 1923 by Adolf Smekal and observed experimentally in 1928 by Sir Chandrasekhara Raman. Initially the effect could not be exploited because only low intensity mercury sources were available; however this situation changed immediately with the advent of the laser. Linear spontaneous Raman scattering, also known as conventional old-fashioned ordinary Raman scattering (COORS), involves either Stokes or anti-Stokes transitions. The COORS process is weak because of its linearity and the radiated photons are incoherent. Consequently, the scattered signal is omni-directional and difficult to detect amongst noise.

A more useful flow diagnostic is provided by stimulated Raman scattering. This effect was first observed in 1962 by Eric Woodbury, when he noticed that a second coherent beam was generated when a pulsed ruby laser beam interacted with a nitrobenzene Kerr cell. In this original experiment, a component of the ruby beam at 694nm was shifted to a longer wavelength of 766nm. In coherent Stokes Raman spectroscopy (CSRS - pronounced scissors), the scattered photons are amplified by stimulated emission and a large proportion of the incident beam energy is converted into the

new spectral component. The emitted beam is more intense than with COORS and the photons are uni-directional because of their mutual coherence. Other stimulated Raman scattering processes exist and they are classified according to the electronic transitions involved:

- Coherent Anti-Stokes Scattering (CARS).
- Inverse Raman Scattering (IRS).
- Coherent Anti-Stokes Raman Ellipsometry (CARE).
- Raman Induced Kerr Effect Spectroscopy (RIKES).
- Photo-Acoustic Raman Scattering (PARS).
- Higher Order Stokes Effect Scattering (HORSES).

Raman scattering can be used to measure species concentration and flow velocity. Rotational and vibrational temperatures can also be determined spectroscopically and translational temperature may be estimated from the Doppler shift of the spectra. The most important coherent Raman processes are IRS, CARS and CSRS. With CARS and CSRS, the Stokes and anti-Stokes scattered radiation is detected, whereas the attenuation of the incident beam is measured in IRS. Raman spectroscopy is relatively simple to perform experimentally and a detailed account is given by Greenhalgh [68]. Basically, the incident laser beam is focussed into the flow volume and scattered radiation is analysed using a monochromator or spectrometer. The Raman scattered spectrum is highly complex for even simple molecules and quantum mechanics is required to provide a rigorous interpretation of the results.

The acquisition of spectroscopic information from Raman scatter is simplified if two lasers are used. The first laser provides a pump beam, which excites the molecules in the flow, while the second laser stimulates a decay of the excited states. Spectroscopic information may be obtained by measuring the scattering cross-section as the wavelength of the probe beam is varied (by using a tunable laser). The angle of the

Raman scattered beam can be controlled by adjusting the orientation of the pump and probe beams.

Stimulated Raman scattering has been used predominantly in combustion research and it has rarely been used as a diagnostic in non-reactive flows. In particular, it is suitable to highly luminous harsh environments such as those present in combustion engines [69] or coal furnaces [70]. Wind tunnel use is likely to increase, because Raman scattering is one of the few techniques that can provide simultaneous measurements of temperature, density and velocity. At present, however, practical systems employ multiple laser sources, which are prohibitively expensive and difficult to maintain. Data analysis is complicated, although it is possible to obtain an accuracy of typically $\pm(3-6)\%$ for instantaneous results and $\pm(1-5)\%$ when temporal averaging is used [71].

2.4 Non-Scattering Whole-Field Techniques

2.4.1 Schlieren Testing

Schlieren are inhomogeneities in a refractive index field and the name is derived from the German word *schlere* meaning streak. The schlieren effect was first reported by Robert Hooke in 1665 when he interpreted the scintillation of stars [72]. In 1859 the first schlieren system was constructed by Focault for examining astronomical mirrors and lenses (Focault knife test). The technique was further developed by Toepler in 1864, for improving the general quality control of optics [73]. Schlieren testing was first used scientifically in aircraft design by Schardin in 1934 [74].

Schlieren testing is now an established technique for producing pictorial information of predominantly two-dimensional and axisymmetric compressible flow fields. Shock waves, expansion fans, stagnation regions, and other features encorporating high density gradients, may be visualised and studied dynamically. Figure 8 shows a schematic representation of a schlieren system. Light from a source of high spatial coherence is collimated by a spherical or parabolic mirror and propagated at right angles through the air flow. Spherical mirrors generate aberrations, but often provide satisfactory imaging and are less expensive. The transmitted beam is subsequently collected by a second identical mirror and the wind tunnel centre-line is imaged by a lens onto a detector or photographic plate. A knife edge, or other aperture function, is placed at the focus of the lens to block refracted rays and thereby modify the image intensity.

Historically, sources of medium temporal coherence (e.g. mercury gas discharge) were used to reduce dispersion, increase image sharpness and ensure good collimation of the illumination. Recently, quasi-monochromatic radiation from lasers has been used, although the high temporal and spatial coherence can often generate diffraction noise and speckle. An advantage of intense laser illumination, is that self-luminous flows may be studied without fogging of the recorded image. Diffraction noise from coherent illumination can be minimised by sharply focussing any edges in the field-

of-view and speckle may be suppressed by eliminating scatter.

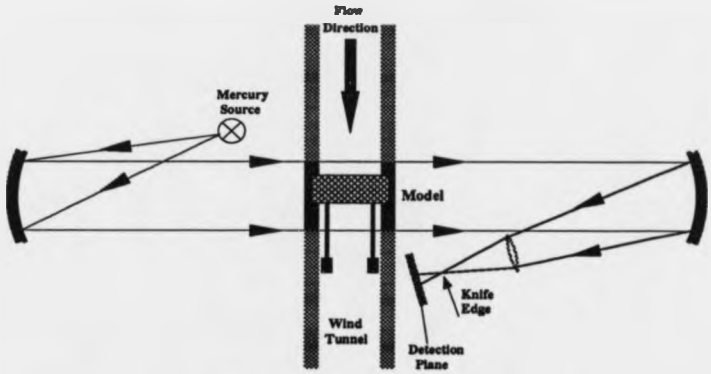


Figure 8: Conventional schlieren system

When the flow is incompressible at low subsonic speeds, the light propagates without deviation and may be imaged unobstructed through an aperture onto a photographic emulsion or video camera. The recorded image exhibits no flow features at this stage, because the light is not refracted and there is no change in the intensity distribution of the image due to blocking of the rays. At higher speeds, the flow becomes compressible and the density of the fluid begins to vary. If the air is calorically perfect so that the specific heats at constant pressure and volume (c_p and c_v) are independent of temperature, and isentropic conditions are assumed (frictionless adiabatic), the density ratio is related to Mach number by:

$$\frac{\rho_0}{\rho} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{\frac{1}{(\gamma-1)}} \tag{25}$$

Where,

- ρ_0 - stagnation density when the fluid is bought adiabatically to rest
- ρ - local static density
- M - local Mach number
- γ - ratio of specific heats (c_p/c_v)

Figure 9 illustrates that at Mach 0.30 the static air density ($\gamma = 1.4$) is only 4.6% lower than the stagnation density. At higher Mach numbers the density follows a fifth

power law and compressible effects become progressively stronger. Although flows are never completely incompressible, density changes for speeds less than Mach 0.1 are negligible ($\ll$1%) and this defines a regime above which schlieren effects become visible.

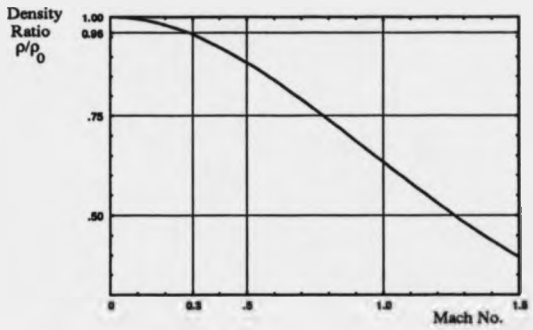


Figure 9: Density ratio ρ/ρ_0 as a function of Mach number

The refraction of light by a flow is governed by the changes in air density that occur at compressible speeds. Light travels slower in dense regions of the flow and these parts of the wavefront are retarded so that the associated ray is bent. The refractive index, defined as the ratio of the speed of light in a vacuum to that in the medium, is used to quantify the amount of refraction experienced during ray propagation. If the flow consists of two uniform regions of different refractive index (figure 10), the angle of refraction at the interface is governed by Snell's law:

$$n_i \sin\theta_i = n_t \sin\theta_t \tag{26}$$

Where,

n_i, n_t - refractive indices of the media on the incident and transmitted sides of the interface

θ_i, θ_t - angles of the incident and transmitted rays with respect to the normal of the interface

The refractive index is determined by the Lorentz force exerted by the electromagnetic optical field on the nuclei and electrons of the gas. A rigorous analysis of

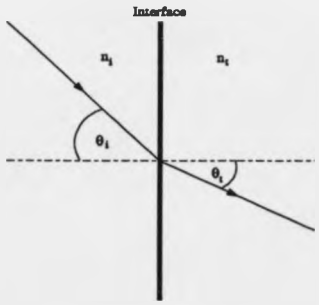


Figure 10: Refraction at an interface

the response to an impinging field requires the use of quantum mechanics, but a satisfactory description is provided by assuming that the electrons experience a driving force of sinusoidal time dependence, which is impeded by a linear restoring force and damping.

The classical model predicts that the refactive index rises monotonically with frequency, except when there is a resonance between the applied field and the molecular vibrations. At resonant frequencies, anomalous dispersion occurs and there is a sharp drop in the real component of the complex refractive index, coupled with strong absorption. Transparent gases do not exhibit strong absorption in the visible spectrum and therefore are unaffected by anomalous dispersion. In this regime, a simple linear relationship, described by the Gladstone-Dale equation, can be used to calculate the density of a gas from its refractive index:

$$n(x,y,z) = 1 + K_g \rho(x,y,z) \tag{27}$$

Where,

$n(x,y,z)$ - refractive index distribution
K_g - Gladstone-Dale constant
$\rho(x,y,z)$ - density field generated by a compressible flow

The Gladstone-Dale constant (defined here as K_g instead of K to avoid confusion with the Boltzmann constant) is a function of gas species and is weakly dependent

on the wavelength of the transmitted radiation. However, it is independent of temperature and pressure under moderate conditions, and therefore flow density may be calculated from equation 27 if quasi-monochromatic radiation is used. In most non-reacting flows, gas composition is uniform and therefore K_g is spatially invariant. Values of K_g for important laser wavelengths, at standard temperature and pressure, are listed below:

514.5nm (green argon ion line) - $0.230\times10^{-3}\mathrm{m^3kg^{-1}}$
632.8nm (red helium-neon line) - $0.228\times10^{-3}\mathrm{m^3kg^{-1}}$
694.3nm (ruby) - $0.227\times10^{-3}\mathrm{m^3kg^{-1}}$

Snell's law is only applicable for refraction at an interface, whereas compressible flows consist of a continuously varying refractive index distribution. Light does not obey rectilinear propagation in these circumstances (figure 11) and rays follow curved paths defined by a vector differential equation founded on Fermat's principle of least times [75]:

$$\frac{d}{ds}\left[n(x,y,z)\frac{d\underline{r}}{ds}\right] = \nabla n(x,y,z) \tag{28}$$

Where,

$\underline{r}$ - position vector of a point on the ray
s - distance along the ray from a fixed origin
$d\underline{r}/ds$ - Unit vector normal to the wavefront (ray direction or Poynting vector)
∇ - Gradient operator

When the refractive index field is uniform, as with an incompressible flow, the solution of equation 28 is a family of linear ray paths described by:

$$\underline{r} = \underline{r}_0 + s\underline{b} \tag{29}$$

Where $\underline{b}$ is a unit vector parallel to the ray direction and $\underline{r}_0$ is a position vector on the ray.

If the refractive index varies throughout the field, the rays follow more complicated curved paths and bend towards regions of higher refractive index. Consequently, the

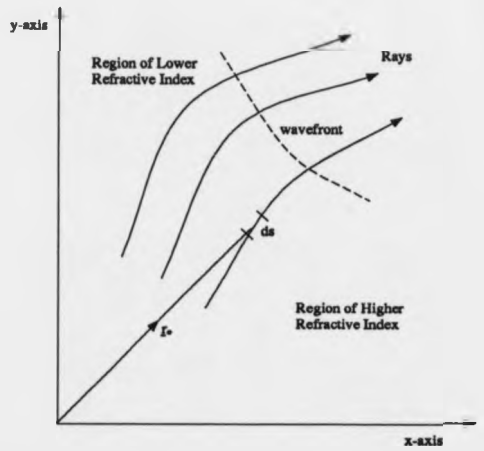


Figure 11: Refraction by a refractive index field

light transmitted through the flow is incident at different angles to the imaging lens of a schlieren system. When the flow contains large density gradients the focus is broadened, with the most highly refracted rays lying furthest from the optical axis. The aperture only passes paraxial rays and therefore, when the test section is imaged, flow regions containing large density gradients (e.g. shocks) appear in shadow. A knife edge is commonly used to block rays refracted parallel to one particular axis, and the orientation of the blade with respect to the focus, determines the direction of sensitivity. Other aperture functions may be employed and a banded colour filter is used in colour schlieren [76, 77, 78] to emphasize the direction, approximate magnitude and sign of the density gradients.

The sensitivity of a schlieren system is governed by the change in intensity ΔI that occurs at a point in the viewing plane, due to a given density gradient in the flow. When the collimated illumination is propagated through the flow parallel to the z-axis, the sensitivity $\Delta I/I$ for small angles of refraction is:

$$\frac{\Delta I}{I} = \frac{f}{w}\int_0^L \frac{1}{n}\frac{\partial n}{\partial x}dz \tag{30}$$

Where,

f - focal length of the imaging lens

w - distance from the optical axis to the knife edge

L - integration length through the flow

x - direction of sensitivity (orthogonal to the knife edge)

Equation 30 shows that the sensitivity improves as more of the peripheral rays in the focus are removed, although this occurs at the expense of optical throughput. Unfortunately other physical processes, such as non-uniformities in the illumination and scatter from optical defects, may also result in intensity variations in the image. Consequently, schlieren testing is difficult to use quantitatively and, at best, it can only give estimates of the integrated density gradient along a particular ray path. Due to the difficulties of data extraction, the technique is usually used qualitatively for locating the position of features with large density gradients (particularly shocks).

2.4.2 Shadowgraphy

At compressible speeds, flow features may become visible without the insertion of an aperture at the focus of the imaging system. Flow visualisation occurs due to non-uniform refraction of the illumination by the phase object and the effect is exploited in a technique known as shadowgraphy. Shadowgraphy is popular due to its inherent simplicity and high sensitivity to large density gradients (especially shock waves). The principles of operation are best understood by referring to figure 12, which represents a uniform density gradient in a flow by a prism. When there are no density variations in the beam path, the rays are undeviated and the plane of observation is evenly illuminated. Introduction of the prism causes refraction of the rays according to Snell's law (equation 26).

The prism refracts light into a region already illuminated by unrefracted light and a bright zone is formed. A complementary shadowed region is created due to loss of the refracted rays. The contrast of shadowgraph images is proportional to the second spatial derivatives of density (equation 31) and sensitivity increases as the viewing plane is moved further from the phase object. It is difficult to extract density

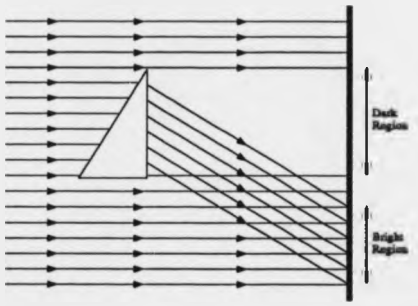


Figure 12: Formation of a shadowgraph image by a density gradient

information using shadowgraphy, because intensity based images are inherently noisy and cannot be integrated accurately.

$$\frac{\Delta I}{I} \propto \int_0^L \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right] \ln[n(x,y,z)]dz \tag{31}$$

Shadowgraph features are weak when the object is sharply focussed and therefore a compromise must be made between image definition and sensitivity. When laser illumination is used, image clarity is further degraded by diffraction at the edges of the model and other scattering centres. Recent research has investigated the use of shadowgraphy with high repetition rate pulsed lasers to study transient flow features. In particular, Brownell *et al* have studied the interaction between a shock wave and a turbine blade [79]. Metal vapour lasers are a popular choice, because they can supply moderate pulse energies ($\sim$ 1mJ), at frequencies exceeding 10kHz.

2.4.3 Interferometry

When a viewing screen is placed in the region of two overlapping beams of light, the observed intensity distribution may not be simply the algebraic sum of the individual intensities. If the two beams are mutually coherent, and their polarisation vectors are not orthogonal, a periodic intensity distribution is produced. This fringe pattern is created by the interference of temporally correlated optical wavefronts. Dark

fringes occur where the component waves are in anti-phase and interfere destructively, whereas bright fringes are formed in regions of constructive interference.

Optical interference was first studied by Robert Hooke, Robert Boyle and Isaac Newton in the 17th century. The first theoretical analysis of interference phenomena in terms of wave properties, was made nearly two centuries later, when Thomas Young investigated the fringe patterns generated by a pair of closely spaced narrow slits. In recognition of this work, the interference patterns formed by this arrangement came to be known as Young's fringes. Subsequent research has led to the evolution of many different interferometers for use in applications as diverse as spectroscopy and lens testing.

Shadowgraph and schlieren images are produced when light rays are blocked or redistributed after refraction through a compressible flow and the techniques may be described by classical geometrical optics. In contrast, the wave properties of light must be invoked to explain the mechanics of interferometry. Also, it is usually assumed when analysing interferograms quantitatively, that the wavefronts (surfaces of constant phase) propagate along linear ray paths. A refractionless approximation is only valid if the density gradients in the flow field are small, so that a linear integration path can be assumed.

Throughout this thesis, transonic flows are studied which contain mainly slowly varying density fields. Consequently, a refractionless approximation is adopted in the analysis of the fringe patterns. Significant refraction occurs in the boundary layer and wake of transonic flows, and therefore these regions may not be analysed using this approach. A rigorous study of the effects of refraction on holographic measurements is presented by Vest [80] and the influence on wind tunnel measurements is discussed by Lanen [81].

Interferometry is most applicable to flows where the density is varying slowly over the whole-field and any shock waves are weak or normal to the viewing direction. Conversely, schlieren and shadowgraphy are most useful when the density field is highly discontinuous, with regions of uniform density that can be measured at a

few discrete points. Interferometry, schlieren and shadowgraphy are complementary techniques, and in chapter 3 it is shown that all three types of image can be retrieved from a single double-exposure hologram.

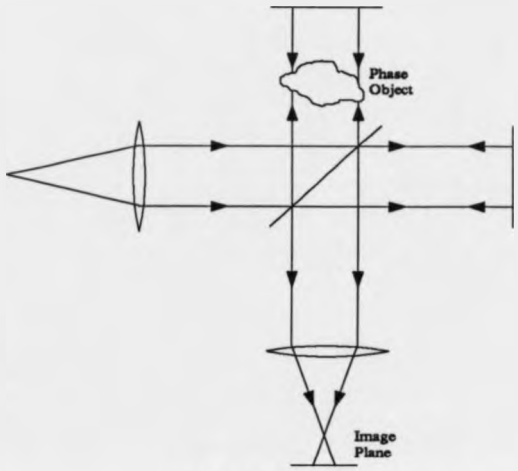
Historically, flow analysis by interferometry was accomplished using a Mach-Zehnder or Michelson interferometer. These systems obey identical principles (figures 13), except that the object wave of the Michelson interferometer traces a double path through the flow and therefore has twice the sensitivity. To understand the operation of an interferometer it is simpler to consider the single-pass Mach-Zehnder configuration.

In a Mach-Zehnder interferometer, the input wavefronts are collimated and divided in amplitude by a beam splitter. One of these components, denoted the object beam, is propagated through the phase object and its wavefronts are retarded or advanced. The second component, known as the reference beam because its planar wavefronts act as a reference to the distorted wavefronts of the object beam, is passed around the flow field unperturbed. The phase retardation imparted by the flow may be measured by interfering the object wave with the planar reference wave, so that a fringe pattern is generated. The intensity distribution of this pattern is modulated by the phase of the object beam, with anti-phase components generating dark fringes and in-phase regions interfering constructively to produce bright fringes. Interferometrically encoded phase information is ambiguous to multiples of 2π, because the sign of the phase change between adjacent fringes is unknown.

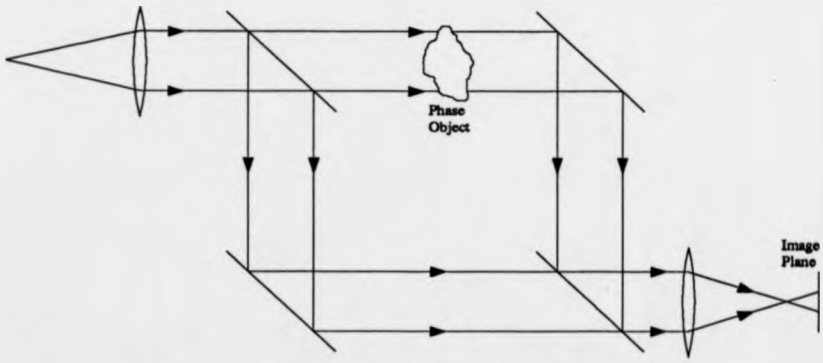
The difference in optical path length $\Delta l(x,y)$ travelled by a wavefront passing through a compressible flow field of length L, compared to one traversing static air of refractive index n_0, is given by the path integral in equation 32.

$$\Delta l(x,y) = \int_0^L [n(x,y,z) - n_0]dz \tag{32}$$

Hence, the phase change experienced by the object wavefront is:



Michelson Interferometer



Mach-Zehnder Interferometer

Figure 13: Michelson and Mach-Zehnder interferometers

$$\Delta\Phi(x,y) = \frac{2\pi}{\lambda}\Delta l(x,y) \tag{33}$$

The fringe pattern $i(x,y)$ generated by the interference of the object and reference waves is given by the coherent superposition of the beams:

$$\begin{aligned} i(x,y) = & \ |r(x,y) + o(x,y)\gamma\exp[j\Delta\Phi(x,y)]|^2 \\ = & \ |r(x,y)|^2 + \gamma^2|o(x,y)|^2 + 2r(x,y)o(x,y)\gamma\cos[\Delta\Phi(x,y)] \end{aligned} \tag{34}$$

Where,

$r(x,y)$ - amplitude of the reference wave
$o(x,y)$ - amplitude of the object wave
γ - degree of coherence between the object and reference waves

Equation 34 may be simplified by grouping terms which describe the bias level and modulation of the fringes:

$$i(x,y) = a(x,y) + b(x,y)\cos[\Delta\Phi(x,y)] \tag{35}$$

Where,

$a(x,y)$ - unmodulated background intensity of the fringe pattern
$b(x,y)$ - fringe contrast

The contrast of the fringes $b(x,y)$ is determined by a range of factors:

- Intensity ratio of the object and reference beams.
- Spatial coherence (degree of phase correlation across a wavefront).
- Temporal coherence (monochromicity of the source and effective length of phase correlated wave trains).
- Levels of vibration present during recording of the fringe pattern.

Unlike schlieren and shadowgraph techniques, interferometry encodes flow information as a sinusoidal intensity modulation (equation 34). It is possible to calculate the phase of the fringes to an accuracy of $\sim \pi/2$ by simple interpolation and, if the fringe number is high, a reasonable estimate for the density variations across a two-dimensional flow field can be obtained. In most instances, however, automated image processing with a higher interpolation accuracy is desirable and this has led to the development of computer aided fringe analysis (chapter 4).

There are two principal reasons why interferometry is more conducive to quantitative analysis than schlieren and shadowgraphy. Firstly, the density information is phase encoded and therefore it can be extracted more accurately than a direct intensity modulation. Secondly, interferometry is sensitive to density changes rather than its spatial derivatives and therefore less data processing is required to relate the measurements to aerodynamic parameters.

A principal disadvantage of optical diagnostics that are not based on localised scattering, is that it is difficult to invert path-integrated data. This problem is exacerbated when the flow is highly three-dimensional and tomographic reconstruction algorithms must be applied to obtain quantitative information. An additional restriction with interferometry is that the two beams must be kept interferometrically stable along the path between the beam splitter and the plane of interference.

Interferometry reveals the phase difference between the object and reference waves. Hence, the exact shape of the reference wave must be known before the absolute phase distortion of the object wave can be determined. A planar reference beam is normally used to simplify analysis; however, it is difficult to produce an extended plane wave without the use of prohibitively expensive mirrors or lenses. In addition, optical surfaces in the object beam path must be planar, or of known shape, to prevent the introduction of systematic phase errors.

The dynamic range of interferometry is restricted by the Gladstone-Dale constant, which determines the fringe densities for a particular set of system parameters. When the fringe densities are high, speckle noise and aliasing at the detector limit the

extraction of data. Conversely, when the fringe densities are low it is difficult to interpolate between fringes, due to variations in fringe contrast and because of the restricted dynamic range of the image sensor.

Initially, Mach-Zehnder interferometry was used infrequently because of the precise path matching (accuracy of typically < 0.2mm) required for sources of medium coherence (e.g. mercury discharge). This situation changed markedly with the advent of the laser which provided coherence lengths in excess of 1m. The development of holographic interferometry since 1965, has extended the possibilities of conventional interferometry, by enabling the recording of complicated reference wavefronts. In the extreme this has meant that diffusely scattered wavefronts may be compared interferometrically to ones that have been retarded by a flow. Consequently, holographic interferometry may be applied to large apertures, without the need for prohibitively expensive high quality optics. Also, because wavefront reconstruction is possible using holography, the beams forming the interferogram do not need to co-exist. These advantages are important in flow analysis where it is necessary to produce an initial 'phase calibration' of the wind tunnel working section, prior to the recording of a second holographic exposure with the flow present.

Holographic interferometry has been used extensively in the analysis of numerous flows encountered in industrial research and development. Important areas of application are:

- Investigation of rotating machinery [82, 83].
- Aerofoil design [84, 85, 86].
- Study of boundary layers [87, 88, 89].
- Shock wave research [90, 91].
- Hypersonics [92, 93].
- Analysis of jets and sprays [94].
- Study of crystal growth in space [95].

- Transonic flow studies of helicopter rotor blades [96].

The automated extraction of three-dimensional data from holographic interferograms is currently the largest single problem which must be overcome before the technique can be used routinely.

2.5 Conclusions and Comparison of Optical Flow Diagnostics

Technological advances in high power lasers, CCD cameras and computer hardware, have led to the evolution of a host of complementary optical flow diagnostics. Flow density may be calculated from refractive index measurements, or the velocity field can be determined by analysing the trajectories of seeding particles. Alternatively, inelastic scattering processes may be used in conjunction with spectroscopic analysis, to determine species concentration, temperature and Mach number. At speeds lower than Mach 0.3 the flow is virtually incompressible and therefore the refractive index distribution is uniform, unless temperature gradients or regions of unmixed gas are present. Consequently, at these speeds, particle seeding is required to extract information non-intrusively. The inertia of micron-sized particles is negligible in incompressible flows and therefore measurement accuracy is high. LDA, L2F and PIV are the principal techniques for non-intrusive flow analysis at these speeds.

Laser Doppler and 'time-of-flight' techniques (L2F) measure flow velocity non-intrusively, by analysing light scattered elastically from the seeding particles. A principal disadvantage of these methods is that they measure Mach number at only a single point. The acquisition of whole-field data requires scanning of the focussed beams, which is time consuming and demanding on the alignment of the system. Also, because the data is acquired sequentially, the measurements must be temporally averaged to ensure that transient events do not produce spatial inconsistencies.

LDA and L2F measure only a single velocity component, unless complicated optical configurations are used that generate three orthogonal beam pairs. High measurement accuracy is possible if mono-disperse seeding particles are used, although it is difficult to achieve uniform entrainment at the correct density. Both techniques are dependent on statistical analysis of the scattered signals and therefore measurement accuracy is highest (typically 1%) when the data is temporally averaged. L2F is unsuccessful in turbulent flows because the local mean velocity can approach zero. LDA and L2F have been used predominantly in the design of turbomachinery.

PIV offers greater potential than LDA and associated techniques for wind tunnel testing, because it can simultaneously record the flow velocity of an entire plane. Two velocity components may be determined from a single photograph and the third component can be obtained by holographic recording or stereoscopic viewing. The penalty paid for the increase in flow information is the greater complexity of the data capture and processing systems. Particle pair identification, and determination of the sign of the velocity vector, complicate the analysis and are topics of current research.

The restrictions imposed by particle seeding may be overcome by studying the kinematics of molecules within the flow. In practice, this may be achieved by spectroscopically examining the inelastic interaction of light with the molecules. The most promising techniques are those which depend on the natural constituents of the flow, instead of artificially introduced substances. For example, LIF and stimulated Raman scattering are used regularly to provide point measurements in combustion research [5, 97, 98, 68].

LIF may be used to determine species concentrations, density, temperature and velocity, from the fluorescent emission spectrum at a point in the flow. LIF is most successful when readily identifiable electronic transitions are available. Consequently, the technique has been used predominantly in reactive flows where the signature from combustion products (e.g. OH and NO) is easily separated from noise. In wind tunnel testing, attempts have been made to exploit LIF using fluorescence of naturally occurring oxygen molecules [65]. This research is still in its infancy and exotic lasers are required to excite the energetic electronic transitions. Alternatively, molecular seeding may be deployed, although this is limited at compressible speeds by the large mass flow rates.

Stimulated Raman scattering, which utilises Stokes and anti-Stokes transitions (CSRS and CARS), is popular in the examination of self-luminous and high enthalpy flows. Temperature, species concentrations and velocity have been measured in hostile environments, ranging from combustion engines to coal furnaces [97, 98, 68]. Rotating machinery are particularly suited for investigation, because the swept measurement

volume circumvents the need for beam scanning. Raman scattering has not been used in production wind tunnel testing, due to the lack of suitable electronic transitions. This situation may change, however, with the increased availability of high power excimer lasers which operate in the UV.

Alternatively, unseeded flows may be studied by measuring the distortion of optical wavefronts as they propagate through a density field. This approach is particularly attractive in non-reactive high speed flows, where particle seeding is impractical and there are no suitable electronic transitions available for studying inelastic scattering. Schlieren, shadowgraph and moiré are intensity based techniques that examine the spatial derivatives of density. Although useful for examining shocks, and other features containing large density gradients, these diagnostics are qualitative and only appropriate for preliminary investigations. In contrast, interferometry is sensitive to absolute density changes in a flow and fringe analysis enables accurate data extraction.

Interferometric measurements are whole-field because the wavefronts are extended in space. In addition, the technique is non-intrusive because the plane of interference is external to the flow volume. Interferometry is extremely sensitive and, in principle, can be used to measure optical path length changes of less than 10nm. Unfortunately, wavefront distortion may also occur due to the movement of optical surfaces and therefore the accuracy is reduced by vibration. The stability requirements, and need for precision optical components, have been largely removed by the development of pulsed holographic interferometry. A more fundamental limitation of interferometry is that the phase retardation of a wavefront is integrated along the optical path. Consequently, depth information can only be retrieved for a general three-dimensional flow, if multi-directional interferometric data is reconstructed tomographically.

In conclusion, LDA and L2F are suitable for studying low speed flows and PIV offers great potential for extending these methods to whole-field testing. At transonic speeds these techniques are still valid, although difficulties arise in seeding the flows at sufficient densities with mono-disperse particles. Also, sub-micron sized particles must be used at high speeds to ensure that they follow the flow accurately and therefore the

volume circumvents the need for beam scanning. Raman scattering has not been used in production wind tunnel testing, due to the lack of suitable electronic transitions. This situation may change, however, with the increased availability of high power excimer lasers which operate in the UV.

Alternatively, unseeded flows may be studied by measuring the distortion of optical wavefronts as they propagate through a density field. This approach is particularly attractive in non-reactive high speed flows, where particle seeding is impractical and there are no suitable electronic transitions available for studying inelastic scattering. Schlieren, shadowgraph and moiré are intensity based techniques that examine the spatial derivatives of density. Although useful for examining shocks, and other features containing large density gradients, these diagnostics are qualitative and only appropriate for preliminary investigations. In contrast, interferometry is sensitive to absolute density changes in a flow and fringe analysis enables accurate data extraction.

Interferometric measurements are whole-field because the wavefronts are extended in space. In addition, the technique is non-intrusive because the plane of interference is external to the flow volume. Interferometry is extremely sensitive and, in principle, can be used to measure optical path length changes of less than 10nm. Unfortunately, wavefront distortion may also occur due to the movement of optical surfaces and therefore the accuracy is reduced by vibration. The stability requirements, and need for precision optical components, have been largely removed by the development of pulsed holographic interferometry. A more fundamental limitation of interferometry is that the phase retardation of a wavefront is integrated along the optical path. Consequently, depth information can only be retrieved for a general three-dimensional flow, if multi-directional interferometric data is reconstructed tomographically.

In conclusion, LDA and L2F are suitable for studying low speed flows and PIV offers great potential for extending these methods to whole-field testing. At transonic speeds these techniques are still valid, although difficulties arise in seeding the flows at sufficient densities with mono-disperse particles. Also, sub-micron sized particles must be used at high speeds to ensure that they follow the flow accurately and therefore the

scattered signal is weak and hard to extract from noise. High enthalpy reactive flows are best analysed using diagnostics based on inelastic scattering. CARS, CSRS and LIF are all promising techniques that provide non-intrusive flow information when the combustion products contain molecules with suitable electronic transitions. In non-reactive compressible flows, holographic interferometry offers the greatest potential, because particle seeding is unnecessary and whole-field data is captured in minimal run-time. Although holographic interferometry offers many advantages over other diagnostics, the practical problems of data acquisition and processing must be addressed. These topics are covered in the remainder of this thesis, with an aim at producing a quantitative diagnostic that can be used in routine testing.

3 Holographic Interferometry

3.1 Principles of Holographic Interferometry

3.1.1 Principles of Holography

Aerodynamic information may be obtained by interrogating a flow with coherent wavefronts from a laser. To record the distortion of optical wavefronts after they have propagated through a flow of varying density, it is insufficient to utilise only the intensity of the transmitted light. In addition, the phase of the wavefronts, which determines the time they have taken to traverse the flow field, must also be recorded. As photographic materials only record intensity information, it is necessary to encode the phase of the light as an intensity modulation.

The time-averaged intensity of a light beam may be modulated by the phase of its wavefronts, if it is interfered with a second beam. Interference occurs provided the beams are mutually coherent and not orthogonally polarised. The phase and amplitude information encoded in this interference pattern may then be recorded by a photographic emulsion. This recording is known as a hologram and was first described in 1948 by Dennis Gabor [99, 100, 101]. Originally, the need for coherent radiation was prohibitive, as there were no light sources which were both highly coherent (monochromatic and point emitting) and of sufficient intensity to expose the high resolution, low speed, recording materials necessary for holography.

In 1960 the invention of the laser provided a convenient source of highly intense, quasi-monochromatic radiation. Due to the high resolution and large dynamic range of photographic recording materials, vast quantities of information, relating to both the shape and reflectivity of an object, could now be stored and reconstructed. The first high quality holograms of three-dimensional diffusely reflecting objects were made by Leith and Upatnieks in 1964 [102], after their pioneering work in the field of off-axis holography between 1961 and 1963 [103, 104].

The information encoded on a hologram may be manipulated by recording multiple exposures, to yield details about the topology or distortion of a surface. This interference of holographic images forms the basis of holographic interferometry and was first reported by Powell and Stetson in 1965 [105]. Double-exposure and real-time holographic interferometry were developed later that year [106, 107]. The extension of holographic interferometry to transparent phase objects, with the associated implications for wind tunnel testing, occurred in 1966 by Brooks *et al* [108]. These techniques have since been applied by Bryanston-Cross *et al* to numerous industrial flow diagnostic problems [82, 109].

The basic arrangement used to produce a hologram of a solid object, in this instance a diffusing screen, is shown in figure 14. Initially, the light from the laser is divided into two components by a beam splitter. The reflected and transmitted beams are subsequently expanded by positive or negative lenses, and pin-hole spatial filters may be used to improve the uniformity of the light. One of these components, which is known as the object beam, is arranged to illuminate the subject of the hologram. The second beam is used to directly illuminate a photographic plate, without scattering from intervening surfaces. This beam has wavefronts with a simple form (usually planar or spherical) so that they can be easily reproduced at a later stage and for this reason it is known as the reference beam.

When the optical system is correctly aligned, the scattered object wavefronts overlap the reference beam at the holographic plate. The object and reference beams are mutually coherent (temporally) at the recording plane, if their path lengths differ by less than the coherence length of the light. The coherent wavefronts interfere, provided they are not orthogonally polarised, and form a fine fringe pattern (fringe spacing is typically 1μm). Fringe visibility is high if the system is stable and the intensity ratio of the object and reference beams is correct (the optimum ratio depends on the illumination characteristics). Information about the shape and reflectivity of the object may be stored by recording this fringe pattern on a photographic plate. Silver halide emulsions contain extremely fine grains (typically 50nm in diameter), so that the high

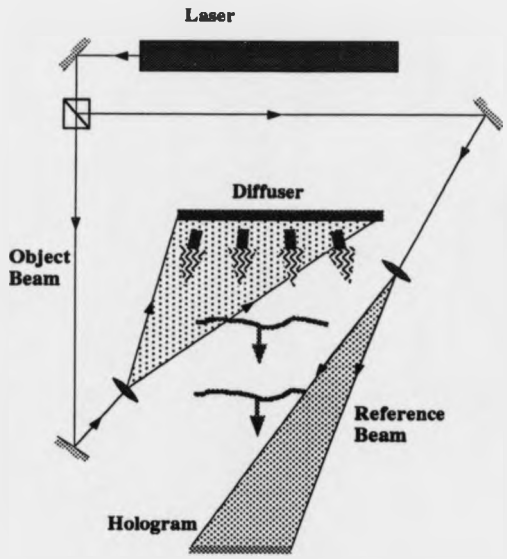


Figure 14: Optical system used to record a hologram of a diffusing screen

spatial frequencies present in fringe patterns can be resolved (>1000lpmm^{-1}). The reference beam is usually inclined to the object beam, to ensure that the holographic reconstruction is spatially separated from the directly transmitted light (known as an off-axis hologram).

The light scattered from the object forms a near-field (Fresnel) diffraction pattern in the plane of the hologram P(x_h, y_h). The amplitude distribution of this pattern may be described in complex notation by:

$$O_D(x_h, y_h)\exp[j\Phi(x_h, y_h)] \tag{36}$$

Where,

O_D - Amplitude of the diffracted object wave
Φ - phase of the diffraction pattern

The intensity distribution $I(x_h, y_h)$ of the pattern exposed on the holographic plate, by interference of the scattered object wave with a reference beam R, is given by the

coherent superposition of the two beams (assuming perfect coherence):

$$I(x_h, y_h) = \left| O_D(x_h, y_h) \exp[j\Phi(x_h, y_h)] + R \right|^2 \tag{37}$$

When the hologram is replaced in the optical system after photographic processing, the holographic image may be reconstructed by re-illuminating the recording with the original reference beam. To obtain a faithful reconstruction of the object, the reference beam wavefronts must be unaltered in shape between recording and replay of the hologram. In practice, this is simplified by using either: a collimated beam, with planar wavefronts; or a divergent beam, with spherical wavefronts. When planar or spherical wavefronts are used, the hologram is easily realigned and, if aberrations are neglected, only a change in image magnification occurs if the curvature of the wavefronts is altered.

When the reference beam reconstructs the hologram, the wavefronts are usually diffracted into three orders (others may occur if the recording is non-linear). These diffraction orders may be explained mathematically by expanding equation 37 and multiplying the result by the reconstructing wave R. The diffraction orders are physically separated if the reference beam is sufficiently inclined to the object wave. After expansion, the term of interest is given by:

$$RR^* O_D(x_h, y_h) \exp[j\Phi(x_h, y_h)] \tag{38}$$

The term R^* denotes the complex conjugate of the reference beam R. Physically this represents a wavefront of opposite curvature and equal amplitude to the original reference beam. When R and R^* are multiplied, the opposing curvatures cancel and inspection of expressions 36 and 38 shows that the reconstructed optical field is identical to within a multiplicative factor (image brightness) of the light scattered from the object. When this diffraction order is viewed, by illuminating the hologram with the original reference beam, the observer sees a virtual image which is visually identical to the original object (figure 15).

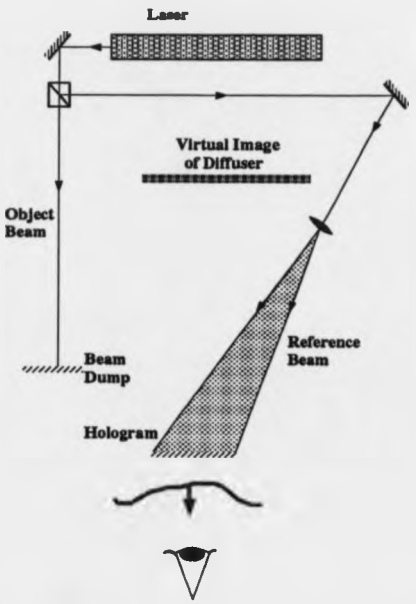


Figure 15: Reconstruction of a diffuse illumination hologram

3.1.2 Double-Exposure Holographic Interferometry

Double-exposure holographic interferometry, as the name implies, is the recording of two holograms on the same photographic plate. This technique may be used as a wind tunnel diagnostic, by taking an exposure of static or incompressible air in the test section, followed by one of a compressible flow. During reconstruction, the wavefronts from the two recordings are replayed simultaneously and an interference pattern is produced (figure 16).

If the flow retards or advances the transmitted wavefronts by an odd integer number of half wavelengths, the reconstructed object beams interfere destructively and the holographic image is darkened by an interference fringe. Conversely, when the light has been retarded by an integer number of wavelengths, the beams interfere constructively and a bright fringe is produced. When the entire flow region is considered, a cosinusoidal fringe pattern is formed which is modulated by variations in fringe visibility and beam uniformity.

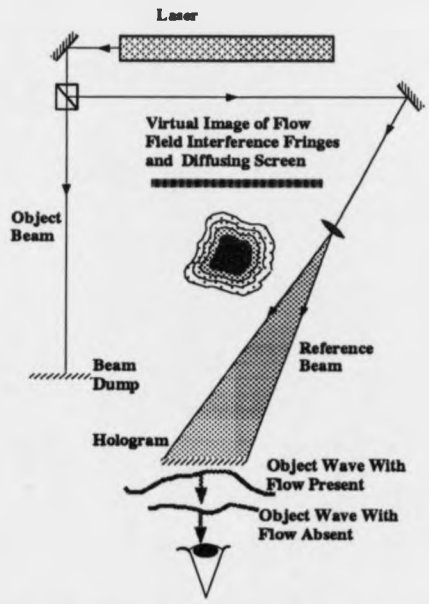


Figure 16: Reconstruction of a double-exposure hologram

The fringe pattern accurately maps out the optical path length changes imparted by a compressible air flow. If the density field is two-dimensional then the fringes define a set of iso-density contours. An attraction of this mode of holographic interferometry, is that the two consecutive exposures may be made in quick succession to reduce the effects of wind tunnel vibration.

3.1.3 Real-Time Holographic Interferometry

Real-time holographic interferometry is similar to the double-exposure technique, except that instead of generating both of the wavefronts holographically, a single holographic reconstruction is combined with light transmitted directly through the flow field. Hence, because one of the wavefronts is updateable, it is possible to visualise an animated fringe pattern.

Real-time interferometry is an extremely powerful technique as it enables an air flow to be studied dynamically, without the need for lengthy intermediate photo-

graphic processing. The evolution of the fringe pattern may be monitored in real-time with a sensitive TV camera, enabling an interactive study of flow fields [110, 111]. Unsteady features, existing for less than 1ms, may be observed if high speed video is used in conjunction with strobed or pulsed illumination.

Unfortunately, in its simplest form real-time holographic interferometry is restricted to systems that are interferometrically stable and therefore usually confined to a laboratory. Vibration may be tolerated if a servo-system is used to sense any movements of the reference hologram, so that it can be realigned to the infinite fringe condition [112]. Alternatively, real-time holographic recording materials may be used to continually update the hologram as movements occur. A review of suitable materials is presented in section 3.2.3.

The stability constraints may be relaxed if a common-path geometry is adopted, or if the object and reference beams are divided close to the hologram plane. Shearing interferometers are particularly stable, as the reference wave is derived from a sheared replica of the object wave. These systems produce fringe patterns that are related to the first spatial derivative of density and are therefore less suitable for quantitative flow analysis.

The radial shearing or point diffraction interferometer, developed by Smartt [113], enables the study of absolute density by spatially filtering flow information encoded on the sheared wavefronts. This technique was developed at NASA Ames to improve the data acquisition rates of routine holographic testing [114]. Although convenient, measurement errors may occur if the flow information has not been completely removed from the sheared wavefronts. Large errors are introduced when the optical system is aberrated, or if the flow field does not contain any regions of uniform density.

3.1.4 Differential Holographic Interferometry

Unlike conventional interferometry, holography separates the interfering wavefronts temporally rather than spatially. This important difference enables the production

of accurate interferograms, even when the test section windows are not of schlieren quality. Another advantage is that it is possible to perform differential, instead of absolute interferometry. In this instance, both holographic exposures are made with a compressible flow present. Consequently, if a steady-state flow field is examined, the reconstructed wavefronts interfere constructively and form a featureless image. If the flow field is temporally unstable, however, destructive interference occurs in the unsteady regions and features become visible. The interference pattern I, reconstructed from a double-exposure hologram recorded at times t_1 and $t_1 + \delta t$, is given by:

$$\begin{aligned} I = & \left| O_D \exp[j\Phi(t_1)] + O_D \exp[j\Phi(t_1 + \delta t)] \right|^2 \\ \approx & \; 2|O_D|^2 \left\{ 1 + \cos\left[\delta t \frac{d\Phi(t_1)}{dt} \right] \right\} \end{aligned} \tag{39}$$

The technique is known as differential holography, because of the time derivative in equation 39, and the interferometric features are strongest in the most unsteady regions of the flow. Unsteady flow features may be located in three-dimensions by differential holographic interferometry, but quantitative analysis of the density changes is limited. Figure 17 shows a sequence of differential interferograms, illustrating unsteadiness in the boundary layer covering the slotted liner of a transonic wind tunnel (see chapter 7 for details of the test). The flow was incident from the right of the picture at Mach 1.10 and the shock structures emanate from a cone-cylinder model located on the wind tunnel centre line (working section height 200mm). A circular collimated beam of diameter 200mm was used to illuminate the flow and the width of the field-of-view is approximately 130mm (magnification of 1.17). The wind tunnel liner is aligned to the top edge of the photographs and part of the flow is not illuminated due to the use of a circular beam.

The series of interferograms in figure 17 show that turbulence in the boundary layer occurs on two different scales. When a pulse separation of 4μs is used (top image), a number of small 'bursts' are seen that have a diameter of (1.5$\pm$0.5)mm. At Mach 1.10, a transient flow feature moves 1.5mm in 4μs and therefore these bursts may be smaller when they are created. Also, it is likely that diffusion causes them

to broaden as they propagate downstream. Although these features appear randomly distributed in space, they are only observed for pulse separations in the range $(4\pm1)\mu$s (0.25MHz bursting frequency). When the pulse separation is increased to 7μs (second image), the features have disappeared and the flow appears steady.

A black fringe is formed in differential holographic interferometry when the cosine argument of equation 39 is equal to π radians. Hence, from equation 39 and the holographic image (mean feature size of 1.5mm), it is possible to calculate that the density field is fluctuating at an average rate of $(1.5\pm0.8)\times10^5 \mathrm{kgm^{-3}s^{-1}}$. If the density is assumed to be uniform within the turbulent bursts, then this result shows that the perturbation of the density field is $(0.7\pm0.5)\mathrm{kgm^{-3}}$.

At longer pulse separations, density fluctuations occur on a larger scale. These features are first observed at a 10μs pulse separation (third image) and become more pronounced when the delay is increased to 20μs. When the pulse separation is increased further, these large scale density perturbations remain. This suggests that unlike the smaller bursts, these disturbances are generated at a range of frequencies. Although these features are an order of magnitude larger, the maximum phase shift is only about 2π radians, which indicates that the density changes in these features is much smaller than those in the turbulent bursts. The results obtained in this investigation were found to be independent of the point at which the shock wave impinged on the boundary layer.

Localised density perturbations are also observed in the turbulent boundary layer over an aerofoil section. This is illustrated by an image captured from a diffuse illumination hologram (figure 18) of a 3.5" chord NACA 0012 aerofoil (see chapter 6 for details). A pulse separation of 10μs was used, with a Mach number of 0.90 and a Reynolds number of approximately 10^6. Absolute interferograms were also recorded of this flow and these results showed that the turbulent bursts were generated after the point of shock induced boundary layer separation. In this image the features are viewed from an elevated position, which proves that they are localised effects and not generated by integration through the flow volume.

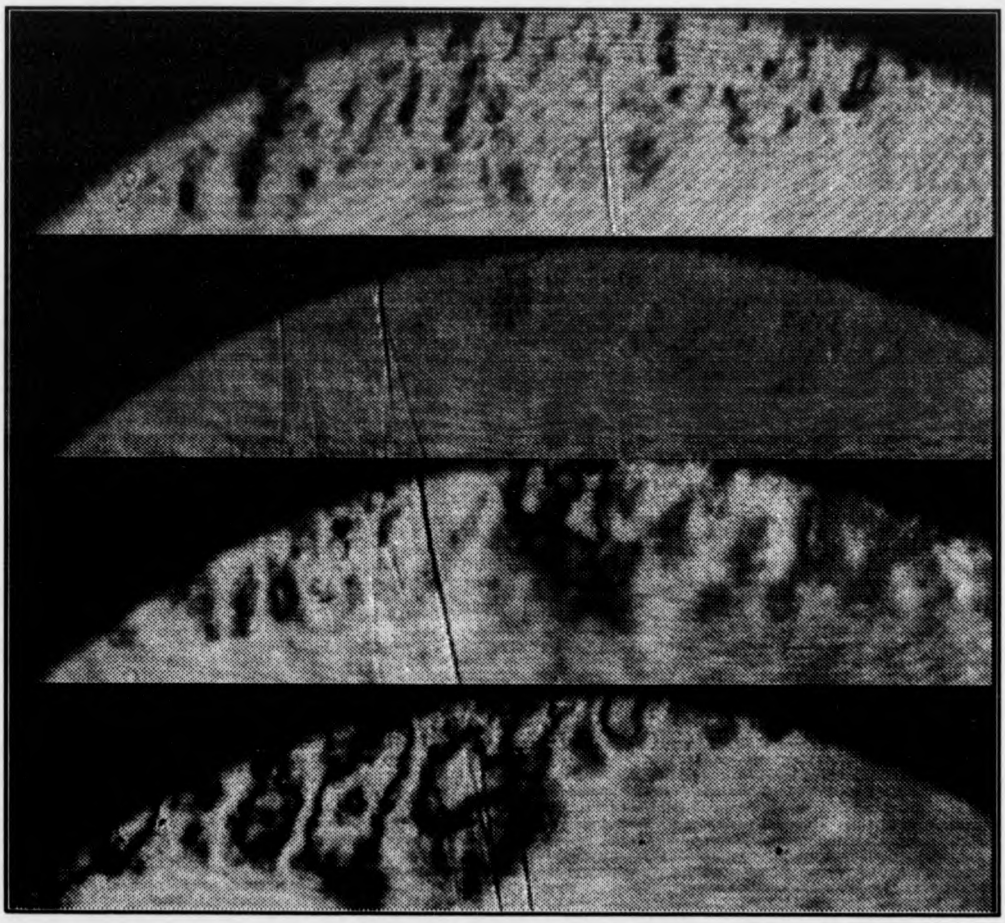


Figure 17: Study of side-wall boundary layer features using differential holographic interferometry

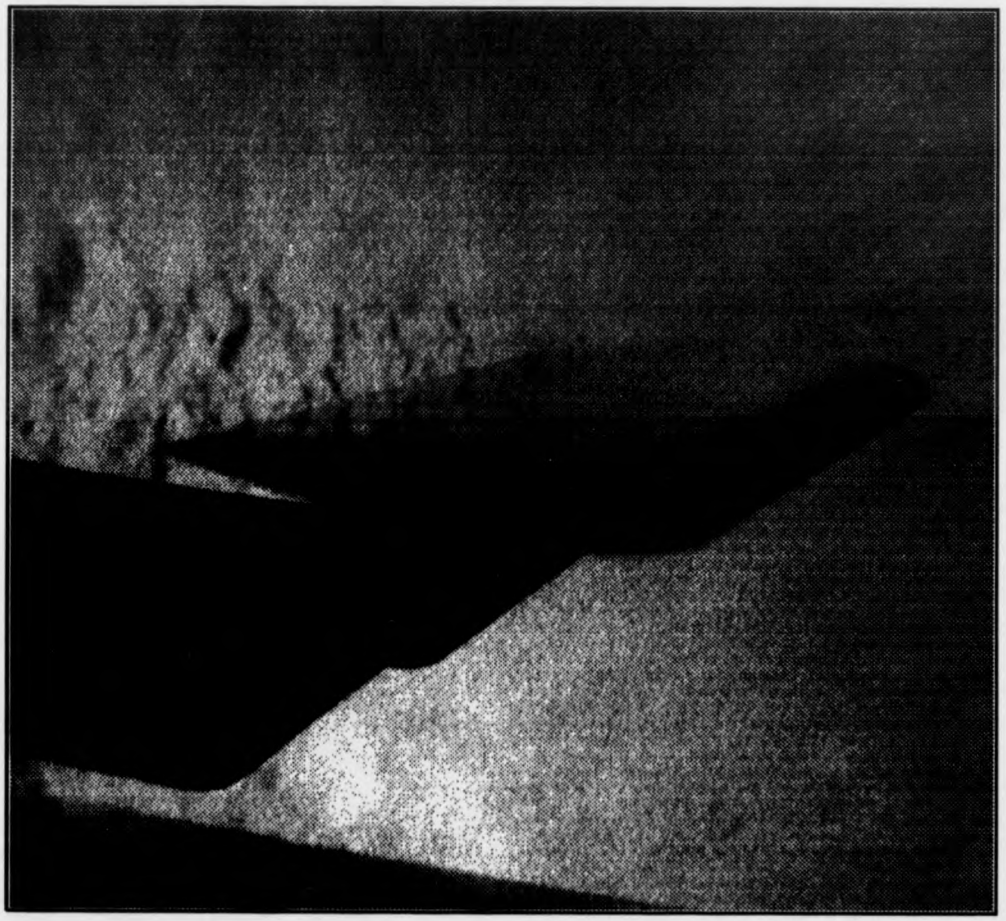


Figure 18: Formation of turbulent features in a region of separated flow

Differential holographic interferometry is a powerful technique for comparing two similar flow fields. For example, it may be possible to study the effect of drag reduction devices that perturb the flow between exposures. In shock tube experiments, where the air flows rapidly through the test section, differential holography enables recording of both 'flow-off' and 'flow-on' holograms within the same double pulse. Ruby lasers may be pulsed twice within a period of 1-800μs and therefore vibration induced error fringes are minimised.

Although differential holographic interferometry is useful for studying unsteady features, and in verifying the position of shock waves [115], it is desirable to determine the magnitude and time dependence of $\Phi(t_1)$ defined in equation 39. The example considered in this section demonstrates the feasibility of this analysis, provided that the phase shift generated by the flow features can be calculated accurately. Also, it is desirable to determine whether the phase changes are positive or negative. In practice these objectives may be achieved by recording the two exposures with separate reference beams. Rapid switching of the beams may be accomplished using electro-optic devices [116, 117].

3.1.5 Time-Averaged Holographic Interferometry

Time-averaged holographic interferometry involves recording a time integrated hologram of an unsteady flow, using a continuous wave laser and prolonged exposure time. If the instabilities have a characteristic frequency spectra, and the exposure-time exceeds the period of these fluctuations, then the features are visualised in the reconstructed interferogram. Decker has shown that if the reference beam is frequency modulated throughout the exposure, then the hologram can be 'tuned' to a particular flow feature [117]. Frequency modulation can therefore be used to make the hologram temporally selective, so that chosen features are seen with improved definition. This technique has been used in conjunction with real-time interferometry to study transonic flow over a flutter cascade [118]. Although potentially useful for studying turbulence, and other unsteady features, this technique has been used infrequently.

3.2 Practical Considerations

3.2.1 Pulsed Holography

When a hologram is recorded, path differences between the object and reference beams must remain constant to within $\lambda/8$ during the exposure, to ensure that the interference fringes remain stationary. This level of stability is difficult to achieve in an industrial environment, unless the two beams are phase-locked by a servo system, or follow equal paths. A simpler solution is to deliver the illumination in a short pulse ($< 1\mu s$), by Q-switching the laser, so that all vibrations are frozen.

Although pulsed lasers enable the recording of holograms in wind tunnel environments, vibration problems still occur when two holographic exposures are taken to form an interferogram. If the optics move between the exposures, the holographic reconstruction exhibits a set of fringes that 'contour' the movement. The required stability can be achieved, however, if the optical system is carefully designed to ensure that components are located at nodal points. Residual rigid body movements, that appear as linear fringes on the interferograms, may be removed during data processing (chapter 5).

Additional practical difficulties are imposed by the high peak optical powers used to record pulsed holograms. Whenever the power density exceeds $10^{11}\mathrm{Wcm}^{-2}$, for example at the focus of a lens, ionisation of the air may occur and the beam can be retro-reflected into the laser cavity [119]. This problem is particularly inconvenient because it prevents the use of pin-hole spatial filters that are needed to improve beam uniformity. If focussed power densities exceed the breakdown threshold, then the beam must be expanded by a negative lens. Spatial filtering may sometimes be performed if a relatively long focal length positive lens (e.g. 200mm), is used in conjunction with a ceramic or diamond pin-hole. If a foil pin-hole is used, there is a danger that if it becomes misaligned, the focussed beam may retro-reflect from the metallic vapour.

Other components which have to be used with care at high optical powers are aluminised mirrors, which suffer damage at typically $10^{6}\mathrm{Wcm}^{-2}$; and multi-element

components, which have to be air-spaced to prevent ablation of the intervening adhesive. Although there are numerous practical difficulties associated with pulsed holography, the use of specialised optical components, and careful system alignment, can ensure that results are of high quality and equipment is not damaged.

3.2.2 Choice of Laser

In controlled environments it is sometimes possible to use continuous wave lasers for holographic flow visualisation. The two most popular lasers for this application are the helium-neon (He-Ne) and the argon ion. The argon ion laser is approximately two orders of magnitude more powerful than the He-Ne and a typical water cooled system provides 5W of monochromatic light at a wavelength of 514nm or 488nm. For holographic applications an intra-cavity etalon is required, which reduces the power to approximately 2W.

In most instances, the flow field or environment is temporally unstable and it is necessary to use a mechanical or electro-optic shutter to strobe the beam [120]. High resolution silver halide holographic emulsions (e.g. Agfa 8E56HD) have a sensitivity of typically $25\mu Jcm^{-2}$. Hence, if 10% of the optical energy is retained after the beam has been over-expanded and spatially filtered to ensure uniformity, and the hologram area is $10cm^2$, the minimum exposure time that can be used is approximately 1ms. This duration is often too long to prevent blurring of the reconstructed interferogram by flow instabilities and vibration of the optical components. Exposure times as short as $1\mu s$ could be achieved if a lower resolution emulsion is used (e.g. Agfa 10E56) and the beam is expanded less. This may enable the recording of interferograms in relatively unstable environments, although it is likely that the reconstructed image quality would be poor.

The ruby laser is the most popular pulsed laser for holographic interferometry, because it can supply high peak powers in the visible wave band (694nm). Recently, the frequency doubled Nd:YAG laser has challenged the role of the ruby laser due to

its high repetition rate, especially for real-time holographic interferometry [121, 122]. However, the peak optical powers obtainable from Nd:YAG lasers are typically an order of magnitude lower than those from ruby lasers.

The most powerful Nd:YAG systems, that provide optical powers approaching that of ruby lasers, use unstable resonators in their design. These lasers produce high order transverse modes which generate wavefronts with non-uniform phase. This is undesirable because it results in the recording of inefficient holograms with poor fringe modulation. Consequently, ruby lasers are normally used when a large expanded beam is required, or if the illumination is diffuse and scattered into a large solid angle. In contrast, Nd:YAG lasers are preferred for pulsed real-time applications, where the illumination is specular. A detailed comparison of ruby and Nd:YAG lasers is given in a review by Pitlak and Page [123].

It is likely that diode pumped Nd:YAG systems will eventually supersede the ruby laser, due to their superior efficiency, coherence length and compactness (320mm×100mm×80mm for the laser head of a 0.1mJ/pulse 421QD II ADLAS laser). Alexandrite and carbon dioxide lasers also have great potential for pulsed holography [124]. At present, however, these lasers and the supporting hardware (e.g. infra-red recording materials) are poorly developed for holographic applications. For high speed schlieren or shadowgraph, metal vapour lasers, that operate at kilohertz repetition frequencies, are becoming popular. Unfortunately, these lasers have poor coherence and therefore are unsuitable for holographic interferometry unless a common path configuration is adopted.

The HLS range of ruby lasers, produced by Lumonics (formerly known as JK Lasers), are the most popular in the UK for pulsed holography. These lasers are dedicated to holographic applications and may be purchased as a 'holographic camera'. In this configuration (figure 19), the object and reference beams are divided and path matched within the laser head. Also, provision is made for expanding the reference beam onto a holographic plate holder positioned on the front of the laser. Most of the HLS laser series have a facility for opening the Pockel cell Q-switch twice during

the discharge cycle. Careful control of the timing circuitry enables the production of two closely matched pulses, with an adjustable delay of 1-800μs between them. Individual pulses may be obtained every ten seconds and customised systems have been designed by Lumonics to enable the flash lamps to be fired in quick succession (delay of between one millisecond to ten seconds).

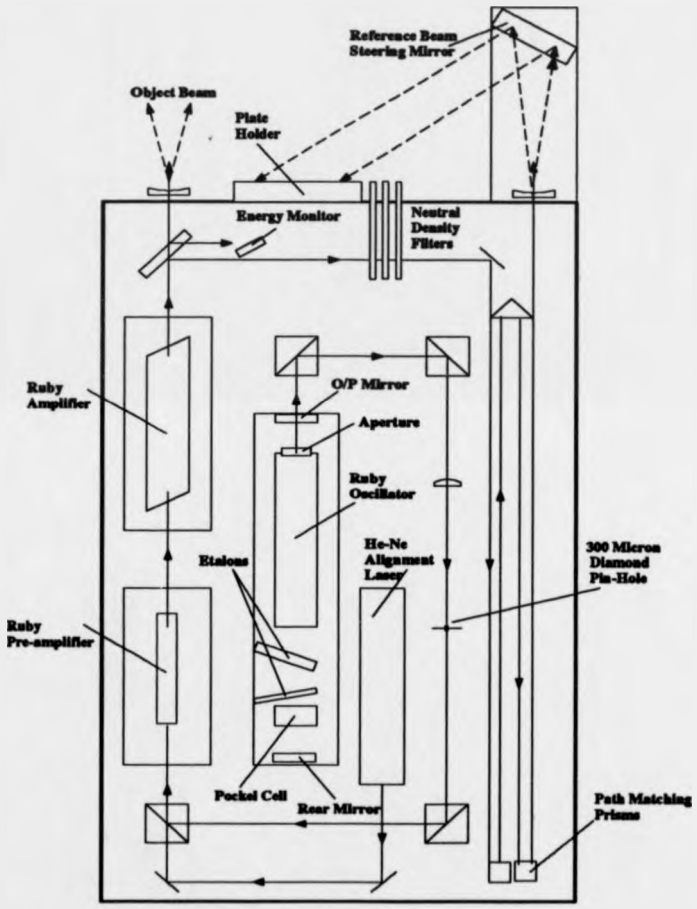


Figure 19: Ruby laser holographic camera

As well as enabling holograms to be made in vibration prone conditions, pulsed lasers supply optical powers far in excess of those produced by normal ambient lighting [124]. Therefore, it is possible to record holograms in all lighting conditions encountered in industry, if ambient light is suppressed by spectral filtering and a high speed

shutter.

3.2.3 Recording Materials

The interference of two plane waves, of wavelength λ and angular separation 2θ, produces a set of fringes (figure 20) which have a spacing d_f given by:

$$d_f = \frac{\lambda}{2\sin\theta} \tag{40}$$

If the reference wave is inclined at an angle of 60° to the object wave, the fringe spacing is equal to the wavelength of the laser (694nm for a ruby laser). A high resolution photographic material is required to record this rapidly undulating intensity distribution. Silver halide emulsions are usually chosen for holographic interferometry, because they have a resolution in the range 1500-4000lpmm^{-1} and are comparatively sensitive (an energy density of 3-10μJcm^{-2} is required for a suitable exposure).

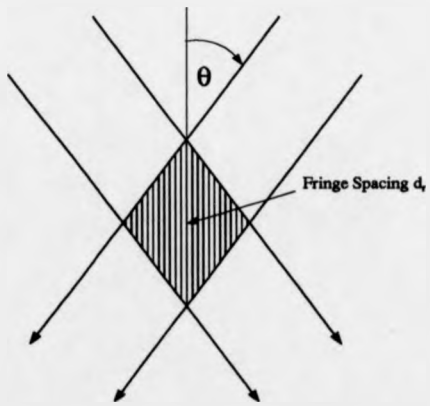


Figure 20: The formation of fringes by the interference of two plane waves

The optical properties of silver halide are not ideal, because the developed silver grains are typically 50nm in diameter and generate Rayleigh scatter. Consequently, diffraction efficiency and contrast of the reconstructed image are reduced. Scatter is prohibitive if the holographic element is being used in a visual system, for example

a head-up display (HUD), and in these circumstances dichromated gelatin [125, 126, 127], photoresist [128] and photopolymer [129] are preferred. Unfortunately, these 'continuous' materials are more difficult to work with and are generally less sensitive. In a holographic interferometer scatter is less detrimental, because a lower resolution is required and the images are digitally post-processed. Consequently, silver halide was chosen for developing the prototype systems described in this thesis.

Recently, real-time holographic recording materials have been developed so that the intermediate photographic processing may be omitted. The most promising substances are thermoplastics [128] and photorefractive crystals [130]. At present, these produce noisier reconstructed interferograms than silver halide and are therefore topics of current research. Another promising technology, which is maturing rapidly with the availability of inexpensive framestores and CCD cameras, is electronic speckle pattern interferometry (ESPI) [131, 132]. This technique uses a TV camera and framestore to record an on-axis image plane hologram, and the fringe patterns are generated by differencing video frames. Although convenient, the speckle patterns used to encode the holograms are coarse, and therefore the resolution of the interferogram is poor compared to conventional holography. Image quality is likely to improve as larger format cameras are implemented.

3.3 Potential Benefits and Disadvantages of Holographic Interferometry

To justify research into holographic interferometry it is important that the technique offers significant advantages over other flow diagnostics. The advantages of holographic interferometry may be summarised as follows:

- The technique is non-intrusive, whole-field, and does not require the use of seeding particles or tracers.
- Recording is effectively instantaneous and therefore unsteady flow features may be studied and run-time is minimised.
- Interferometry is sensitive to density and not its spatial derivatives (with the exception of shearing interferometers).
- Holography enables the recording of interferometric data without the need for schlieren quality optics.
- Shadowgraph, schlieren and interferometric images might be obtained from a single hologram (schlieren images require the use of high quality optics).
- Quantitative density measurements can be made from the reconstructed interferograms.
- Differential holographic interferometry may be used to study the temporal properties of unsteady flows.
- Holographic interferometry may be extended to three-dimensional density measurement, if multi-directional illumination is used in conjunction with tomographic data reconstruction.
- Flow data may be extracted accurately by phase sensitive methods (e.g. heterodyning or phase shifting).

- Pulsed lasers enable the recording of holograms in environments containing vibration and intense ambient lighting.

Although these advantages are significant, holographic interferometry suffers from a number of drawbacks that must be addressed if the method is to be used routinely. These disadvantages may be summarised as follows:

- Data extraction is difficult for three-dimensional density fields because the measurements are integrated through the flow. Computer aided Tomography is required to reconstruct projected views and aliasing errors are produced if the data is sparse or unevenly sampled.
- The dynamic range is small, because at a fixed wavelength there is only a limited range of fringe densities that can be processed by a digital image processing system.
- The accuracy of holographic measurements is compromised by air currents and vibration.
- Holographic systems are non-portable and usually customised to each specific wind tunnel facility.
- The lasers used for holography are bulky and prone to misalignment.
- Conventional holography uses inconvenient 'wet' photographic processing.
- Pulsed holography, although eliminating single exposure stability constraints, requires the use of optics that can withstand high peak powers.
- Diffraction and speckle noise can be severe for holographic systems because of the high coherence of the light.
- The optical system is bulky unless the beam paths are folded.

The purpose of the remainder of this thesis is to address these problems that have prevented the widespread use of holographic flow diagnostics.

3.4 Illumination, Imaging and Recording Configurations

3.4.1 Holographic Interferometry using Diffuse Illumination

The simplest method of visualising a compressible flow holographically is to make a double-exposure hologram of a diffusing screen, with the flow passed between the observer and the diffuser during the second exposure (section 3.1.2). When the hologram is reconstructed, a fringe pattern appears localised in front of the screen which exhibits the parallax of a three-dimensional object. Initially, it seems that this is a useful method for studying flow features in three-dimensions by simply focusing on a plane of interest. Unfortunately this assumption is incorrect, and this chapter shows that the effects of fringe localisation and phase integration must be considered.

Whenever a laser beam or other coherent source is shone onto a rough surface, the spatially coherent wavefronts are scattered randomly and interfere to form a grainy speckle pattern. The intensity distribution of these speckles is governed by exponential statistics [133] and the mean speckle diameter d_s is given approximately by equation 41. The depth-of-field of an imaging system is proportional to the square of the f-number. Consequently, when diffuse illumination holograms are imaged onto a detector to enable data extraction, there is a rapid trade-off between speckle noise and fringe contrast.

$$d_s \approx 1.5 f^{\#} \lambda \tag{41}$$

Where $f^{\#}$ is the f-number of the imaging system

Diffuse illumination double-exposure holograms of flow fields, also exhibit a phenomena known as fringe localisation. When this occurs, the fringes appear localised in space, even though they are a product of a delocalised path integration through the flow. Fringe localisation may result in the misinterpretation of interferograms, because the position of best focus of a flow feature does not necessarily correspond to its true location. For example, if two spatially separated shocks of equal strength are located

along the line-of-sight, then the holographic interference fringe will localise midway between them. Consequently, large positional errors may be introduced when locating the features if they are not viewed at the correct angle. Further complications, such as spatial decorrelation, occur when studying turbo-machinery, due to rotation of the flow or reflective background between exposures [115, 82]. These problems mean that the three-dimensional qualities of a diffuse illumination hologram cannot be fully exploited, unless the interferometric data is gathered from multiple directions and reconstructed tomographically (chapter 8).

Fringe localisation may be understood by considering the propagation of a cone of rays through the holographic interferometer and human visual system. The trajectory of a ray through a refractive index distribution $n(x, y, z)$ may be derived from *'Fermat's principle of least time'* as shown by Born and Wolf [75]. A vector differential equation defining this path is given by equation 28 and is restated here for convenience.

$$\frac{d}{ds}\left[n(x,y,z)\frac{d\underline{r}}{ds}\right] = \nabla n(x,y,z) \tag{42}$$

The analysis of fringe localisation is simplified if the medium is assumed to be 'refractionless' (known as a phase object). This approximation is extreme, but remains valid for weakly compressible transonic flows, or regions of supersonic flow that are devoid of shocks and large density gradients. Refraction is strong in the boundary layer and therefore this limits the capture of information close to a wind tunnel model's surface. An analysis of fringe localisation in the presence of refraction has been made by Schumann [134]. When the refractive index is only slowly varying with position, equation 28 reduces to:

$$n\frac{d^2\underline{r}}{ds^2} = 0 \tag{43}$$

The solution of this equation is a family of linear ray trajectories:

$$\underline{r}(x,y,z) = \underline{r_0}(x_0,y_0,z_0) + S\hat{\underline{k}} \tag{44}$$

Where,

$\underline{r_0}(x_0,y_0,z_0)$ - fixed vector from the origin to a point on the ray

$\hat{\underline{k}}$ - unit vector along the ray path

S - scalar

A set-up for double-exposure diffuse illumination holography is illustrated in figure 21 and the following analysis is based on that by Vest [135]. For convenience, the diffuser is located in the plane $z = 0$ and the line of sight through the air flow is centred at a point $(x_d, y_d, 0)$ on the diffuser. The vector from the origin to this point on the diffuser is denoted by $\underline{r_0}$ and the vector along the ray path is denoted by $\hat{\underline{k}}$ in accordance with equation 44. Fringes localise in space, because the lens has a finite collecting area which samples a pencil of rays that have taken different routes through the flow. If the optical path lengths of these rays are significantly different from the central ray $\hat{\underline{k}}$, then there is not complete constructive interference when they are recombined by the imaging system. Consequently, fringe contrast is poor, unless the limiting aperture is stopped down to reduce the spread of angles propagating through the system. The fringes have best visibility, and therefore appear localised, when the variation in optical path length over the aperture is minimised.

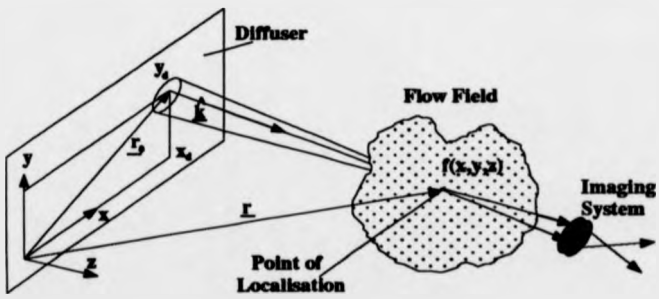


Figure 21: Fringe localisation in diffuse illumination holographic interferometry

The optical path length change ΔL, between the two exposures of a double-exposure hologram, is given by the following line integral:

$$\Delta L = \int_l \Delta n(x,y,z) ds \tag{45}$$

When a differential holographic interferogram is recorded, $\Delta n(x,y,z)$ describes the differential refractive index distribution $n(x,y,z,t+\Delta t) - n(x,y,z,t)$, where Δt is the time between exposures. For an absolute double-exposure hologram (flow-on and flow-off), this expression is simplified to $n(x,y,z) - n_0$, where n_0 is the refractive index of static air. To simplify the notation, the change in refractive index is denoted by $f(x,y,z)$.

The fringe visibility is maximised when the optical path length variation $\mathrm{d}(\Delta L)$ across the imaging aperture is minimised. As the movement across the aperture is associated with a complimentary movement $\delta x_d, \delta y_d$ across the diffuser, the condition of localisation may be defined by the following minimisation of ΔL:

$$d(\Delta L) = \frac{\partial(\Delta L)}{\partial x_d} dx_d + \frac{\partial(\Delta L)}{\partial y_d} dy_d = 0 \tag{46}$$

The partial differentials may be evaluated by considering the projection of the ray's position vector $\underline{r}(x,y,z)$ onto the gradient of the refractive index change $\nabla f(x,y,z)$. Hence, the differential with respect to x_d (and similarly for y_d) is given as:

$$\frac{\partial(\Delta L)}{\partial x_d} = \int_{-\infty}^{\infty} \nabla f \cdot \left(\frac{\partial \underline{r}}{\partial x_d} \right) ds \tag{47}$$

The solution of equation 46 for a particular gradient function ∇f defines the surface of fringe localisation. Vest shows [135] that when the aperture has similar dimensions in the x and y directions (circular or square), equation 46 reduces to two simple equations defining the z co-ordinate of localisation:

$$z_l = \frac{\int(\partial f/\partial x)z dz}{\int(\partial f/\partial x)dz} \tag{48}$$

$$z_l = \frac{\int(\partial f/\partial y)z dz}{\int(\partial f/\partial y)dz} \tag{49}$$

Where the integration limits are over the extent of the flow field.

These equations show that the fringes localise at the centroid of the transverse gradient in the perturbed refractive index field.

If the change in density between the two exposures is highly three-dimensional then the surface of localisation becomes extremely convoluted. This is illustrated in figure 22, which shows three images captured from a diffuse illumination hologram of transonic flow around a NACA 0012 aerofoil (see chapter 6 for details of the test). Although the flow is two-dimensional, the symmetry is destroyed by viewing the flow from an upstream position. The first image was recorded with the camera focussed on the end of the model butted against the near-side test section window. The fringes around the separated boundary layer region, in the middle of the model, are defocussed and low in contrast. In the second image, where the focal plane is centred in the middle of the flow, the fringe contrast is higher in this region, although it has deteriorated for planes closer to the observer. Finally, the third image was recorded with the camera focussed on the far-side test section window and the fringes are badly defocussed, except towards the rear of the model.

These results demonstrate that a diffuse illumination holographic interferometer must have a large depth-of-field to ensure that all of the fringes viewed by the imaging sensor are sharply focussed. Unfortunately, this requires the use of a small aperture stop (lens of large f-number), which reduces the level of illumination and increases speckle size (equation 41). This is illustrated by an interferogram shown in figure 23, that was recorded with an f/8 aperture stop, opposed to the f/0.95 lens used to record the images in figure 22. Hence, it is difficult to improve the depth-of-field of the interferogram without incurring a loss in image resolution. Schemes to reduce

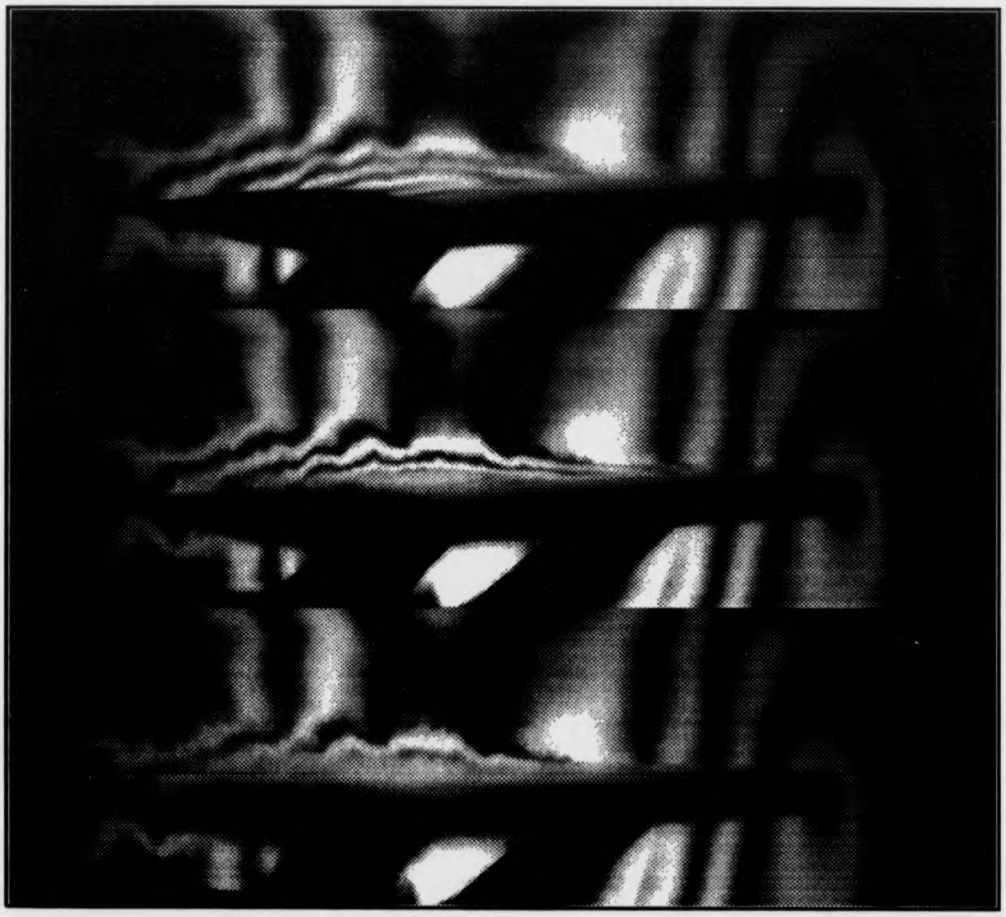


Figure 22: Fringe localisation in a transonic flow

speckle noise by image averaging have been devised (summarised by Dainty [136]), however these methods use moving diffusers and apertures [137], which complicate data extraction.

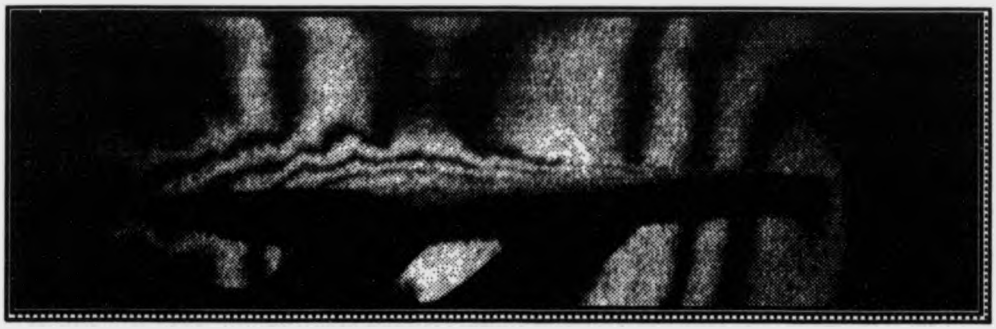


Figure 23: The production of speckle noise when the f-number is increased

In conclusion, although diffuse illumination holographic interferometry is simple to apply, and robust to macroscopic noise, it is not conducive to the capture of high quality interferograms that can be further analysed by image processing. The technique is useful for qualitative inspection of complicated three-dimensional flow field holograms; although caution must be exercised when trying to accurately locate particular flow features.

3.4.2 Holographic Interferometry using Specular Illumination

Section 3.4.1 has demonstrated that interferograms recorded in diffuse illumination are unsuitable for automatic digital processing, because image contrast is poor due to fringe localisation, and resolution is reduced by speckle noise. To eliminate fringe localisation, the rays must follow a one-to-one mapping onto the image plane, rather than the many-to-one mapping that occurs in a diffuse system. This may be achieved using collimated specular illumination, so that a point in the flow has only a single ray direction propagating through it.

Highly collimated specular illumination may be produced by placing a coherent point source at the focus of a spherical or parabolic mirror. To interrogate the flow in a wind tunnel, these rays are transmitted through the working section at right angles to the longitudinal axis and preconverged by a second mirror onto an imaging lens (figure 24). An image of the working section can then be recorded on a holographic plate, with a reference beam superimposed so that the phase information of the transmitted light is encoded as an intensity modulation. This system was originally proposed by Trolinger in 1969 [138], and is essentially a schlieren system with the knife edge removed and a reference beam added. Bryanston-Cross *et al* have subsequently used similar configurations in harsh engineering environments [139, 84, 140].

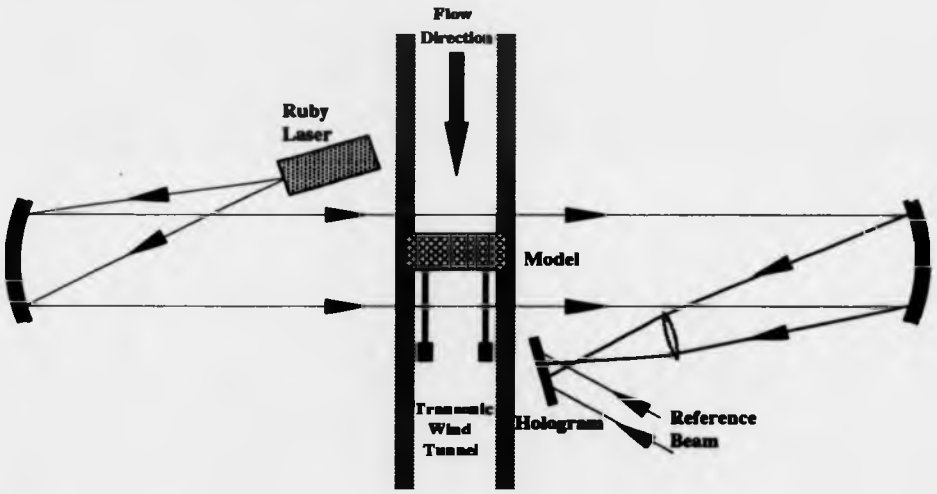


Figure 24: Conversion of a schlieren system into a holographic interferometer

Interferometers which use specular illumination produce delocalised fringe patterns and speckle noise is substantially reduced because there are no optically rough surfaces scattering the wavefronts. In addition, all of the light is retained within the system and therefore a specular configuration is typically three orders of magnitude more efficient than a diffuse geometry. A disadvantage of this approach is that because the rays propagate in only a single direction, the system does not have built-in redundancy and therefore it is prone to diffraction noise from dust or scratches on the optics.

Diffraction also occurs around the silhouette of the wind tunnel model and, because long beam paths are required to reduce optical aberrations, air currents can result in the formation of 'cloud-like' noise fringes.

The use of imaging optics solves many of the disadvantages associated with specular illumination holography. Diffraction noise can be almost entirely eliminated by imaging the centre of the working section onto the holographic plate, so that any imperfections on the optics or wind tunnel windows are defocused. The use of an image plane configuration is also advantageous during reconstruction, because the coherence requirements of the light source are reduced.

During the reconstruction of image plane holograms, optical aberrations are minimised when the replay wavefront is identical in shape and wavelength to the reference beam used during recording. In practice, the hologram has usually been produced with a pulsed laser and therefore a continuous wave laser of identical wavelength must be found. Due to the principals of laser operation, an exact match in wavelengths cannot be achieved unless a tunable laser is used. However, these complicated systems are expensive and produce speckle noise because of their high coherence.

An attractive alternative to laser illumination is to use a white light source to reconstruct the hologram. With three-dimensional diffuse illumination holograms this is not feasible because of the dispersive properties of the holographic fringe pattern, and the limited spatial coherence of the source. In contrast, an image plane hologram recorded in specular illumination, may be reconstructed in white light. Chromatic aberrations cause significant blurring around the periphery of the interferogram, but this may be alleviated by spectral filtering, or by placing a slit in the back focal plane of the hologram.

Figure 25 shows a diagram of the optical system used to reconstruct some of the holograms recorded during the wind tunnel tests described in chapters 6 and 7. The reconstruction wavelength was chosen to maximise the product of the radiance of the source with the sensitivity of the camera.

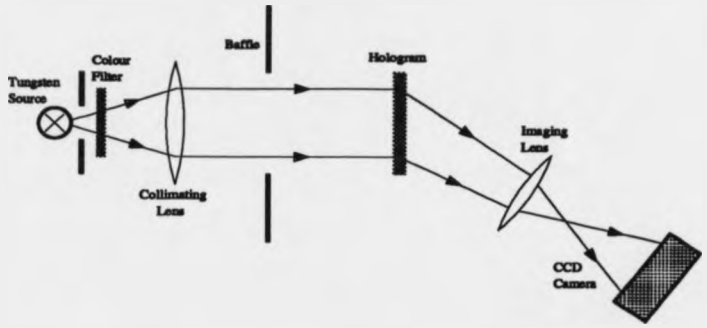


Figure 25: White light reconstruction system

The advantages of a white light reconstruction system, in the context of production wind tunnel testing, are as follows:

- The reconstruction system does not require the use of an expensive laser.
- Laser safety problems are avoided.
- The system is less prone to coherent noise which is prevalent in a dusty wind tunnel environment.

The associated disadvantages are:

- The reconstruction system is more difficult to align because the beams are not so clearly defined.
- The two foci from a dual reference beam hologram are difficult to separate in white light, unless a high carrier frequency is used (see chapter 5 for details). This is inconvenient because it restricts optical pre-processing of the interferograms prior to fringe analysis.

In summary, a white light reconstruction system is preferred for use in a wind tunnel environment, where coherent noise levels are high and laser safety is harder to control. When a laboratory is available, laser illumination is advantageous because it enables optical pre-processing of the images.

3.4.3 Multiplexed Specular Illumination

Sections 3.4.1 and 3.4.2 demonstrate that interferograms recorded with specular illumination exhibit higher fringe contrast and lower noise, than those produced using diffuse illumination. Also, diffraction noise is suppressed if the centre-line of the wind tunnel is imaged onto the hologram. Unfortunately, specular illumination image plane holograms are inherently two-dimensional and the three-dimensional qualities of diffuse illumination holography are lost. The benefits of both recording geometries may be obtained if a holographic optical elment (HOE) is used to generate a discrete fan of specular beams.

If transmission HOE's are used to replace the bulk schlieren optics in a specular illumination holographic interferometer, the resulting design is more compact, and less susceptible to vibration and air currents. HOE's may be superimposed during recording and therefore a single element may be used to generate multiple wavefronts from a single input beam. Interferometric systems encorporating HOE's in their design have been developed for both flow visualisation [135] and metrology [141]. Promising results have been obtained with elements recorded in silver halide and it is likely that further improvements would be obtained by utilising DCG [142, 143], photopolymer [144], or other continuous materials. In particular, diffraction efficiencies approaching theoretical limits could be achieved and Rayleigh scatter would be reduced. A review of the latest developments in HOE design, production, and implementation, is given by Pappu [145, 146].

The transmission HOE used to multiplex the input beam to a wind tunnel, may be formed by interfering a collimated or spherically divergent reference beam, with a 'fan' of collimated object beams. The intensity of the orders diffracted from the resulting hologram may be estimated using transparency theory. This approach is valid for thin gratings and was originally proposed by Raman and Nath in 1935 [147]. Experimental studies in holography have shown that the analysis is applicable provided that the grating does not exhibit strong volume properties. After the hologram is developed, the amplitude transmittance $t(x, y)$ of the grating is given by:

$$t(x,y) = t_0 - \beta T\,|(\mathbf{O_1} + ... + \mathbf{O_n}) + \mathbf{R}|^2 \,\mathrm{rect}\left(\frac{x}{w_x}\right) \mathrm{rect}\left(\frac{y}{h_y}\right) \qquad (50)$$

Where bold font indicates that the terms are complex quantities and the other quantities are defined by:

- x, y - co-ordinates in the plane of the grating
- t_0 - uniform bias transmittance
- β - gradient of the transmittance/exposure curve
- T - exposure time
- $\mathbf{O_i}$ - amplitude of the *ith* object wave
- $\mathbf{R}$ - amplitude of the reference wave
- w_x - width of the hologram in the x direction (assumed rectangular)
- h_y - height of the hologram in the y direction
- n - total number of object waves

with the function rect(z) defined by:

$$\mathrm{rect}(z) = \begin{cases} 1 & \text{if } |z| \leq 1/2 \\ 0 & \text{otherwise} \end{cases}$$

Equation 50 may be expanded to determine the separate interference terms recorded on the grating. If several beams are required from a single HOE then the unmodulated bias term contained in equation 50 increases and the overall diffraction efficiency is reduced. This problem may be resolved by bleaching the element so that the unmodulated exposure does not absorb any of the incident light. In this instance, the transmittance function is complex with unity amplitude:

$$t(x,y) = \exp\left\{ jk_\phi \left[t_0 - \beta T\,|(\mathbf{O_1} + ... + \mathbf{O_n}) + \mathbf{R}|^2 \,\mathrm{rect}\left(\frac{x}{w_x}\right) \mathrm{rect}\left(\frac{y}{h_y}\right) \right] \right\} \qquad (51)$$

Where k_ϕ is related to the modulation depth of the grating.

Equation 51 may be expanded as a series of Bessel functions to determine the strength and interaction of the diffraction orders. A rigorous analysis of diffraction by thin and thick bleached HOE's is given in a book by Syms [148].

In the wind tunnel tests described in this thesis, the number of beams interrogating the flow was limited and therefore multiple elements were deemed unnecessary for the prototype optical systems. A more detailed configuration, encorporating multiplexed elements, will be designed once a feasibility study into the tomographic reconstruction of interferometric data has determined how many separate optical channels are required.

3.4.4 Separate Image Encoding by Dual Hologram and Twin Reference Beam Methods

Single reference beam double-exposure holographic interferometry, suffers from two major limitations. Firstly, it is common for rigid body movements to occur between the exposures, resulting in the superposition of randomly orientated linear fringes on the interferogram. Secondly, infinite fringe patterns cannot be analysed accurately, because the images are degraded by amplitude noise and the phase relationship between adjacent fringes is unknown. These limitations may be overcome if the individual exposures are encoded uniquely, so that the holographic images may be replayed or modulated separately. Separate image encoding may be achieved by recording the holograms on individual plates, or by using two angularly spaced reference beams.

Sandwich holography was invented by Abramson in 1974 as a means of recording the separate object waves on a pair of closely spaced holographic plates [149, 150, 151]. Initially, the first exposure of static air is recorded on a 'sandwich' of two holographic plates (emulsion facing the test section). An index matching fluid (e.g. xylene or white spirit) is trapped between the plates to ensure that they are in intimate contact. After this first exposure, the holograms are removed from the holder and an exposure of the compressible flow is made on a second sandwich.

During replay, the front plate of the first sandwich is combined with the rear plate of the second sandwich, and the composite hologram is viewed in a conventional reconstruction system. Fringes that appear on the reconstructed image, due to a tilting

of the optical system between exposures, may be removed by simply rotating the sandwich relative to the spherical reconstruction wave. In practice, sandwich holography is time-consuming and substantial practical experience is required to apply the technique without damaging the holograms. Sandwich holography must be considered, however, if a continuous wind tunnel is being used. A detailed guide for using the technique is outlined in a paper by Abramson [152].

Dual plate holography can also be used to remove the finite fringes produced by rigid body motion. The technique is similar to sandwich holography, but in this instance a single hologram is made of the quiescent state and then all subsequent holograms are made of the perturbed density field. Unlike sandwich holography, the reconstructed holograms are not placed in physical contact, but are realigned in the system using a kinematic plate holder.

The convenience of not having to use index matching fluid between the plates, and not requiring hologram pairs for each exposure, is offset against the need for submicron repositioning accuracy. When diffuse illumination holograms are made, the positional accuracy must be maintained for all translational and rotational degrees of freedom. The positional tolerances are partially relaxed when specular illumination is used, because a single wavefront extends over the whole aperture, and only rotational alignment is necessary. Numerous researchers have used this method effectively [153, 154] and Trolinger has combined the technique with thermoplastic recording materials [155]. A limitation of dual plate holography is that it can only remove linear or parabolic phase errors and it is often difficult to judge whether this has been successful if the flow generates fringes that extend over the whole field-of-view.

An alternative method to using two holographic plates is to use two different reference beams to record the exposures. This approach was suggested by Ballard in 1968 [156] and is an elegant solution to a difficult practical problem. The object wave $\mathbf{O_1}$ passing through the quiescent test section, is recorded with a reference beam $\mathbf{R_1}$, and the second exposure of the perturbed field $\mathbf{O_2}$ is recorded with a reference beam $\mathbf{R_2}$ inclined to the first. If the recording is linear, the amplitude transmittance of the

developed hologram is given by:

$$t(x,y) = t_0(x,y) - \beta(\mathbf{R_1}^*\mathbf{O_1} + \mathbf{R_1}\mathbf{O_1}^* + \mathbf{R_2}^*\mathbf{O_2} + \mathbf{R_2}\mathbf{O_2}^*) \quad (52)$$

Where t_0 and β are constants that depend on the photographic process.

The identity and angular spacing of the wavefronts diffracted from the hologram may be determined by multiplying equation 52 by the reconstruction waves ($\mathbf{R_1} + \mathbf{R_2}$). Eight separate terms encorporating the phase and amplitude of the object wave are reconstructed. To select one of these diffraction orders it is necessary to adjust the reference beam angles during hologram production, so that the diffracted wavefronts do not overlap. Dual reference beam techniques may be used in both diffuse and specular systems. However, the results in this report were obtained with a specular arrangement, to avoid fringe localisation (section 3.4.1). A detailed analysis of the diffraction orders generated in a diffuse illumination system is given by Dändliker *et al* [157].

When both of the object waves are reconstructed simultaneously, linear (tilt) fringes may be introduced or removed from the interferogram by carefully changing the angle of one of the two replay beams (provided that the holographic grating does not exhibit strong volume properties). Also, the fringes may be scanned over a detector by phase modulating one of the reference beams in a technique known as heterodyne detection (see section 4.2.2).

Research has shown that the angular re-alignment of the developed hologram is critical when the angle between the two reference beams is more than a few degrees [158]. Instead it is better to have a small inter-beam angle ($\sim 1°$) so that the angular tolerance is reduced (re-alignment accuracy is reduced from 0.01° to 10°). When the angle is small, and a diffuse illumination hologram is made, the individual diffraction orders are superimposed and the observed fringe pattern is noisy and has poor contrast. This problem may be alleviated by removing unwanted images by optical spatial filtering or by digital post processing.

A better approach is to use a specular illumination holographic interferometer, with a narrow angle between the two reference beams. The resulting interferogram has high contrast and is amenable to Fourier transform fringe analysis (chapter 5). Also the hologram can be reconstructed in white light to eliminate laser speckle. In unstable environments, finite fringe interferograms can be recorded with a pulsed laser, with an electro-optic or mechanical shutter used to rapidly switch the reference beams between exposures. A theoretical and practical study of pulsed dual reference beam holographic interferometry is given by Lai [159].

4 Fringe Analysis

4.1 Introduction

4.1.1 Origination of Fringe Analysis

When a double-exposure hologram of a flow field is replayed with a replica of the original reference beam, the object waves that passed through the quiescent and perturbed density fields are reconstructed simultaneously. These wavefronts are mutually coherent and interfere to produce a sinusoidal intensity distribution, which encodes the phase difference between the two beams. Figure 26 shows a fringe pattern reconstructed from a double-exposure hologram of a vertical glass column venting carbon dioxide (CO_2) to the atmosphere. The first hologram was taken shortly after the tube had been opened and the second exposure was recorded with the CO_2 displaced by air. The object wave that propagates through the regions containing CO_2 is delayed with respect to the wavefront travelling through air, and therefore a fringe pattern is produced when these wavefronts interfere during reconstruction.

If the reconstructed object waves are diffracted in equal directions from the hologram, then the interferogram is known as an infinite fringe pattern. This name originates from the fact that a bright fringe of infinite extent is observed if the flow does not modify the density field. Alternatively, a finite fringe pattern may be produced if the reconstructed wavefronts are inclined to each other, so that the flow information is superimposed on a series of linear carrier fringes.

The hologram used to produce figure 26 was recorded using diffuse illumination and therefore speckle noise is prominent in the image. The interferogram also shows that the transmitted illumination is depleted around the edges of the glass container due to refraction (shadowgraph effect). In these dark regions, the fringe visibility is reduced because of the mismatch in intensities between the object and reference beams.

Figure 26: Double-exposure interferogram of CO_2 venting from a glass column

When two object waves o_1 and o_2 are reconstructed from a hologram and recombined to form an interferogram, constructive interference occurs for points on the superimposed wavefronts where the net phase difference is a multiple of 2π. In the regions of constructive interference a bright fringe is observed, whereas a dark fringe occurs where the wavefronts are in anti-phase. Figure 27 illustrates the replay of a double-exposure hologram recorded with an off-axis spherical reference wave. The reconstructed object waves are diffracted from the hologram at equal angles and may be imaged to form an infinite fringe interferogram, where the fringes appear 'fluffed-out'.

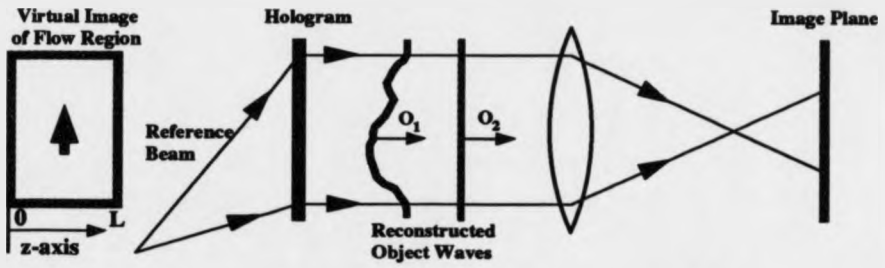


Figure 27: Reconstruction of a double-exposure hologram

In the image plane, the position of the Nth bright fringe centre is defined by a locus of optical path difference. When the viewing direction is parallel to the z-axis, and the flow field is confined within the bounds $0 \leq z \leq L$, the integral equation describing the locus is:

$$\int_0^L [n(x,y,z) - n_0]dz = N\lambda \tag{53}$$

Where,

$n(x,y,z)$ - refractive index distribution of the flow field
n_0 - refractive index of static air
λ - wavelength of light used to record the hologram

Conversely, the condition of maximum destructive interference is defined by intermediate values of optical path difference:

$$\int_0^L [n(x,y,z) - n_0]dz = (N + 1/2)\lambda \tag{54}$$

If an interferogram is imaged onto a two-dimensional imaging sensor, for example a charge coupled device (CCD), the intensity distribution $i(x,y)$ observed when the finite sampling properties of the detector are neglected, is of the form:

$$i(x,y) = a(x,y) + b(x,y)\cos[\phi(x,y)] \tag{55}$$

Where,

$a(x,y)$ - background intensity of the image (bias level)
$b(x,y)$ - modulation envelope of the fringes
$\phi(x,y)$ - phase distribution encoded in the fringe pattern

The function $\phi(x,y)$ describes the entire phase distribution of the interferogram, in addition to the values at the fringe centres. If refraction is neglected, the phase difference between a wavefront passing through static air and one perturbed by a compressible flow, is directly proportional to the path integrated change in refractive index. The refractive index of a gas $n(x,y,z)$ is related to its density $\rho(x,y,z)$ by the Gladstone-Dale relation (equation 27), which is restated here for convenience:

$$n(x,y,z) = K_g\rho(x,y,z) + 1$$

The Gladstone-Dale constant K_g is a function of gas species, but is only weakly dependent on gas pressure and illumination wavelength. An expression relating flow density to the phase of the interferogram is therefore given by:

$$\phi(x,y) = \frac{2\pi}{\lambda}\int_0^L K_g[\rho(x,y,z) - \rho_0]dz \tag{56}$$

Where ρ_0 is the density of static air.

Equation 56 provides a means for calculating three-dimensional density information from holographic interferograms. If isentropic flow is assumed it is also possible to derive local Mach number from holographic density measurements (figure 9). Fringe analysis involves calculating the phase distribution $\phi(x,y)$ from interferograms and

the accuracy of this process determines the feasibility of quantitative holographic flow visualisation. In the period 1964-1970 there was an explosive development in holographic research for metrological applications, but the interest subsided because of the inherent difficulties in processing interferometric data. Subsequently, there has been a prolonged effort to automate fringe analysis, and methods have now been devised that can extract the phase to an accuracy exceeding 0.01 radians.

Fringe analysis originated in the early 19th century, when Thomas Young measured the spacing of fringes to calculate the wavelength of light. Concerted efforts at automating the process began in the 1960's, when analogue electronic devices were developed for locating fringe centres. To automate data acquisition, commercial densitometers were used to scan a photographic transparency of the interferogram. This method was crude and the phase values could only be determined to an accuracy of approximately one radian. Also, the detection of fringe peaks required extensive manual intervention and the images took several minutes to process.

The next development was initiated by advances in digital computer technology, most notably in terms of processing speed and memory capacity. In addition, inexpensive framestores and high quality solid state CCD cameras became available for digitising the interferograms. However, although improvements in hardware substantially decreased processing time, the fringe orders still had to be labelled interactively using a *priori* knowledge about the test object or flow field.

Recent advances in automated fringe analysis have been dominated by the development of electro-optic reconstruction systems, in conjunction with the implementation of knowledge based image processing software. Algorithms have been devised that operate on data which has been encoded onto a spatial or temporal carrier wave. Consequently, amplitude noise that is not modulated at the correct frequency, can be removed by digital or analogue filtering. Also, because the data is biased by the carrier wave, it is possible to distinguish between elevation and depression in the phase surface. The phase distribution may be determined with greater accuracy when a carrier wave is used and the data is not constrained to fringe maxima or minima.

Phase sensitive fringe analysis reveals the sign of the phase change between adjacent fringes and therefore fringe numbering does not require the use of *a priori* knowledge. Regions of high amplitude noise in the image are readily identifiable so that data extraction may be fully automated. Although these methods are computationally intensive, a parallel growth in digital computing technology has meant that 512×512 8-bit images may be processed at speeds approaching video rates. A survey of the latest fringe analysis techniques has been made by numerous authors (e.g. [160]) and a review of the most important results is presented in the remainder of this chapter.

4.1.2 Fringe Tracking and Counting

Digital image processing was first applied to fringe analysis with the development of tracking and labelling algorithms that locate the fringe centres defined by equations 53 and 54. In practice, these techniques are partially interactive and frequently unsuccessful when applied to complicated or looped interferograms. Fringe tracking is also error prone when parts of the image are occluded or noisy. These restrictions are particularly relevant in the analysis of holographic flow data, where the density field may be extremely convoluted and the wind tunnel model casts a shadow. Fringe tracking is currently used when it is difficult to encode the phase information onto a carrier wave; for example in the study of isochromic and isoclinic fringes in photoelasticity [161]. The basic principles of fringe counting and tracking are reviewed in this section to illustrate the deficiencies of these techniques when compared to phase sensitive strategies.

The success of fringe tracking and counting algorithms is dependent on the characteristics and signal-to-noise ratio of the interferogram. Correct identification of fringe centres is non-trivial as the intensity profile $i(x, y)$ is not a simple sinusoid, but rather of the form defined by equation 55. The main difficulties experienced when automating fringe analysis are due to:

- Black areas produced by opaque obstructions (e.g wind tunnel model and sting).
- Broken or discontinuous fringes.
- Non-uniform fringe contrast resulting from incomplete interference.
- Extraneous fringes produced by parasitic interference.
- Diffraction noise from dust and other scattering centres.
- Presence of both broad and narrow fringes.
- Variations in fringe orientation.
- Lack of a reference position.
- Ambiguous fringe sign and order.
- Uneven background intensity due to non-uniform illumination and reflectance.
- Speckle noise (ameliorated by using specular illumination).
- Fragmentation of fringe edges during digitisation.
- Errors introduced when post-processing with analogue or digital electronics.

The simplest method of locating fringe centres is to detect the loci of maximum or minimum intensity (equations 53 and 54). If intensity extrema are used, however, variations in fringe modulation or bias level sometimes result in the incorrect assignment of fringe positions. These errors are accentuated by the presence of coherent noise in the image (speckle and parasitic interference). An alternative is to binarise the interferogram about a fixed threshold value. The fringes may then be skeletonised by digital erosion techniques until their centres are located [162, 163, 164, 165]. This method is more accurate because it is less susceptible to local variations in intensity and the erosion algorithm may be tailored to a particular fringe width. Nevertheless, large positional errors are still incurred due to variations in bias level and fringe contrast. A more robust technique is to assign the values about a floating threshold, to compensate for variations in the background intensity [166, 167]. Alternatively, the gradient of the fringes may be calculated to detect turning points [168].

Attempts at analysing complicated fringe patterns, for example in calculating the topology of turbine blades [169], have involved tracking algorithms which describe the orientation of each fringe segment by a normal vector. The position of a fringe at the edge of the field-of-view is mapped onto adjacent fringes using these vectors until a boundary is reached. Funnell has successfully applied similar methods to the processing of circular and other 'closed' fringe patterns [163]. The most recent work in this field is by Krishnaswany [170], who has used both normal and tangential fringe vectors, in conjunction with a 'predictor-corrector' scheme, to map the fringe positions.

One of the first custom-built interactive fringe tracking systems was developed in 1979 by Glassman and Or [171] on a PDP-11/34 minicomputer. This system correctly identified the co-ordinates of fringes, provided they had high contrast and their curvature was small. In practice, when complicated interferograms were processed, the fringe centres had to be assigned interactively by a manual on-screen cursor. Similar systems have been built by Ben-Dor *et al* [172] and a more refined semi-automatic processor, which uses a convolution mask to detect fringe peaks and a light pen for manual intervention, has been developed by Yatagai [173].

Once the fringe centres have been located it is necessary to assign fringe orders so that the phase relationship between adjacent contours is known. For example, a series of parallel linear fringes may represent a phase surface which is increasing or decreasing in height monotonically, or one that is corrugated. When a single infinite fringe interferogram is available, the sign of the phase change between fringes can only be determined if the general topography of the reflecting surface or transparent phase object is already known. This confusion between gradients of opposite sign is illustrated in figure 28: the one-dimensional fringe patterns generated by these two functions are identical, although their phase profiles are clearly different.

The ambiguity in the sign of the phase change between adjacent fringes, may be resolved without the use of a *priori* knowledge, if the interferometric data is encoded onto a spatial carrier wave. Physically this entails adding a slope of uniform gradient to

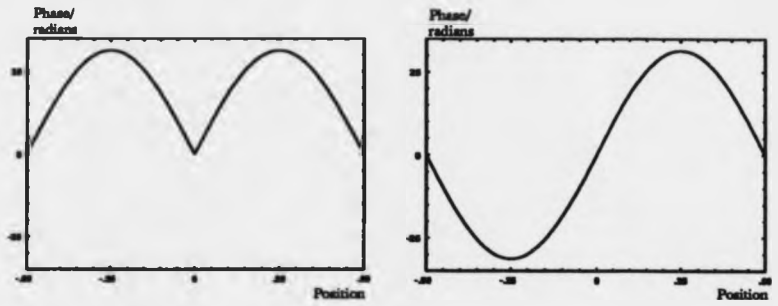


Figure 28: Production of identical infinite fringe patterns by different phase profiles

the phase surface. If the superimposed gradient exceeds all opposing gradients present in the data, then the phase will always change monotonically between fringes and the ambiguity is removed. This is illustrated in figure 29, where the addition of a spatial carrier wave results in a markedly different fringe pattern (alternate fringes are plotted for clarity). A phase gradient may be added to the interferogram by encoding the two holographic exposures with different reference beam angles. The phase relationship between contours may then be determined by examining the deviation of nominally straight fringes.

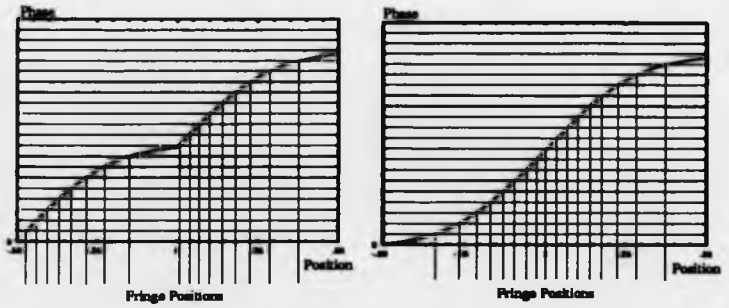


Figure 29: Removal of phase ambiguities by superposition of a carrier wave

After the fringe orders have been assigned, the intermediate phase values may be estimated by interpolation. In practice, the maximum interpolation accuracy is approximately $\pi/4$ radians, because the intensity profile of the fringes is strongly dependent on variations in modulation and bias level. Consequently, fringe tracking is only successful if there are many fringes across the image field, so that a satisfactory

dynamic range is obtained. A detailed study of the latest methods in fringe pre-processing, tracking and labelling is given by Hunter [174].

In summary, fringe tracking is a useful technique for extracting data from double-exposure holograms that have been recorded with only a single reference beam angle. Anomalies in the interferogram, such as regions of high or low object reflectivity, may however, be misinterpreted as fringe information. This can result in the incorrect assignment of fringe orders and these errors can only be corrected by substantial user intervention. Consequently, the best results are obtained when the processing algorithm is dedicated to a particular type of fringe pattern. Unfortunately, this loss in generality has meant that fringe tracking and labelling algorithms are often specific to individual research groups and further technical progress has been restricted.

When fringe tracking is successful, the data is constrained to the skeletonised fringe centres. Hence, if the fringe density is low, the phase map is sparsely sampled and the positional errors of the fringe centres is high. Conversely, if the fringe density is too high, the Nyquist frequency of the detector may be exceeded (see section 5.2.1) and aliasing errors are introduced. Phase sensitive fringe analysis, which is discussed in the following section, ameliorates many of these problems and facilitates fully automated data extraction.

4.2 Phase Retrieval by Spatial or Temporal Heterodyning

4.2.1 Principles of Phase Sensitive Fringe Analysis

The physical data contained in the phase of an interferogram is encoded as an intensity distribution, which is subject to systematic amplitude noise. To extract quantitative information from this image requires the removal of these amplitude variations so that the phase can be decoded without errors. After amplitude noise has been filtered from the interferogram, the extraction of phase information involves two separate processes: firstly, the 'wrapped' phase map, consisting of principal values, must be obtained; and secondly, the absolute values must be calculated by an unwrapping algorithm.

The principles of fringe analysis may be illustrated by considering a single line of data through a hypothetical phase surface. Figure 30 shows a parabolic phase profile $\phi(x)$, which models the retardation of an optical wavefront after passing through a compressible flow.

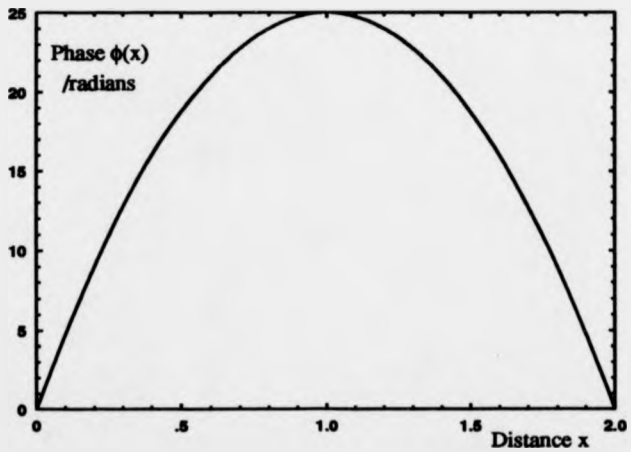


Figure 30: Line through a parabolic phase surface

The profile is described mathematically by:

$$\phi(x) = 25[1 - (x-1)^2] \tag{57}$$

A peak phase retardation of 25 radians is experienced by the wavefront, which is equivalent to an optical path length increase of approximately four wavelengths. Consequently, an interferogram of this wavefront would consist of four fringes either side of the central peak value (eight complete fringes in total). An experimental fringe pattern of this object would also exhibit variations in background intensity and fringe modulation. These features are shown schematically in figure 31.

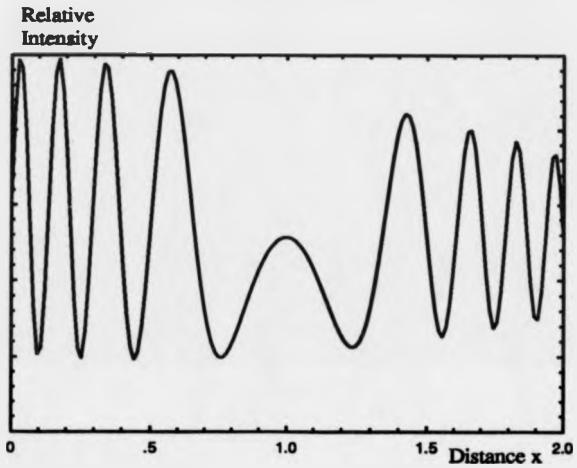


Figure 31: Line through the simulated fringe pattern

Initially, amplitude noise introduced by spatial variations in bias level and fringe modulation, must be removed from the interferogram. This may be achieved by image subtraction, with the amplitude distribution held constant while the phase is varied. Alternatively, the phase data may be encoded onto a carrier wave, so that the amplitude noise remains unmodulated and may be removed by spatial filtering. Figure 32 shows the fringe profile after the multiplicative and additive amplitude noise terms have been removed.

Once a fringe pattern of uniform amplitude has been obtained, the phase values may be calculated by an inverse trigonometric operation. A unique solution is not obtained, however, because only the principal values of the phase argument are retrieved (figure 33). The phase values ramp continuously over a range of 2π and are then decremented discontinuously (wrapped) by 2π at the end of each cycle. Consequently, a

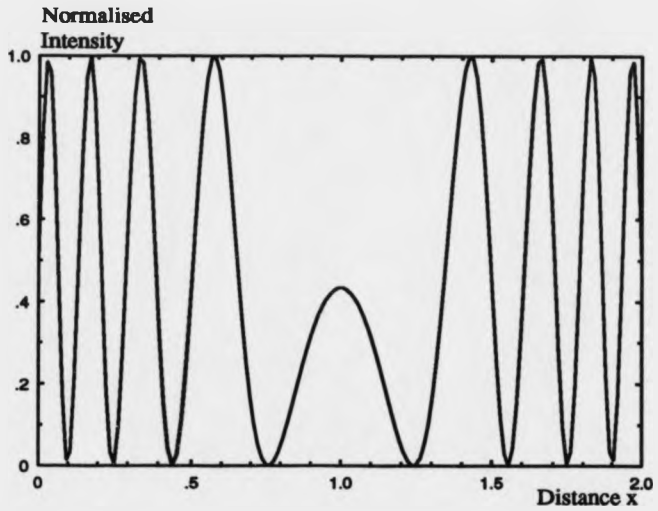


Figure 32: Fringe pattern with the amplitude noise terms removed

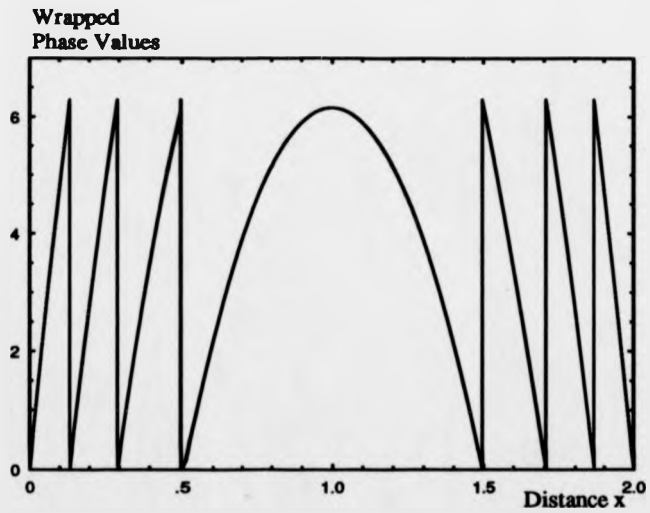


Figure 33: Wrapped phase map

phase unwrapping process is required to convert principal phase measurements into absolute values that may be related to flow parameters. Essentially this may be achieved by the following operations:

- Edge detection of the discontinuities.
- Inspection of the sign of the gradient prior to a discontinuity.
- Modification of phase values after the discontinuity by a multiple of 2π, where the precise multiple is determined by keeping a running total of the number of positive and negative discontinuities that occur between a fixed reference point and the current value.
- Calibration of the unwrapped phase map by determining two density values from an independent measurement technique.

The phase unwrapping process is non-trivial due to aliasing, discontinuous fringes, missing data and noise. Recently, algorithms have been devised which unwrap the phase along a path that exhibits a minimum number of inconsistencies [175, 176]. Throughout this study, software originally developed by Judge [177] was used to phase unwrap the images. This algorithm prevents the propagation of phase errors across the interferogram by dividing the image into an array of smaller tiles. These are phase unwrapped separately and 'sewn together' by determining the phase relationship between adjacent tiles. The path used to join the tiles is determined by computing the noise levels and inconsistencies of each tile boundary, and assigning an appropriate figure of merit. The optimum unwrapping path is then decided by calculating the 'minimum spanning tree' of these weights.

Fringe analysis techniques which operate solely on intensity information, for example fringe tracking and counting, are chiefly restricted to the calculation of phase values at the extrema of the fringes. Interpolation between fringes is limited, because the intensity distribution of the pattern is not solely a function of the encoded phase information, but also dependent on fringe contrast and bias level. To improve the

interpolation accuracy, and determine whether the phase is increasing or decreasing between fringes, it is necessary to operate on the phase of the interferogram directly, while all other intensity dependent factors are held constant.

The phase difference between two reconstructed object waves can only be modified if the two holographic recordings are encoded uniquely, so that they may be replayed individually (section 3.4.4). The principle techniques used for separate image encoding are:

- Real-time reconstruction.
- Dual plate recording or sandwich holography.
- Dual reference beam methods.

Real-time holographic interferometry combines a holographically reconstructed object wave of the quiescent flow state, with a 'live' wavefront that propagates through the flow. The holographically reconstructed wavefront is generated from an off-axis hologram using a reference beam which is spatially separate from the live scene beam (figure 34). One of the two wavefronts forming the interferogram may be phase modulated to enable phase sensitive fringe analysis. The principal attraction of this configuration is that the live object wavefront enables the study of dynamic flows in real-time. However, the penalty paid by real-time operation is that the system must remain interferometrically stable to retain the infinite fringe condition.

Holographic flow visualisation is normally performed with a pulsed laser to prevent decorrelation of phase information due to vibration. It is impractical to use a real-time interferometer in a production wind tunnel unless a 'common path' is used for the object and reference beams. Hence, a method is required to encode the object waves separately, so that real-time phase modulation may be performed during reconstruction in a stable laboratory environment.

The quiescent and perturbed density fields may be recorded separately if either the reference beam angle or recording medium are changed between exposures. When

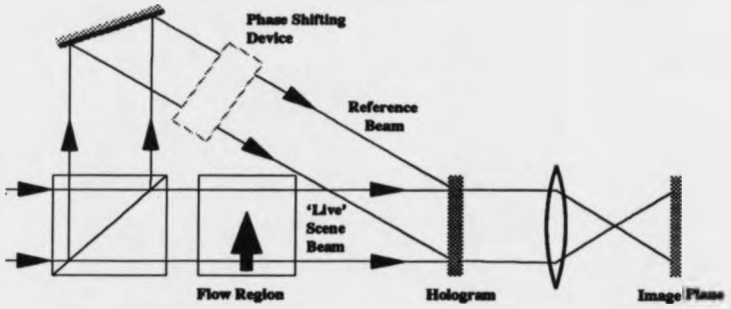


Figure 34: Real-time holographic interferometer

the dual reference beam method is used, the object waves are multiplexed on the hologram at different spatial frequencies, so that during reconstruction they may be separated and modulated independently. Alternatively, individual holographic plates may be used to record the phase of the two wavefronts, with a kinematic holder used to realign the holograms in the reconstruction system.

During replay of dual reference beam or twin plate holograms, the angle between the reconstructed wavefronts may be adjusted to remove error fringes introduced by rigid body motion of the recording apparatus. Similarly, the phase of one or both beams may be modulated in either the spatial or temporal domains, to facilitate phase sensitive fringe analysis. Fourier transform fringe analysis operates in the spatial frequency domain and entails the encoding of flow information onto a set of equally spaced linear fringes. In practice, this finite fringe pattern is formed by inclining the two object waves with respect to each other. If a phase object is encoded onto one of the object beams, the fringes are deviated from their ambient positions and the carrier frequency is modulated. A comparison of finite and infinite fringe patterns is made in figure 35. The flow consists of a supersonic carbon dioxide jet emanating from a 2mm diameter convergent-divergent nozzle.

An alternative strategy is to process the phase information in the temporal domain by a method known as heterodyne holographic interferometry. When a temporal carrier wave is used, the flow information is encoded onto an optical wavefront which is continuously phase modulated. The absolute frequency of the oscillating electric

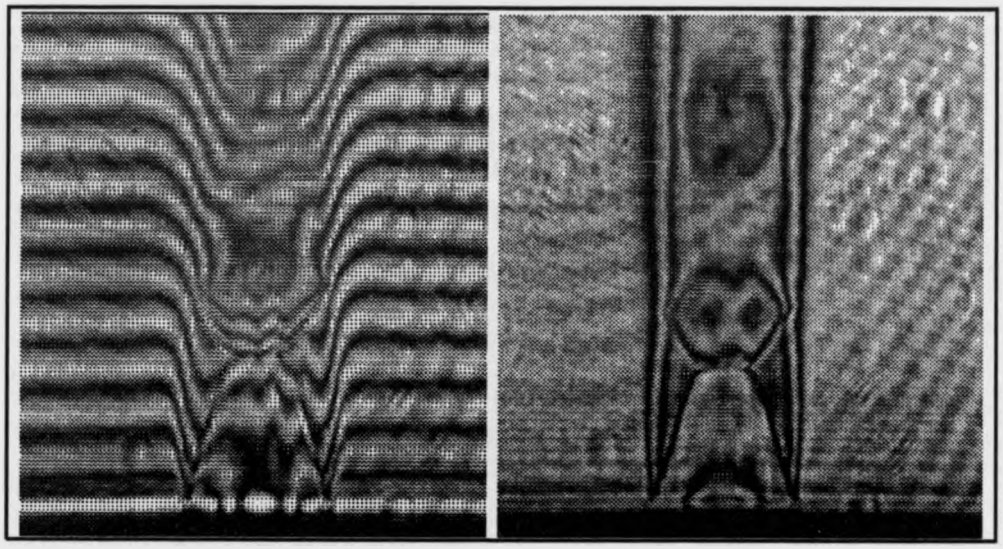


Figure 35: Finite and infinite fringe patterns of a transonic carbon dioxide jet

field exceeds the bandwidth of current detectors ($> 6 \times 10^{14}$Hz) and therefore the wavefront must be heterodyned with the unmodulated object wave of the quiescent flow state, to scan the fringes at the beat frequency. The interferometric data may then be extracted electronically by phase sensitive detection of the heterodyned signal. Intensity variations that are independent of phase differences between the component object waves remain unmodulated and may be removed by high-pass filtering.

A discretised version of heterodyne fringe analysis, known as quasi-heterodyne holographic interferometry, has been developed to simplify automated data extraction. Instead of using continuous phase modulation, this method involves the recording of a series of phase stepped interferograms. The absolute phase difference between the object waves may then be calculated by solving a set of simultaneous equations, with each image contributing a single equation containing the unknowns $a(x,y)$, $b(x,y)$ and $\phi(x,y)$. To solve for these unknowns it is usual to record at least three images.

Takeda has recently unified the methods of temporal and spatial phase modulation

to capitalise on the increased space-bandwidth-product afforded by processing the data in both domains [178]. A hybrid of the Fourier transform and discrete phase stepping techniques has also been developed, to enable calculation of phase stepped information from a single interferogram [179]. Details of these various phase sensitive fringe analysis techniques are given in a comparative study presented in the remainder of this chapter.

4.2.2 Phase Extraction by Temporal Heterodyning

A detailed description of heterodyne fringe analysis is given in several papers by Dandliker *et al* [180, 181]. The technique entails the real-time processing of a periodic temporal signal and therefore the system must be kept interferometrically stable to enable observation of the 'live' interference fringes. The stability requirements are less restrictive if a pulsed dual reference beam method is used to record the data, followed by real-time reconstruction and phase extraction under laboratory conditions.

During replay of a dual reference beam hologram, two object waves are reconstructed which may be described mathematically by:

$$\begin{aligned} \mathbf{O_1}(x,y) &= a_1(x,y)\cos[\phi_1(x,y) + w_1 t] \\ \mathbf{O_2}(x,y) &= a_2(x,y)\cos[\phi_2(x,y) + w_2 t + \psi] \end{aligned}$$

Where the variables describing the waves are defined as follows:

$\mathbf{O_1}, \mathbf{O_2}$ -	complex amplitudes
a_1, a_2 -	real amplitudes
ϕ_1, ϕ_2 -	phase distributions
w_1, w_2 -	angular frequencies
t -	time dependence
ψ -	constant phase difference between the object waves

When the hologram is replayed using two unmodulated wavefronts, the reconstructed object waves are of identical frequency and a stationary (time independent) fringe pattern is observed. However, due to variations in fringe contrast and non-uniformity of the unmodulated background illumination, the phase difference between the object

waves can only be determined to an accuracy of approximately π radians. Also, without the addition of a spatial or temporal carrier wave to the pattern, it is impossible to determine whether the phase is increasing or decreasing between fringes.

Fringe interpolation accuracy may be greatly enhanced by frequency shifting one or both of the object waves during reconstruction. Suitable devices for continuous phase modulation include rotating polarisers [182] and moving diffraction gratings [183]. A dynamic diffraction grating may be formed, without mechanical movement, by generating a travelling wave of refractive index modulation in an acousto-optic Bragg cell [184]. These devices operate at megahertz frequencies and therefore both object waves are usually modulated at slightly different frequencies to produce a beat signal (at typically 100kHz). An alternative to extrinsic phase modulation, is to create two phase correlated wavefronts, separated in frequency by approximately 1MHz, using a Zeeman laser [185].

When two coherent object waves of slightly different frequency interfere, the resulting intensity distribution is given by equation 58, where the notation of equation 55 has been adopted.

$$i(x,y) = a(x,y) + b(x,y)\cos[(w_2 - w_1)t + (\phi_2 - \phi_1) + \psi] \tag{58}$$

The animated fringe pattern may be analysed using phase sensitive detection. Unfortunately, the bandwidth of the carrier wave required to encode the flow information exceeds that of current CCD cameras and therefore it is necessary to use conventional photodetector technology. The interferogram may be processed by holding a single photodiode stationary in the fringe pattern, while a second detector is scanned over the entire field. The signals from the two sensors can be analysed by a lock-in amplifier which uses part of the signal fed to the acousto-optic cell as its reference.

Once the intensity variation at a point in the interferogram $i[\phi(t)]$ has been converted into an electrical waveform, analogue electronics may be used to demodulate the signal and filter the noise terms $a(x,y)$ and $b(x,y)$. Subsequently, the wrapped

phase $\phi(x,y)$ may be measured from the two processed signals using an electronic phasemeter. The sensitivity of heterodyne fringe analysis is governed by the magnitude of the signal at the beat frequency compared to noise levels. The fundamental accuracy is limited by thermal Johnson noise of the electronics and quantised shot noise of the light. These noise sources may be defined as noise equivalent powers at the output of the detectors:

$$P_{jn} = 4KTB\left(\frac{R_0}{R_l}\right) \tag{59}$$

$$P_{sn} = \frac{2R_0^2 e^2 B\eta\overline{P}}{R_l h f} \tag{60}$$

Where,

- P_{jn} - noise equivalent power of Johnson noise
- P_{sn} - noise equivalent power of shot noise
- B - temporal bandwidth of the detector
- R_0 - feedback resistance of the detector
- R_l - load resistance of the detector
- e - electronic charge
- η - quantum efficiency
- $\overline{P}$ - average optical power incident on the detector

With these definitions, the signal power P_s at the heterodyne frequency is given by:

$$P_s = \frac{2P_o P_r}{R_l}\left(\frac{R_0 m_f \eta e}{hf}\right)^2 \tag{61}$$

Where,

- m_f - fringe modulation
- P_o, P_r - optical power of the object and reference beams (at the detector)

Willemin *et al* [186] have investigated the effects of shot and Johnson noise on the accuracy of phase determination for a conventional interferometer using specular illumination. They speculate that because the reference wavefront is free from signal noise, it may be made more intense than the object beam so that the measurements

are dominated by shot noise. With this assumption, the signal-to-noise ratio for fringes of unity modulation, is given by:

$$SNR = \frac{\eta P_o}{hfB} \tag{62}$$

Equation 62 indicates that phase measurement accuracy is highest for a narrow band sensor, with a peak sensitivity and quantum efficiency at long reconstruction wavelengths. Willemin has applied non-holographic heterodyne interferometry to the study of submicroscopic vibrations within the human inner ear and has demonstrated, both theoretically and experimentally, that a vibration amplitude of 3×10^{-13}m may be detected [186]. This result illustrates that a fundamental phase measurement accuracy of approximately 10^{-5} radians (millionth of a fringe) is attainable for a specular illumination interferometer maintained in a perfectly stable environment. For practical holographic systems that use extended wavefronts, speckle noise and environmental disturbances impose greater constraints on measurement precision, and a limiting accuracy of typically 10^{-2} radians (one thousandth of a fringe) is generally accepted [187].

4.2.3 Quasi-Heterodyne Phase Stepped Fringe Analysis

Stringent stability requirements, and the inconvenience of lengthy sequential data acquisition, have limited the use of heterodyne holographic interferometry to environmentally controlled laboratories. To overcome these restrictions, quasi-heterodyne fringe analysis was developed in 1974 by Bruning *et al* to automate optical shop testing [188]. Unlike heterodyne interferometry, this method uses a CCD camera and framestore to record the entire intensity distribution of the fringe pattern. Subsequent image processing is performed digitally and the output data is produced in a convenient format for direct comparison to computer predictions. Also, quasi-heterodyne fringe analysis algorithms operate on individual pixels, without the use of neighbouring information, and therefore can be implemented on a parallel computer system to

facilitate frame-rate processing.

Quasi-heterodyne fringe analysis retains many of the benefits associated with phase sensitive detection, while offering greater potential for automated data processing. The method entails the recording of several phase stepped interferograms, from which a set of simultaneous equations are formed containing the unknown phase distribution. Amplitude noise, introduced by variations in fringe modulation, bias level and diffraction, is eliminated by phase stepping because the fringes move across the image, whereas the noise remains stationary. Quasi-heterodyne interferometry is sometimes referred to as phase stepped fringe analysis, due to the use of discrete phase shifts; or as digital fringe analysis, because of the implementation of digital electronics.

Quasi-heterodyne interferometry enables parallel processing of fringe patterns, unlike full heterodyne interferometry which operates sequentially. However, parallel data acquisition is achieved at the expense of a marked reduction in the sampling rate of the signal and a corresponding loss in phase measurement accuracy. Also, the algorithm that computes the phase distribution from the series of interferograms assumes that the recording process is linear, the intensity of the fringes varies sinusoidally and the phase steps are implemented correctly. Nevertheless, a phase measurement accuracy of 0.2 radians ($\sim$1/30th of a fringe) is generally accepted, although this figure is highly dependent on the presence of systematic errors. This accuracy is considerably poorer than for continuous temporal modulation; however the data may be acquired and processed at near video rates ($\sim$80Mbit/s). In practice, this means that if the reconstruction system is unstable for periods greater than approximately one second, then the phase stepped technique may actually yield more accurate results than continuous heterodyne interferometry.

Phase stepped fringe analysis is normally implemented in a real-time holographic interferometer, or in a reconstruction system that uses two angularly separated reference beams. If a dual reference beam configuration is chosen, the angle between the reconstruction waves must be sufficient to enable insertion of the phase shifting device into only one of the beams. A fringe pattern is characterised by three functions

of the spatial co-ordinates (equation 55). To calculate these unknowns algebraically it is necessary to produce at least three phase stepped interferograms, so that three independent simultaneous equations can be formed. These images may be obtained by successively incrementing the phase of one of the two replay beams by 120°, with a CCD camera and framestore used to capture and store the fringe patterns. Data extraction accuracy may be improved if more than three phase steps are used. Equation 63 describes the intensity distribution of the kth phase stepped interferogram:

$$i_k = a(x,y) + b(x,y)\cos[\phi(x,y) + k\delta_p] \tag{63}$$

A general solution, derived in appendix 1, for calculating the phase from these images is:

$$\phi(x,y) = \tan^{-1}\left[\frac{\beta_f(x,y)}{\alpha_f(x,y)}\right] \tag{64}$$

$$\phi(x,y) \in 0 \text{ to } 2\pi$$

Where $\alpha_f(x,y)$ and $\beta_f(x,y)$ are the constant Fourier coefficients:

$$\alpha_f(x,y) = \sum_{k=0}^{N} i_k(x,y)\cos(k\delta_p) \tag{65}$$

$$\beta_f(x,y) = \sum_{k=0}^{N} i_k(x,y)\sin(k\delta_p) \tag{66}$$

and N is defined such that $2\pi/\delta_p > N > 2\pi/\delta_p - 1$.

One particular solution for three phase steps is given by equation 67, where the spatial dependence of the variables has been omitted for clarity:

$$\phi(x,y) = \tan^{-1}\left\{\frac{1}{\sqrt{3}}\left[\frac{i_3(4\pi/3) - i_2(2\pi/3)}{i_1(0)}\right]\right\} \tag{67}$$

An alternative algorithm for analysing three phase stepped images is given by:

$$\phi(x,y) = \tan^{-1}\left\{\frac{\sqrt{3}[i_1(2\pi/3) - i_2(4\pi/3)]}{[-i_1(2\pi/3) - i_2(4\pi/3) + 2i_3(2\pi)]}\right\} \tag{68}$$

Similarly, four steps may be combined to yield an improved estimate for $\phi(x,y)$:

$$\phi(x,y) = \tan^{-1}\left\{\frac{1}{2}\left[\frac{i_4(3\pi/2)}{i_1(0)} + \frac{i_2(\pi/2)}{i_3(\pi)}\right]\right\} \tag{69}$$

or alternatively,

$$\phi(x,y) = \tan^{-1}\left\{\frac{i_1(0) - i_3(\pi)}{-i_2(\pi/2) + i_4(3\pi/2)}\right\} \tag{70}$$

The most popular method used by commercial interferometers is the five step technique developed by Hariharan [189]. The images are produced with phase steps of $\pi/2$ over an interval $[-\pi,\pi]$, and the phase of the interferogram is given by:

$$\phi(x,y) = \tan^{-1}\left\{\frac{2[i_2(-\pi/2) - i_4(\pi/2)]}{2i_3(0) - i_5(\pi) - i_1(-\pi)}\right\} \tag{71}$$

The redundancy incorporated when using more than three steps reduces the sensitivity of the algorithm to systematic errors. Nevertheless, Tyrer *et al* [190] have shown that when the phase step is calibrated accurately, and the system is isolated from environmental disturbances, it is possible to obtain acceptable results with a single phase shift of π radians. When only two anti-phase images are used, however, there are insufficient equations to determine the three unknowns and therefore the unmodulated bias level $a(x,y)$ must be removed by digital spatial filtering.

The single phase step technique is particularly useful if the anti-phase components can be generated simultaneously, or acquired at video rates to minimise vibration and air currents. The system developed by Tyrer *et al* in 1988 was based on electronic speckle pattern interferometry (ESPI) and therefore the image processing time was

reduced by using a single band-pass filter to remove both the unmodulated bias term and high frequency speckle noise. A danger in using spatial filtering is that phase errors may occur if signal information is accidentally discarded due to an overlap of the spectra from $a(x,y)$ and $\phi(x,y)$. A detailed discussion of spatial filtering operations is presented in chapter 5, in the context of Fourier transform fringe analysis.

An alternative to phase stepping is the phase integrating method. This involves linearly increasing the phase of one of the object waves during image capture, so that the fringes move continuously across the field-of-view. With the integrating four-bucket method, four images are captured with the phase shifted linearly by $\pi/2$ during each frame capture. An identical algorithm to that employed in the analysis of phase stepped images is then used to calculate the wrapped phase map. The advantages of the phase integrating method are:

- The Fourier expansion of the phase object (equation 66) contains a greater number of terms (Fourier transform rather than series). Consequently, noise is reduced because more statistically independent information is utilised.
- The phase shifting element moves continuously and therefore the data can be acquired without the danger of overshoot and ringing.
- Synchronisation of the framestore and phase shifter is easily implemented by applying a ramped voltage from a signal generator.

A disadvantage of the phase integrating technique is that, because the phase of each image is lost, the fringe contrast is reduced by the constant velocity of the fringes over the detector.

The accuracy of quasi-heterodyne fringe analysis is subject to numerous sources of systematic error. The largest errors are produced by:

- Air turbulence.
- Vibration.

- Inaccurate calibration of the phase shifter.
- Non-linear response of the phase shifter.
- Non-linear response of the imaging sensor.
- Intensity quantisation by the detector and framestore.

It is possible to minimise or compensate for all of these sources of error, except those due to intensity quantisation. Creath has shown [187] that for fringes of intensity modulation m_f, the rms quantisation phase error q_p, for Q discrete levels is:

$$q_p = \frac{1}{\sqrt{3} m_f Q} \tag{72}$$

If the fringes have unity modulation, and an 8-bit image is used with 256 grey levels, the phase error is 2.26mrads. This measurement uncertainty is equivalent to an interpolation accuracy of 1/2500th of a fringe. In practice, other sources of noise predominate and quantisation errors only place an upper limit on data extraction accuracy. This example illustrates the futility of using sensors and dedicated framestores with a dynamic range exceeding 8-bits, if the images are noisy or contain systematic errors.

Unless precautions are taken, air turbulence, mechanical relaxation, and vibration, can induce phase shifts that exceed those being measured. These errors may be minimised by locating the holographic recording and reconstruction systems in a stable environment, and allowing them to reach thermal and mechanical equilibrium. Phase disturbances created by air currents can be prevented if the optical path lengths are minimised and the beams are enclosed to restrict air circulation. Vibration isolation is most easily accomplished by building the system on a floating optical table, which is situated in a room with acoustic damping to prevent the coupling of sonic waves into the optical components. Higher levels of vibration may be tolerated if a feedback system is used to phase-lock the replay beams.

Errors are introduced into phase stepped fringe analysis if inaccurate phase steps are used to generate the multiple images. The phase shift $\delta\phi'(x,y)$, imparted by an actuator, may be expressed by a quadratic approximation, where the constants ϵ_1 and ϵ_2 denote calibration error and device non-linearity respectively.

$$\delta\phi'(x,y) = \delta\phi(x,y)[1 + \epsilon_1 + \epsilon_2\delta\phi(x,y)] \tag{73}$$

Lanen has shown [81] that when the phase step is mis-calibrated by an amount $\epsilon_1\delta\phi$, the peak-to-peak phase error $\Delta\phi$ is:

$$\Delta\phi(\gamma_n) = 2\tan^{-1}\left[\frac{\gamma_n - 1}{2\sqrt{\gamma_n}}\right] \tag{74}$$

Where γ_n is tabulated in table 2 as a function of the number of phase steps.

Step Number	3	4	5
$\delta\phi$	$2\pi/3$	$\pi/2$	$\pi/2$
γ_n	$\frac{\sqrt{3}}{\tan[\delta\phi(1+\epsilon)/2]}$	$\frac{1}{\tan[\delta\phi(1+\epsilon)/2]}$	$\frac{1}{\sin[\delta\phi(1+\epsilon)]}$

Table 2: Error parameter γ_n as a function of the number of phase steps

Phase step accuracy is governed by the method used to increase the optical path length of one of the two replay beams. Numerous devices have been used and the most common methods are summarised below:

- Mirror mounted to a piezo translator.
- Optical fibre delay line wrapped around a piezo cylinder [191].
- Electro-optic modulator (Pockel cell or electro-optic crystal [192]).
- Tilted glass plate (Doerband *et al* [193]).
- Rotating polariser or half wave plate [182].

- Acousto-optic modulator: e.g. a Bragg cell [184].
- Moving diffraction grating [183].

The most popular phase shifting element consists of a mirror mounted on a piezo translator (known as a PZT due to the composition of lead zirconate titanate). The phase shift $\Delta\phi$ is related to the mirror displacement d_m, and the angle of the incident beam θ, by:

$$\Delta\phi = \frac{4\pi d_m}{\lambda \cos\theta} \tag{75}$$

Most PZT's consist of a stack of ceramic discs, separated by a series of flat electrodes that supply the operating voltage. When a voltage is applied across a PZT, an expansion is produced by the piezoelectric effect. The relationship between applied voltage and PZT extension is non-linear, and hysteresis occurs when the voltage is varied. In addition, there is a residual drift of the PZT surface after the applied voltage has reached a constant value. A strain gauge may be incorporated with the actuator to compensate for these effects and a displacement accuracy of 5nm can be obtained for a PZT with a maximum extension of approximately 5μm (dynamic range of 1000).

A PZT may be used to produce optical path length changes with an accuracy of $\sim$10nm when the incident beam is aligned to the normal of the mirror. A He-Ne laser, operating at a wavelength of 633nm, is often used as a continuous-wave reconstruction source and therefore a displacement precision of 10nm is equivalent to a phase shift accuracy of 0.10 radians. If a three phase step algorithm is used, the peak-to-peak error in the wrapped phase map is 0.12 radians (from equation 74), which is equivalent to an interpolation accuracy of $\sim$1/60th of a fringe. When a four-bucket method is used, the peak-to-peak phase error is reduced to 0.10 radians, which is equal to the calibration accuracy of the phase shifting device.

Agreement between the desired phase step and the one actually produced by the PZT, can be calculated from the phase stepped images to provide a check on the

accuracy of a phase map. This method is particularly useful for determining whether environmental phase disturbances have occurred during image capture (air currents or vibration). Also, it is possible to devise whether the PZT has undergone a tilting motion, in addition to an out-of-plane displacement. To solve for $\delta\phi$, and the three other unknowns $a(x,y)$, $b(x,y)$ and $\phi(x,y)$, it is necessary to capture at least four phase stepped images. When four or five images are produced, with successive phase shifts of $\pi/2$, the respective expressions for $\delta\phi_4(x,y)$ and $\delta\phi_5(x,y)$, are:

$$\delta\phi_4(x,y) = \cos^{-1}\left\{\frac{i_1 - i_2 + i_3 - i_4}{2(i_2 - i_3)}\right\} \tag{76}$$

$$\delta\phi_5(x,y) = \cos^{-1}\left\{\frac{i_5 - i_1}{2(i_4 - i_2)}\right\} \tag{77}$$

Equations 76 and 77 are only valid when the argument of the inverse cosine function is in the interval [-1,1] and the denominator is non-zero. If these conditions are infringed, these erroneous phase values must be discarded to avoid computational errors.

Calibration errors may be avoided if a self-calibrating four step algorithm is used, that was originally developed by Carré in 1966 [194]. The magnitude of the phase step is not required for this method, provided it is constant for each successive image. The wrapped phase map may be calculated from the four images by use of equation 78, where the phase arguments of the images have been omitted for clarity:

$$\phi(x,y) = \tan^{-1}\left\{\frac{\sqrt{|[(i_1 - i_4) + (i_2 - i_3)][3(i_2 - i_3) - (i_1 - i_4)]|}}{|(i_2 + i_3) - (i_1 + i_4)|}\right\} \tag{78}$$

Self-calibrating algorithms are insensitive to variations in the response between individual pixels. Also, if the phase shift is non-uniform over the field, due to poor collimation of the beam incident on the PZT, the results are still accurate provided the phase is stepped linearly at each pixel. If the phase shifted images are acquired at video rates, environmentally induced phase shifts are less detrimental with this

algorithm because these disturbances vary approximately linearly over a period of $\sim$ 100ms.

Further systematic errors are introduced if the response of the phase shifting device is non-linear or exhibits hysteresis. Creath has investigated these errors for different phase stepping algorithms [187], when a non-linear coefficient ϵ_2 of 0.10 is used (10% error in each phase step). The results show that the errors are negligible for the five step method, whereas a peak-to-peak error of 0.16 radians ($\sim$ 1/60th of a fringe) is produced when three steps are used. Higher order non-linearities generate errors of increasing magnitude for equal coefficients, with an error of 0.23 radians ($\sim$1/25th of a fringe) when a quadratic coefficient of 0.10 is used with three phase steps.

In conclusion, phase stepped fringe analysis is a powerful technique for extracting wrapped phase information from interferograms. Phase extraction accuracy is dependent on the stability of the reconstruction system, and linearity of the imaging sensor and phase shifting device. When a phase shifting system is correctly calibrated, and operated in a stable environment, a measurement accuracy of 0.1 radians (1/30th of a fringe) can be expected. Although this is significantly poorer than for continuous heterodyne interferometry, it is sufficient for the extraction of aerodynamic data. For example, a typical transonic flow in a wind tunnel, produces density perturbations which generate between ten and one hundred fringes. Hence, with this technique, the integrated density field may be measured with an accuracy exceeding 0.3%.

4.2.4 Fourier Transform Fringe Analysis

If a linearised fringe pattern is available (figure 36), the phase distribution $\phi(x, y)$ may be retrieved from a single image using a Fourier transform method originally proposed by Takeda *et al* in 1982 [195]. Fourier transform fringe analysis is analogous to heterodyne interferometry in that two signals of differing frequency are mixed to produce a periodic heterodyned signal. The Fourier transform method, however, uses the spatial domain to encode the information, opposed to the temporal domain of

phase stepped and heterodyne techniques. In practice, the spatial carrier wave is produced by recording the two object waves with separate reference beams that differ in angle by approximately one degree.

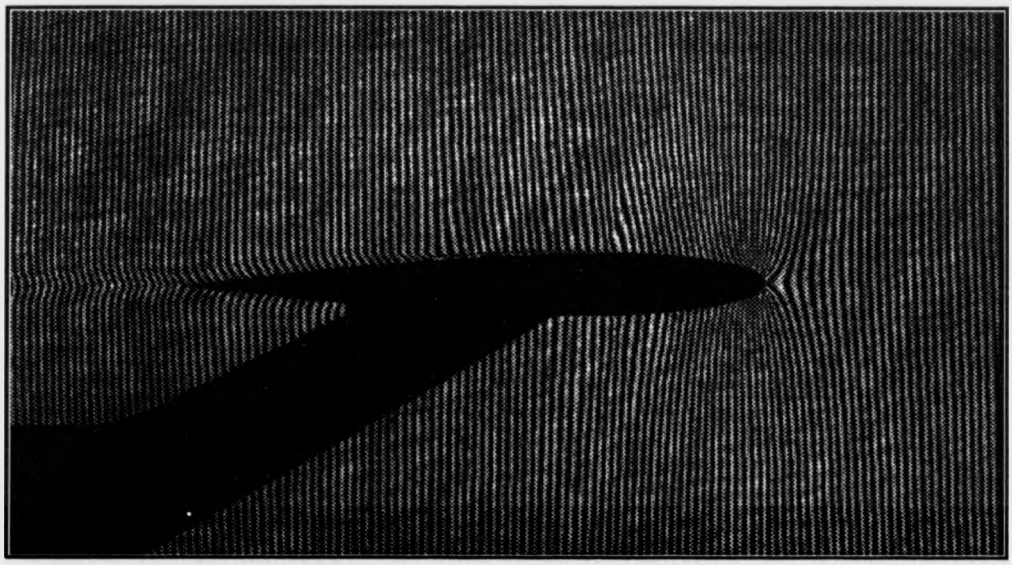


Figure 36: Finite fringe pattern of transonic flow around a NACA 0012 Aerofoil

Fourier transform fringe analysis imposes different demands on the characteristics of the imagery than temporal methods. In particular, because the information is processed in the spatial domain, the image should have a high resolution so that a large carrier frequency can be used to encode the phase of small flow features. The principles of the technique may be described by considering a generalised interferogram, which has been linearised by the addition of a carrier wave of spatial frequency f_x in the x direction:

$$i(x,y) = a(x,y) + b(x,y)\cos[\phi(x,y) + 2\pi f_x x] \tag{79}$$

The cosine term of equation 79 may be expressed as the sum of a complex exponential and its conjugate, by the use of De Moivre's theorem:

$$i(x,y) = a(x,y) + c(x,y)\exp[2\pi j f_x x] + c^*(x,y)\exp[-2\pi j f_x x] \tag{80}$$

Where,

$$c(x,y) = \frac{b(x,y)}{2}\exp[j\phi(x,y)] \tag{81}$$

In Takeda's original algorithm, a one-dimensional Fourier transform was taken of each line in the image. This technique is unsatisfactory when the carrier wave is encoded in both orthogonal directions, as may occur due to camera misalignment, or from rigid body motion of the optics between the two exposures. A better method, first proposed by Macy in 1985 [196], is to compute a two-dimensional Fourier transform of the whole image. The two-dimensional Fourier transform operation, defined by equation 82, converts an image $i(x,y)$ into its spatial frequency spectrum $I(u,v)$:

$$I(u,v) = \int\int_{-\infty}^{\infty} i(x,y)\exp[-j2\pi(ux+vy)]dxdy \tag{82}$$

Other equivalent definitions exist, but this expression is widely used in optics due to the work of Bracewell and Goodman [197, 198]. A consequence of this particular definition is that the inverse operation is of a symmetrical form:

$$i(x,y) = \int\int_{-\infty}^{\infty} I(u,v)\exp[j2\pi(ux+vy)]dudv \tag{83}$$

The two-dimensional Fourier transform of equation 80 may be expressed implicitly as:

$$I(u,v) = A(u,v) + C(u-u_x,v) + C^*(u+u_x,v) \tag{84}$$

Where the use of upper case characters signifies that a two-dimensional spatial Fourier transform has been taken of the functions defined by the lower case characters, * denotes complex conjugation and u_x is the spatial carrier frequency.

The complex Fourier transform of an intensity distribution is Hermittian, with an even real component and an odd imaginary component. A result of this property is that the amplitude spectra is point symmetric. If the carrier frequency exceeds all of the spatial frequencies contained in $a(x,y)$, $b(x,y)$ and $\phi(x,y)$, then the Fourier transform consists of three distinct spectral orders. This is illustrated schematically in figure 37 for a line through the origin of the transform orthogonal to the fringe direction.

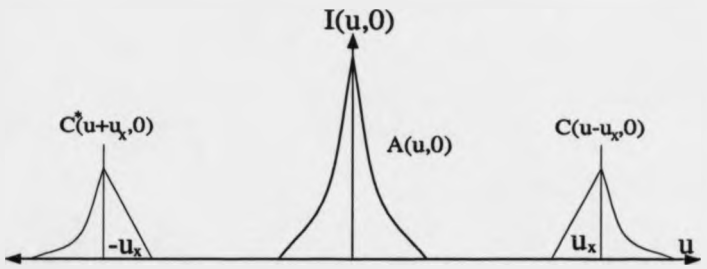


Figure 37: Schematic representation of the three peaked Fourier transform

The zero order component $A(u,v)$, centred on the origin, contains information related to the unmodulated bias level in the image. The zero order term may be removed by spatial filtering, provided that it does not overlap with the first order $C(u-u_x,v)$. Once $A(u,v)$ has been filtered, the Hermittian properties of $I(u,v)$ may be destroyed by discarding the negative frequencies of the Fourier transform. The flow information is subsequently demodulated by translating $C(u-u_x,v)$ back to the origin. All intensity distributions can be described in terms of real numbers and therefore they possess Fourier transforms with Hermittian properties. However, the half spectrum $C(u,v)$ cannot be created by a real image and therefore, when the inverse Fourier transform of $C(u,v)$ is calculated, a complex image $c(x,y)$ is generated (equation 81).

The dependence of equation 81 on fringe contrast $b(x,y)$ may be removed by taking the ratio of the imaginary and real components of the demodulated image $c(x,y)$. This process is analogous to that used in phase stepped fringe analysis (equation 65), where the ratio of the sine and cosine coefficients in the Fourier series is taken. The phase $\phi(x,y)$ may be determined from this result by calculating the inverse tangent function:

$$\begin{aligned} \phi(x,y) &= \tan^{-1}\{\mathcal{I}[c(x,y)]/\mathcal{R}[c(x,y)]\} \\ \phi(x,y) &\in \quad -\pi \text{ to } \pi \end{aligned} \tag{85}$$

Where $\mathcal{R}$ and $\mathcal{I}$ denote the real and imaginary components of $c(x,y)$ respectively.

An equivalent strategy is to take the imaginary component of the complex natural logarithm of $c(x,y)$. The inverse tangent or logarithmic operations produce only the principle phase values, which are indeterminate to modulo π or 2π depending on whether the signs of the sine and cosine coefficients have been utilised. The unwrapped phase values are calculated from this wrapped phase map by inspection of the local phase gradients prior to the discontinuities. A detailed review of phase unwrapping techniques is given in chapter 5.

The accuracy of Fourier transform fringe analysis is more difficult to quantify than quasi-heterodyne strategies. However, the technique only requires the use of a single image for data extraction and therefore it is particularly suited to implementation in harsh environments, where it is difficult to capture a series of correlated interferograms. The processing errors incurred when using this algorithm are assessed in chapter 5 by analysing a combination of experimental and computer generated data.

4.2.5 Phase Stepped Fringe Analysis with a Spatial Carrier Wave

An alternative to Fourier transform fringe analysis is to use a hybrid technique which uses a spatial carrier wave to provide a set of phase shifted images from a single interferogram. This technique was first proposed by Mertz as a means of reducing computational effort to enable frame-rate processing [179]. At present, the Fourier transform of a 512×512 8-bit image can be calculated in 15s CPU time on a SPARC 2 workstation. Consequently, the wrapped phase map takes longer than 30s to compute by Fourier analysis and interactive inspection of flows is impossible unless expensive dedicated hardware is used.

The phase shifted spatial carrier wave technique requires the recording of a single finite fringe pattern, with a predesignated carrier frequency. A linearised interferogram has the phase information encoded on a slope of constant gradient and therefore adjacent pixels are successively phase shifted. If a single period of the carrier wave extends over exactly four pixels, for example as shown in figure 36 where 128 fringes occupy the full 512 pixel image width, the four phase step algorithm (equation 69) may be applied to strings of four pixels which lie orthogonal to the carrier fringes. This calculation is only valid if the variables $a(x,y)$, $b(x,y)$ and $\phi(x,y)$ remain constant over this length and, in practice, the results are spatial averages.

After the phase map has been calculated, it is necessary to subtract the carrier wave which was superimposed on the data. This may be achieved after phase unwrapping by subtracting the appropriate phase ramp, although errors often arise during phase unwrapping because of the large number of discontinuities in the image. A better method is to remove the carrier wave prior to unwrapping, to obtain the demodulated wrapped phase map $\phi'(x,y)$. When the phase difference between pixels is δ_x in the x-direction, the demodulated phase map may be obtained from the wrapped phase map $\phi(x,y)$ using the expression:

$$\phi'(x,y) = \phi(x,y) - i\delta_x \tag{86}$$

Where i is the x co-ordinate of the ith pixel.

For this procedure to work correctly, the direction of the slope must be known. Also, subtraction of the inclined phase surface involves an additional wrapping process which may result in errors at discontinuities in the image.

A disadvantage of phase shifting with a spatial carrier, compared to Fourier analysis, is that it is difficult to remove a surface of unknown gradient, whereas in the Fourier domain this operation is a simple translation in frequency space. This problem may be alleviated by using an indexed mirror to generate the two reference beams,

although rigid body motion between the exposures can result in a slope being left on the data.

4.3 Conclusions on Fringe Analysis for Quantitative Flow Visualisation

The success of a holographic flow diagnostic system is strongly dependent on the performance of the fringe analysis algorithm used to process the data. The relative merits of the various algorithms are application specific and it is necessary to identify which attributes are important when studying flows. A successful system should strike a compromise between phase extraction accuracy and processing speed. The working environment is also an important consideration as wind tunnels are usually dusty and prone to vibration and air currents.

Fringe tracking and counting techniques are primitive and provide insufficient accuracy, reliability and autonomy for a viable system. These methods are still of value, however, when it is impossible to encode the data onto a spatial or temporal carrier wave. Also, as only a single image is required for processing, the reconstruction system does not require interferometric stability. In general, however, tracking methods should be used as a 'last resort' when other techniques have failed.

In contrast to fringe tracking, heterodyne interferometry provides extremely accurate phase extraction ($\sim$ 0.006 radians). The principal drawbacks of this method are that the reconstruction system must be completely isolated from environmental disturbances and the phase data is acquired sequentially. Although highly accurate, these restrictions make heterodyne interferometry unsuitable for implementation in production wind tunnel testing.

Quasi-heterodyne phase stepped fringe analysis offers greater potential for use in a fully automated reconstruction system. Unlike full heterodyne interferometry, this method acquires the image in parallel using established CCD camera and framestore technology. Although processing accuracy is compromised ($\sim$ 0.2 radians), these errors are small enough for the technique to be competitive with other flow diagnostics at transonic and supersonic speeds.

The principal disadvantage of phase stepped fringe analysis is that the reconstruc-

tion system is complicated by the use of two separate replay beams. These spatially divided wavefronts must remain interferometrically stable with respect to one another and must be incident at the correct angles to the hologram. In practice, it is extremely difficult to maintain sub-micron stability between two widely separated reconstruction beams. Consequently, phase extraction accuracy is typically poorer than one radian unless the data acquisition time is reduced to less than one second.

In addition to stability problems, the infinite fringe condition is hard to obtain during reconstruction of a hologram produced with wide angle reference beams, due to differences between the recording and replay geometries. Linear and parabolic phase differences between the reconstructed wavefronts may be removed by translating the collimating lenses of one of the replay beams, but it is difficult to identify when the infinite fringe condition has been achieved. Residual errors also remain due to shrinkage of the holographic medium during 'wet-processing' and because of aberration differences between the object waves. These problems are not unsurmountable and have been overcome in optical shop testing. Nevertheless, the operation of a phase stepping reconstruction system in a wind tunnel facility would require substantial dedicated hardware.

Fourier transform fringe analysis is analogous to heterodyne fringe processing, with the difference that the carrier wave is implemented in the spatial instead of the temporal domain. The attraction of this method is that it capitalises on the sophistication of current CCD, framestore and computer technology. At present, the space-bandwidth-product of CCD's is insufficient to process data as accurately as with temporal heterodyning, but current trends indicate that the size of CCD arrays will increase significantly in the next decade.

Unlike phase stepped fringe analysis, the Fourier technique requires only a single beam to replay the hologram. Consequently, the stability requirements of the reconstruction system are relaxed, and because the object waves diffracted from the hologram follow almost identical paths, a filtered thermal source may be used instead of a laser. In addition, rigid body movement of the recording apparatus may

be removed during fringe analysis as an integral part of frequency demodulation and therefore substantial wind tunnel vibration may be tolerated. Historically, failure to maintain the infinite fringe condition has restricted the use of interferometry in most working environments, but the use of specular illumination and Fourier analysis promises to overcome this limitation.

From this discussion it is apparent that Fourier transform fringe analysis is the most suitable candidate for processing interferometric wind tunnel images. The related spatial carrier phase shift technique also offers potential, with the additional advantage that it uses images of identical format to the Fourier method. The following section explores the potential of Fourier processing in greater depth and examines the accuracy of the resulting phase maps using simulated imagery.

5 Detailed Study of Fourier Transform Fringe Analysis

5.1 Introduction

The relative merits of the fringe analysis strategies reviewed in chapter 4 are summarised in table 3. The Fourier transform method is particularly attractive for processing holographic measurements in a wind tunnel test environment, because the associated reconstruction system has a common path geometry and therefore does not require vibration isolation. Also, finite fringe interferograms may be reconstructed from a double-exposure hologram using an inexpensive filtered thermal source, if an image plane configuration is adopted. The disadvantages of Fourier transform fringe analysis are that spatial resolution is low and the measurement accuracy is poorer than with phase stepped and heterodyne methods. However, as the algorithm operates in the spatial domain, it is likely that both of these limitations will be overcome as the space-bandwidth-product of CCD cameras increases.

Technique	Tracking	Heterodyne	Phase Stepped	Fourier
Accuracy	low	high	medium/high	medium
Need for Recording Linearity	low	low	high	medium
Resolution	low	high	high	medium
No. of Images	1	N/A	3+	1
Stability Required on Reconstruction	low	high	medium	low
Coherence Required on Reconstruction	low	high	high	low
Speed of Analysis	high	low	medium	medium
Potential for Future Improvements	low	low	medium	high
Complexity	low	high	medium	low
Capital Cost	low	high	high	low
Processing Domain	N/A	temporal	temporal	spatial

Table 3: Relative merits of fringe analysis strategies

This chapter quantifies the accuracy of Fourier transform fringe analysis. In particular, the effect of the following conditions on the resolution and accuracy of the

phase data is investigated.

- Spatial frequency of the carrier wave.
- Number of pixels in the imaging sensor.
- Geometry of the light sensitive elements in the sensor.
- Recording linearity.

The principal aim of this thesis is to determine the feasibility of obtaining three-dimensional density information, by reconstructing multi-directional interferometric data tomographically. Error analysis of the processed phase data is a vital step in achieving this objective.

5.2 Sampling Errors in Data Capture and Fourier Transform Fringe Analysis

5.2.1 Sampling Theory and Aliasing

The capture and analysis of fringe patterns involves numerous stages where data is sampled. For example, sampling occurs when the interferogram is detected by a discrete array of CCD sensors and when the analogue video signal is digitised by the analogue-to-digital converter in the framestore. The introduction of sampling errors may be illustrated by considering a line through a bandlimited image $i(x)$, which is sampled periodically with a spacing a (figure 38).

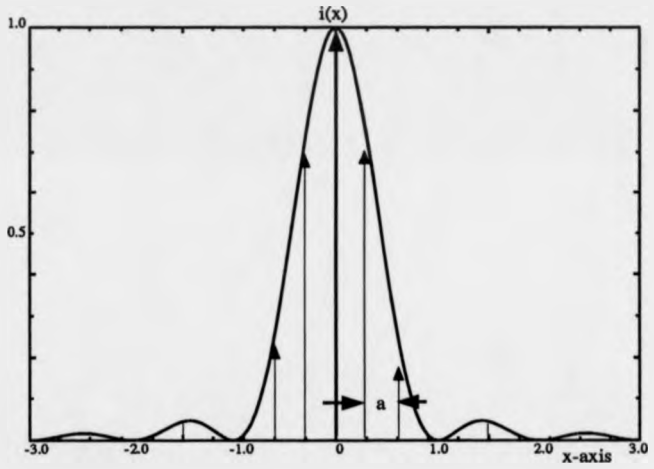


Figure 38: Periodic sampling of a bandlimited function

A squared sinc function (equation 87) has been chosen for illustrative purposes, because of the simplicity of its Fourier transform; however the analysis is valid for any bandlimited function.

$$i(x) = \text{sinc}^2(x) \tag{87}$$

Where the definition of $\text{sinc}(x)$ follows that used by Goodman [198]:

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \tag{88}$$

The sampled function $i_s(x)$ may be denoted mathematically by:

$$i_s(x) = i(x)\text{comb}(x/a) \tag{89}$$

Where the sampling function $\text{comb}(x)$ is defined by:

$$\text{comb}(x) = \sum_{i=-\infty}^{\infty} \delta(x - i)$$

From the convolution theorem, the Fourier transform of the sampled function $I_s(u)$ is the Fourier transform of the continuous function $I(u)$, convolved with the Fourier transform of the sampling function. This is described mathematically by equation 90, where $\otimes$ denotes the convolution operation.

$$I_s(u) = I(u) \otimes \text{comb}(au) \tag{90}$$

The convolution operation replicates the spectrum (figure 39), which in this instance is of triangular profile due to the choice of a squared sinc function for $i(x)$. Consequently, the bandlimited spectrum of the continuous function $I(u)$ is unrecoverable from $I_s(u)$ if adjacent harmonics overlap (first and second graphs of figure 39). When this occurs, high spatial frequency information is mapped back into regions of lower spatial frequency. This phenomena of spectral reflection is known as aliasing and it results in the visualisation of features with spatial frequencies exceeding the bandwidth of the imaging sensor. In fringe analysis, aliasing produces a moiré pattern between the interferogram and the pixels of the camera. If aliasing is ignored, these spurious fringes may be inadvertently processed so that errors are introduced into the final phase map.

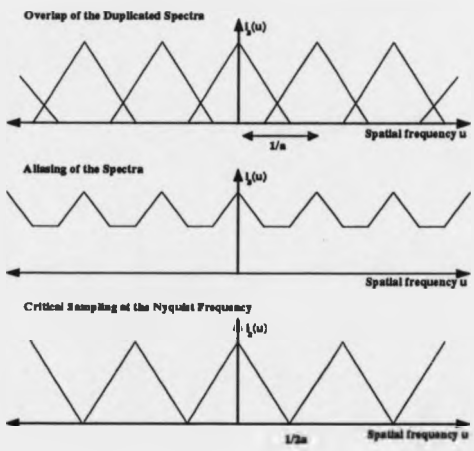


Figure 39: Aliasing of a signal due to overlap of the replicated spectral orders

If the sampling rate is chosen so that the replicated spectra are separated (last graph of figure 39), it is possible to retrieve the original function by spatial filtering in the Fourier domain, or by interpolation in image space. This is only possible if the function is bandlimited so that the spectra are of finite extent and do not overlap. When the function has a bandwidth f_b, it may be fully recovered if the spacing a between sampled values does not exceed $1/2f_b$. This result is stated formally in the Whitaker-Shannon sampling theory; and the critical sampling rate, known as the Nyquist frequency f_N, is given by:

$$f_N = \frac{1}{2a} \tag{91}$$

The accuracy of Fourier transform fringe analysis is reduced by aliasing, and therefore the magnification of the reconstructed image should be tailored to the pixel spacing of the imaging sensor. Aliasing also occurs if the bandwidth of the phase object exceeds the carrier frequency of the finite fringes. A theoretical investigation of the accuracy of Fourier transform fringe analysis has been made by Green [199]. A one-dimensional computer generated interferogram was used in the study (equation 92), with gaussian noise added to simulate the effect of speckle [199]. The zero spectral order was removed by a simple windowing operation as in Takeda's original algorithm

[195].

$$g(x) = 1 + \cos[2\pi f_x x + \phi(x)] + n(x) \tag{92}$$

Where,

f_x - carrier frequency

$\phi(x)$ - random phase function

$n(x)$ - gaussian noise distribution

The random phase function encoded by the interferogram, was synthesized by superimposing one hundred sinusoids, with spatial frequencies governed by a gaussian distribution of variable standard deviation. A gaussian distribution was also used in the generation of the additive noise.

The results from Green's study show that when the interferogram has a signal-to-noise ratio of two, the phase function can be retrieved accurately, provided that its bandwidth orthogonal to the finite fringes does not exceed 5% of the carrier frequency. When noise was omitted, phase information was retrieved with gradients approaching 30% of the carrier frequency. These results demonstrate the limitations of Fourier transform fringe analysis, when the zero order is removed by a simple windowing operation.

During Fourier transform fringe analysis, the high spatial frequencies in the phase object may become corrupted if the interferogram is not captured or processed correctly. For this discussion, positive spatial frequencies are defined as gradients lying in the same direction as the superimposed phase ramp. The three principal mechanisms for data corruption are:

- Overlap and subsequent interference of the symmetrically disposed positive and negative first orders (aliasing of the negative spatial frequencies).
- Corruption of the negative spatial frequencies during removal of the zero spectral order.

- Aliasing of positive spatial frequencies when the bandwidth of the imaging sensor is exceeded.

The degradation of encoded flow information, by overlap of the symmetrically disposed positive and negative first orders in the Fourier domain, is illustrated in figure 40. If the negative bandwidth of the phase object f_{-b} exceeds the carrier frequency by δf, then frequencies at f_{-b} interfere with those at $f_{-b} - \delta f$. Physically, a negative spatial frequency corresponds to a phase gradient which opposes the superimposed slope, so that the fringe spacing increases. When the phase object contains a negative phase gradient with a magnitude comparable to that of the finite fringes, a plateau is formed and an infinite fringe condition is obtained. If the phase gradient is increased further, so that the carrier frequency is exceeded, the slope reverses in direction and the fringe density begins to increase again. An ambiguity now arises because it is unknown whether the increased fringe density has been produced by a negative spatial frequency which is higher or lower than the carrier frequency. In Fourier space this ambiguity occurs when negative spatial frequencies from the positive first order, extend past the origin and are aliased back into the positive frequency domain.

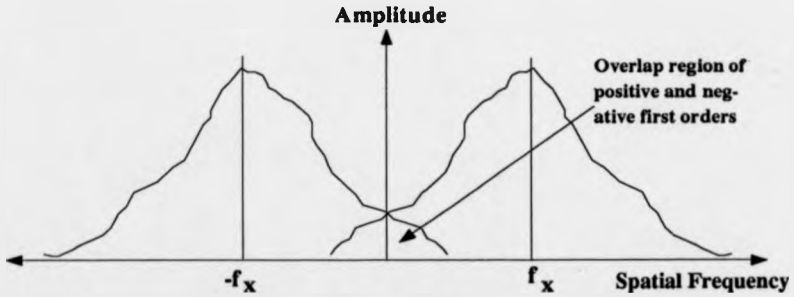


Figure 40: Corruption of high spatial frequencies by overlap of the first orders

Corruption of the negative spatial frequencies, by overlap of the positive and negative first orders, can be prevented if the phase object has a negative cut-off frequency with a magnitude less than the carrier frequency ($|f_{-b}| < |f_x|$). Hence, it is important that the carrier frequency is carefully matched to the negative bandwidth of the phase object, so that frequencies almost as high as f_x may be retrieved accurately. Aliasing

errors are most likely to occur around shocks and stagnation regions, where density gradients are largest.

Negative spatial frequencies in the phase object are also corrupted when the bias image $a(x, y)$ is filtered in the Fourier domain. Windowing methods are indiscriminate and result in the removal of flow information (figure 41). When the encoded phase information is subsequently demodulated, by translating the first order back to the origin, the corrupted data is located near to $|f_x|$. This noise appears as a large amplitude ripple in the phase map, with a periodicity closely matching the original carrier frequency. The severity of the noise indicates how well the noise term $A(u, v)$ has been removed, and the precision to which the carrier frequency has been matched to the bandwidth of the phase object. A low-pass filter may be used to remove the noise, and phase errors are minimised by choosing a cut-off frequency that suppresses most of the rippling, without appreciable loss of the higher frequencies. However, a reduction in image resolution is unavoidable and therefore a better solution is to subtract an estimate for $A(u, v)$. A detailed study of this approach is presented later in this chapter.

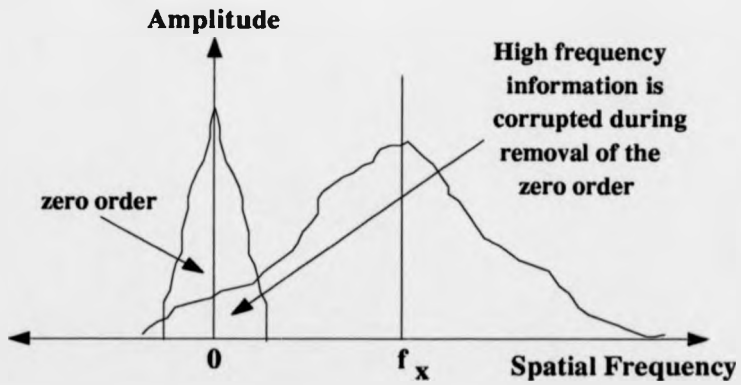


Figure 41: Corruption of high spatial frequencies during removal of the zero order

Windowing, and aliasing of flow information above the carrier frequency, affect only the negative spatial frequencies, which oppose the slope encoded on the data by the finite fringes. It would therefore appear that the phase gradients in the field should be aligned in the same direction as the phase ramp, so that only positive frequencies

are present. Unfortunately, this approach generates high fringe densities which are aliased by the imaging sensor. If the positive bandwidth of the phase object is f_b, the unperturbed finite fringes should be sampled at a frequency exceeding $f_x + f_b$ (four pixels/fringe if $f_x = f_b$). The sampling rate should be increased significantly when the interferogram is noisy due to laser speckle and parasitic interference.

Phase information is corrupted most severely around discontinuities in the interferogram. The worst affected regions are those surrounding the silhouette of a wind tunnel model and at the periphery of the fringe pattern. In these areas the amplitude noise terms $a(x, y)$ and $b(x, y)$ contain the highest spatial frequencies. This broadens the zero and first order spectra, so that they are more difficult to separate. Data corruption may be alleviated by interpolating into shadowed zones prior to Fourier processing. However, this process is complicated by the rapid undulations of the finite fringes and by irregularities in the boundary of the silhouette.

Errors around the perimeter of a processed interferogram originate because the finite window produces a convolution of the Fourier spectra with a two-dimensional sinc function. When the Fourier transform $I(u, v)$ is spatially filtered to remove the zero order term during fringe analysis, the negative side lobes of the sinc function exaggerate errors in the high frequencies of the phase object. In a review of generic image processing [200], Bates suggests that the effects of the image frame may be ameliorated by superimposing a window function over the data, so that the periphery is reduced smoothly to zero. In the Fourier domain this is equivalent to suppressing the side lobes of the sinc function, so that noise in the high spatial frequencies is reduced. However, windowing broadens the impulse response function and therefore, if the field-of-view is large, it is often simpler to discard the noisy data around the border.

Green's results [199], for the extraction of phase information from noisy interferograms, suggest that the carrier frequency of the finite fringe pattern should be twenty times larger than any frequencies contained in the phase object. Hence, if the carrier fringes have a period of four pixels, to support a positive bandwidth of f_x, the

smallest features which may be observed in the processed interferogram have a linear dimension of eighty pixels. If the captured image occupies 512×512 pixels, a size which is consistent with current CCD standard formats, these results suggest that the maximum flow area which may be studied simultaneously with a resolution of 1mm (typical width of an imaged shock wave) is 6.4mm×6.4mm. Therefore, it appears that Fourier transform fringe analysis is of limited use on whole-field wind tunnel imagery, unless the space-bandwidth-product of the recording system is increased by at least two orders of magnitude (providing a field-of-view of 32cm×32cm).

Suitable CCD cameras are now becoming commercially available with 4096×4096 elements (models marketed by Kodak and Loral Fairchild Imaging Sensors). However, these devices are expensive and require dedicated framestores. Also, large data arrays are generated when processing images of this size (134Mbyte for a 4096×4096 single precision complex floating point array). These data files are difficult to store, manipulate and process using readily available computer hardware.

A preferable solution to using elaborate hardware is to optimise the Fourier transform fringe analysis algorithm. This may be achieved by carefully matching the carrier frequency and pixel spacing, to the positive and negative cut-off frequencies of the phase object. If this is not achieved then the space-bandwidth-product of the interferogram and imaging sensor may be under-utilised or, at worst, the phase data may be corrupted by aliasing.

The conclusions drawn by Green are highly specific to the imagery used during his study and the computer generated noise is unrepresentative of that present in specular illumination fringe patterns. Also, the bias image $a(x, y)$ was removed by a windowing operation in the Fourier domain, whereas greater accuracy could have been obtained by subtracting an estimate for $A(u, v)$. Therefore, in order to obtain a greater understanding of the limitations of Fourier transform fringe analysis, an investigation was made using imagery recorded during a wind tunnel test. Phase errors resulting from this analysis are more realistic than those obtained when processing completely artificial interferometric data. The results from this study are presented in section 5.3,

alongside a discussion of the sources of experimental and computational error.

5.2.2 Fringe Degradation due to Sampling by a CCD Camera

The intensity distribution of a finite fringe pattern may be expressed by:

$$i(x,y) = a(x,y) + b(x,y)\cos[\phi(x,y) - k_x x] \tag{93}$$

Where in this definition, the carrier frequency f_x has been expressed as an angular spatial frequency k_x in the x-direction.

When this image is captured and digitised by a CCD camera and framestore, the intensity distribution is modified due to recording non-linearity, intensity quantisation, and spatial sampling. With a Cartesian co-ordinate system centred on the middle of the image, the actual function recorded $i'(x,y)$ is more accurately described by:

$$i'(x,y) = h(x,y) \otimes \mathcal{L}[i(x,y)]\text{rect}(x/ma)\text{rect}(y/na) \tag{94}$$

With $\text{rect}(x)$ defined by:

$$\text{rect}(x) = \begin{cases} 1 & \text{if } |x| \leq 1/2 \\ 0 & \text{otherwise} \end{cases}$$

and,

$h(x,y)$ - point spread function
$\mathcal{L}$- operator describing the non-linear response of the CCD elements
m,n - number of pixels in the x and y directions respectively
a - spacing between pixels (assumed equal in both directions)

This mathematical model denotes the non-linear response of the camera by a quantised operator $\mathcal{L}$, which is assumed to be spatially invariant (all elements respond equally). Detector non-linearity encompasses the effects of saturation, dark level response, blooming and incomplete charge transfer. The resolving power of a CCD

array may be characterised by the point spread function $h(x, y)$, which quantifies how the phase and amplitude of an image point are modified by the sensor.

During Fourier transform fringe analysis, the two-dimensional Fourier transform of the sampled image is calculated, and by use of the convolution theorem, the result may be described by:

$$I'(u, v) = OTF(u, v)\mathcal{F}\{\mathcal{L}[i(x, y)]\} \otimes \text{sinc}(mau) \otimes \text{sinc}(nav) \tag{95}$$

The effect of the image frame is to convolve the spectra with a two-dimensional sinc function as described by Bates [200]. The quantity $OTF(u, v)$ is the optical transfer function of the system and is mathematically the Fourier transform of the point spread function. This complex valued function is separable into amplitude and phase components, which are known as the phase transfer function (PTF) and the modulation transfer function (MTF). The MTF is useful for quantifying the spatial frequency response of optical systems and devices. In practice, it is determined by measuring the output contrast $MTF(k_x)$ of a sinusoidal input function $i_{ip}(x)$, as a function of spatial frequency:

$$MTF(k_x) = \frac{i_{max}(k_x) - i_{min}(k_x)}{i_{max}(k_x) + i_{min}(k_x)} \tag{96}$$

Where the input function is defined by:

$$i_{ip}(x) = 1 + \sin k_x x \tag{97}$$

and i_{max} and i_{min} signify the extreme intensity values of the output signal.

The definition of MTF is ambiguous, however, when discrete imaging devices are considered. This ambiguity is particularly relevant to image processing, because discrete element CCD's are gradually replacing beam-scanned photoconductive Vidicon

cameras in imaging systems. CCD's are usually fabricated using silicon MOS technology and the light sensitive elements occupy a finite area so that sampled values are spatial averages. The two most common CCD configurations are frame transfer and interline transfer geometries (figure 42).

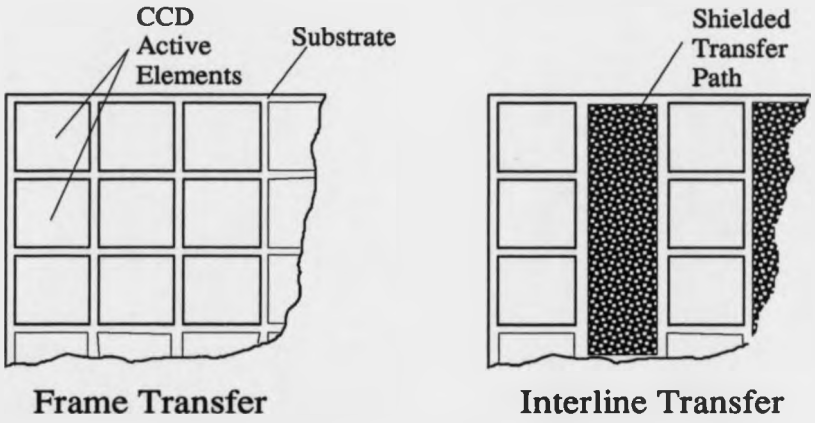


Figure 42: Frame transfer and interline transfer CCD substrate geometries

The frame transfer CCD has a sensor substrate that is divided into two identical areas. One of these, known as the image section, consists of closely packed photo-sensitive elements. The remainder of the device forms a masked storage area of parallel CCD shift registers, separated by insulating walls. After the charge has accumulated at the active sites in the image section, clock voltages are applied to the sensor electrodes, to transfer the picture to the store. With the frame transfer geometry this process is relatively slow and takes several hundred microseconds.

With the interline transfer geometry, the masked CCD shift registers are located adjacent to the active elements. This results in a reduced spatial resolution along one of the camera's axes and a lower sensitivity. The principal advantage of this configuration is that the charge may be transferred more quickly ($\sim 1\mu s$). Consequently, the interline CCD camera is potentially better suited to real-time image processing applications, such as on-line fringe analysis, where rapid data acquisition is necessary. However, the effect of sampling the image at alternate pixels must be considered.

Manifestations of discretised image capture may be illustrated by considering a one-dimensional sinusoidal intensity profile, which has a maximum centred on the middle of a pixel and a minimum located mid-way between pixels (curve B of figure 43). If the spatial frequency of the input sinusoid B is increased or decreased infinitesimally, the quantised location of the minima jumps discontinuously between pixels (curves A and C of figure 43). Hence, a multi-valued MTF curve is generated, where the MTF at a particular spatial frequency is a function of the phase relationship between the input sinusoid and the pixel array.

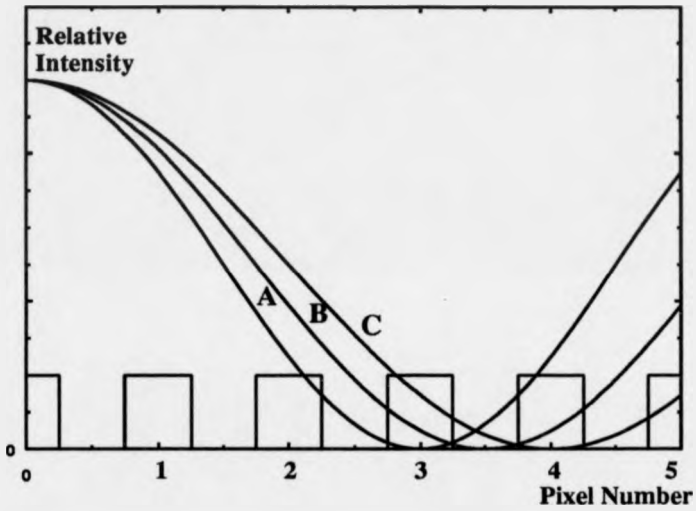


Figure 43: Formation of a multi-valued MTF curve due to spatial discretisation

The sampling of a one-dimensional sinusoidal fringe pattern, by a linear pixel array, may be examined theoretically by considering the charge accumulated by the nth pixel. If the total charge is assumed to be proportional to the integrated light intensity falling on the one-dimensional pixel, the output signal $i_{op}(n)$ is given by:

$$i_{op}(n) = \int_{na-b/2}^{na+b/2} ci_0[1 + \sin(k_x x + \psi)]dx \tag{98}$$

Where the origin has been arranged to bisect the pixel designated n=0, and the other parameters are defined by:

a - inter-pixel spacing
b - width of a pixel ($b \leq a$)
c - proportionality constant
i_0 - amplitude of the input sinusoidal intensity pattern
k_x - spatial frequency of the fringe pattern in lines per unit length
ψ - phase difference between the pixel array and the periodic test function

The solution of equation 98, for pixels spaced by a along the x-axis, is given by:

$$i_{op}(x) = ci_0 b\,\mathrm{comb}(x/a)\left\{1 + \frac{2}{k_x b}\left[\sin\left(\frac{k_x b}{2}\right)\sin(k_x x + \psi)\right]\right\} \tag{99}$$

Inspection of the input function (equation 97) and the response (equation 99), reveals that when fringes are sampled by pixels of finite area, the fringe profile remains unchanged but its contrast is reduced by a factor $m(k_x)$:

$$m(k_x) = \mathrm{sinc}\left(\frac{k_x b}{2\pi}\right) \tag{100}$$

In the limit where the width of the CCD elements tends to zero, the fringe modulation envelope is unity for all spatial frequencies and the input function may be fully recovered if the sampling rate exceeds the Nyquist frequency.

With a spacing a between sampled points, the angular Nyquist frequency k_N is π/a. Consequently, because the pixel width b is less than or equal to the pixel spacing a, frequencies which are not aliased lie in the principal lobe of the sinc function defined by equation 100 (figure 44). Hence, when a detector is used with discrete elements, the spatial frequency spectra is low-pass filtered, with a cut-off frequency at π/b. If the Nyquist frequency is exceeded, the modulation envelope $m(k_x)$ is oscillatory due to the variation in phase between the pixels and the input fringe pattern. This oscillatory behaviour gives rise to moiré fringes and the visualisation of periodic functions with frequencies exceeding the Nyquist limit [201].

Although the modulation term $m(k_x)$ provides an indication of the MTF of a fringe acquisition system, this analysis does not consider the effect of varying the ratio a/b,

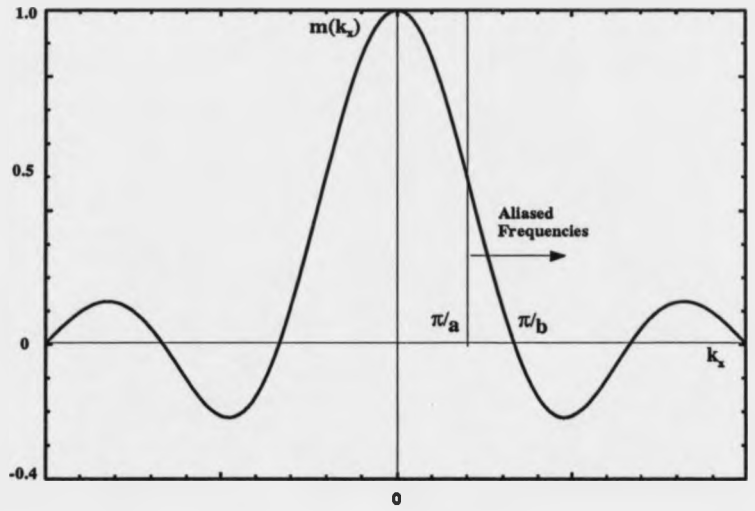


Figure 44: Reduction of fringe modulation by discrete pixel sampling

or the importance of the phase relationship ψ between the input function and the pixel array. To determine the MTF explicitly it is necessary to calculate the maximum and minimum values of the modulated sinusoidal fringes for these different conditions.

A rigorous analysis of how the discrete sampling process results in a multi-valued MTF function is given by Feltz and Karim [202]. In their study they calculate the extrema of these MTF curves and speculate that the mean of these graphs is a good guideline for describing the resolution of a device. Their results show that for a frame transfer CCD, with no gaps between pixels, the mean MTF is 40% when the input sinusoid is sampled at the Nyquist frequency (two pixels/fringe). In comparison, for the interline transfer CCD, with the input sinusoid spanning two pixels and two vacancies, the mean MTF is increased to 57%.

Hence, the interline geometry is better suited to image processing in terms of both speed and sampling accuracy. The only disadvantage is that the pixel density is bi-directional and therefore an image may need to be over-sampled in one direction to ensure adequate sampling in the orthogonal direction. However, in Fourier transform fringe analysis this configuration is desirable, because a higher sampling rate is required in the direction orthogonal to the fringes.

5.2.3 Conclusions on the Sampling of Finite Fringe Interferograms

In conclusion, there are two conditions which must be satisfied to enable accurate extraction of phase data from an interferogram. Firstly, the carrier frequency should exceed the negative bandwidth of the phase object so that the phase is monotonically increasing across the interferogram. Aliasing occurs when this requirement is not met, and an ambiguity arises between phase gradients that are higher or lower than the carrier frequency. Aliasing may be prevented by ensuring that the fringe spacing does not increase by more than 50% of the carrier fringes, thereby limiting the negative cut-off frequency of the phase object to $0.5f_x$. An additional benefit of imposing this bandlimit is that the zeroth and first orders are distinct, so that they may be separated without data corruption.

The second requirement is that the interferogram should be sampled at a rate which exceeds the sum of the carrier frequency and the positive bandwidth of the phase object. Aliased moiré fringes are observed if the pixel sampling rate is insufficient, and processing errors may result. Fringe analysis accuracy is higher if an interline CCD geometry is used, so that there are gaps between the sampled values. The bi-directional resolution of this device may be tailored to the orientation of the finite fringes and data acquisition rates are higher than with frame transfer geometries. When an interferogram is sampled at the Nyquist frequency, the fringe contrast is reduced to approximately 57% due to the finite area of the pixels. Consequently, the sampling rate should be increased if fringe contrast is reduced by noise. Aliasing of positive spatial frequencies in a phase object may be avoided by ensuring that the fringe contrast exceeds 60% at all points in the interferogram.

To extract the most information from an interferogram for a particular imaging sensor, and thereby maximise the field-of-view that may be analysed simultaneously, the Fourier transform fringe analysis algorithm should be optimised so that image information can be retrieved with frequencies of approximately $0.5f_x$. Green's results [199] suggest that when coherent noise is present in the interferogram, the minimum feature size that may be resolved in the phase map is eighty pixels in length. How-

ever, if noise can be filtered without corrupting the encoded data, it is feasible that a resolution approaching the theoretical limit set by information theory can be obtained (two pixels). Optimisation of the phase encoding process forms the basis of section 5.3, which investigates the accuracy of Fourier transform analysis using simulated interferometric data.

5.3 Characterisation of Fourier Transform Fringe Analysis using a Synthetic Interferogram

5.3.1 Production of a Synthetic Interferogram Combining both Experimental and Computational Data

The accuracy of Fourier transform fringe analysis, may be assessed by calculating the phase errors that occur when the amplitude noise terms $a(x,y)$ and $b(x,y)$ are removed by spatial filtering. It is difficult to model the characteristics of $a(x,y)$ computationally, because the background intensity distribution of an interferogram is a complicated function of illumination uniformity, object reflectance or transmittance, and coherence between the object and reference beams. Also, any synthetic image must incorporate speckle noise, diffraction and optical aberrations. In practice, it is simpler to use an experimental interferogram, which has been modified by image processing, as a test case for Fourier transform fringe analysis.

The corruption of phase information by Fourier transform fringe analysis, can be calculated by analysing the interferogram produced by a known phase object. Experimental fringe patterns, which exhibit typical variations in bias level and fringe contrast, are unsuitable because the phase object is not known sufficiently accurately from independent measurements. A better solution is to generate the phase function computationally, so that its values may be compared accurately to those obtained by the processing algorithm. A fringe pattern can be produced that exhibits experimental noise characteristics, while possessing a pre-defined phase function, by combining experimental amplitude data with a computer generated phase object.

A synthetic fringe pattern may be generated by considering the formation of an interferogram in a holographic interferometer. When reconstructed, a double-exposure specular illumination hologram of an air flow, diffracts two optical wavefronts from the incident beam: one which has been encoded with phase information describing the optics of the wind tunnel, and a second that has experienced an additional phase retardation due to the flow field. These co-axial waves may be recombined by an

imaging system to produce an infinite fringe interferogram. Alternatively, a finite fringe pattern may be generated by encoding the object waves o_1 and o_2 with different reference beam angles, so that the phase data is superimposed on a ramp. The interference pattern formed by these beams may be described by a sinusoidal, phase modulated, intensity distribution:

$$i(x,y) = |o_1|^2 + |o_2|^2 + 2o_1o_2\cos[\phi(x,y) + k_x x] \tag{101}$$

Where the sum of $|o_1|^2$ and $|o_2|^2$ constitutes the bias level $a(x,y)$, and the factor $2o_1o_2$ describes the fringe modulation $b(x,y)$.

A finite fringe pattern of the transonic flow over a NACA 0012 aerofoil is shown in figure 36 (see chapter 6 for details of the test). The reconstructed interferogram was viewed by a CCD camera and converted into a digital data file by a framestore. The carrier fringes were sampled at twice the Nyquist frequency (128 fringes across the 512 pixel image width) to enable the encoding of positive phase gradients up to a maximum spatial frequency of f_x. The negative bandwidth of the phase object was restricted to approximately -0.5f_x, except in the stagnation region at the leading edge of the aerofoil, to prevent aliasing and facilitate removal of the zero spectral order during Fourier processing.

A region of low fringe density, for instance as shown at the leading edge of the aerofoil in figure 45, indicates that the opposing phase gradient exceeds that of the carrier fringes. It is likely that phase ambiguities exist in this interferogram and the carrier frequency would need to be increased, or the direction of the superimposed phase ramp reversed, to prevent aliasing. Figure 45 also shows fringes that are inclined at angles less than 45° to the horizontal. This suggests that phase gradients parallel to the finite fringes also exceed the carrier frequency. Although this condition is acceptable, it provides further evidence that large phase gradients exist orthogonal to the finite fringes.

If there are very large density gradients in a flow field, then an excessively high

imaging system to produce an infinite fringe interferogram. Alternatively, a finite fringe pattern may be generated by encoding the object waves o_1 and o_2 with different reference beam angles, so that the phase data is superimposed on a ramp. The interference pattern formed by these beams may be described by a sinusoidal, phase modulated, intensity distribution:

$$i(x,y) = |o_1|^2 + |o_2|^2 + 2o_1o_2\cos[\phi(x,y) + k_x x] \tag{101}$$

Where the sum of $|o_1|^2$ and $|o_2|^2$ constitutes the bias level $a(x,y)$, and the factor $2o_1o_2$ describes the fringe modulation $b(x,y)$.

A finite fringe pattern of the transonic flow over a NACA 0012 aerofoil is shown in figure 36 (see chapter 6 for details of the test). The reconstructed interferogram was viewed by a CCD camera and converted into a digital data file by a framestore. The carrier fringes were sampled at twice the Nyquist frequency (128 fringes across the 512 pixel image width) to enable the encoding of positive phase gradients up to a maximum spatial frequency of f_x. The negative bandwidth of the phase object was restricted to approximately -0.5f_x, except in the stagnation region at the leading edge of the aerofoil, to prevent aliasing and facilitate removal of the zero spectral order during Fourier processing.

A region of low fringe density, for instance as shown at the leading edge of the aerofoil in figure 45, indicates that the opposing phase gradient exceeds that of the carrier fringes. It is likely that phase ambiguities exist in this interferogram and the carrier frequency would need to be increased, or the direction of the superimposed phase ramp reversed, to prevent aliasing. Figure 45 also shows fringes that are inclined at angles less than 45° to the horizontal. This suggests that phase gradients parallel to the finite fringes also exceed the carrier frequency. Although this condition is acceptable, it provides further evidence that large phase gradients exist orthogonal to the finite fringes.

If there are very large density gradients in a flow field, then an excessively high

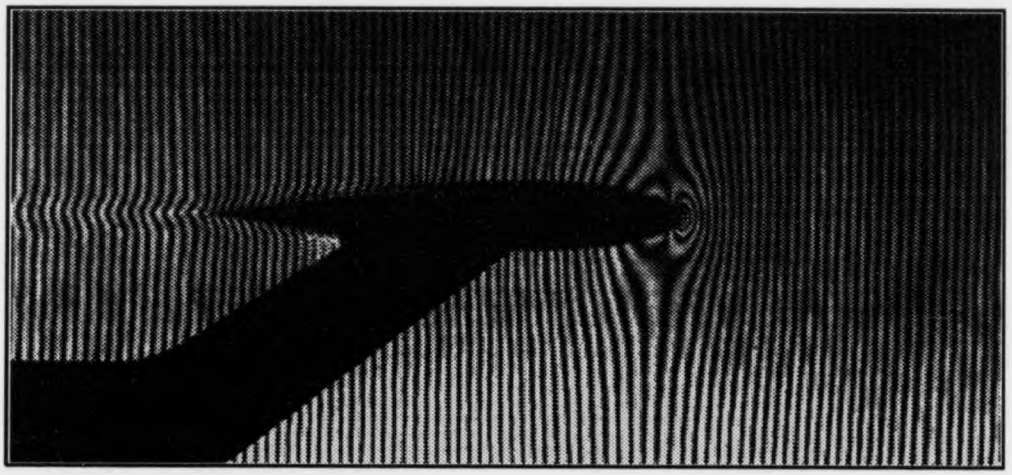


Figure 45: Aliasing of phase information around the leading edge of an aerofoil

carrier frequency is required to encode the phase data. However, significant refraction occurs in these regions and therefore the optical path length data is erroneous if rectilinear propagation is assumed. Consequently, zones containing highly refracted rays, such as the stagnation region at the leading edge of an aerofoil, should not be used in determining the optimum carrier frequency. After fringe analysis, the phase data calculated from these regions should be discarded. Refracted rays can be identified in the finite fringe reconstruction by inserting a knife edge into the back focal plane of the hologram to produce a schlieren image (figure 46).

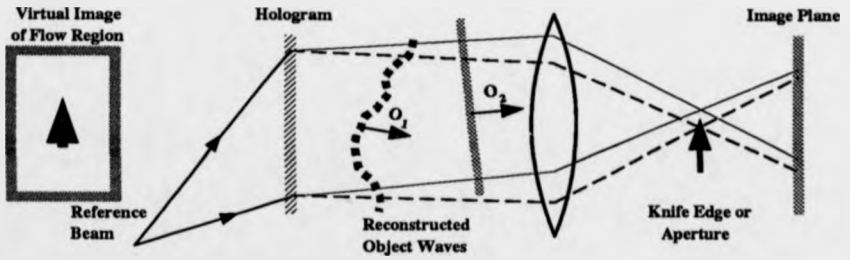


Figure 46: Optical system used to obtain schlieren images and the bias image $a(x, y)$

A synthetic interferogram, comprised of experimental amplitude and computer generated phase data, can be generated by removing the finite fringes from figure 36. This was achieved in the optical reconstruction system, by spatially filtering one of the component object waves in the back focal plane of the hologram (figure 46). The intensity distribution $|o_1|^2$ of the remaining wavefront was projected onto a CCD camera and digitised by a framestore. This process was repeated for the second object wave and an estimate for $a(x, y)$ was obtained by adding these two images.

The resulting bias image (figure 47) contains information about the model's silhouette, variations in illumination, and the spatial distribution of noise artefacts. A residual finite fringe pattern remains, because it was difficult to completely filter each of the component beams due to astigmatism of the reconstruction system. The fringe modulation is weak, however, and parasitic interference of this magnitude is typical in experimental interferograms.

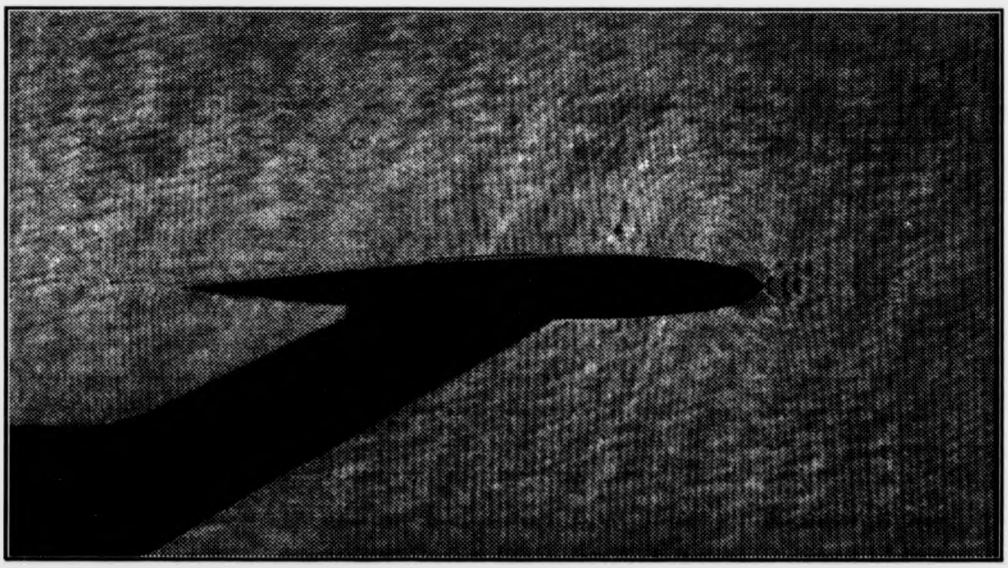


Figure 47: Optically derived bias image $a(x, y)$

A Panasonic WV-CD50 CCD camera, comprised of 500×582 elements, was used to view the interferogram. The analogue video signal from the camera was subse-

quently captured and converted into a 512×512 8-bit digital image by an ITEX PC Vision Plus framestore, controlled by a personal computer (PC) with an 80286 Intel microprocessor. This image size was chosen because its dimensions are a power of two, which is a requirement for the Fast Fourier transform (FFT) program, and it is the largest size that can be processed in convenient run-times on a SUN4 workstation ($<$ 1min).

The software controlling the framestore produced 0.25Mbyte image files with an ITEX header. These images were sent from the host PC to a SUN4 workstation via an Ethernet link and converted into standard SUN raster files by substitution of the correct header. A SUN raster format was chosen because of its compatibility with a suite of existing image processing routines on the local network. In particular, this was useful for displaying the images and for interfacing to an IBM FFT program. Single precision (4 byte) floating point complex (FPC) numbers were used at intermediate processing stages, to minimise rounding errors and enable complex arithmetic. The captured images were stretched by a factor of 1.55 in the vertical direction due to the aspect ratio and number of pixels in the CCD array. This distortion was removed by linear interpolation during image processing.

The offset and gain of the framestore were programmed to give a zero grey level output for a dark input condition, and an output of 255 when the sensors were saturated. CCD elements do not respond equally when the camera is uniformly illuminated, and the output intensity values are spread over a normal distribution curve [203]. Consequently, the extremes of the dynamic range were set to the centre of this distribution. Calibration of the camera using extreme illumination conditions ensures that the full dynamic range is used; although this is obtained at the expense of recording linearity. Fourier transform fringe analysis computes the phase of the flow field from the deviation of the finite fringe centres and therefore it was thought that recording linearity would be less important than for phase stepped fringe analysis. Also, saturation of a periodic signal results in the generation of spectral harmonics, which can be filtered in the Fourier domain. Nevertheless, some loss in accuracy is likely due to the increased

noise in the Fourier domain and therefore these errors are investigated as part of this study.

A computer generated phase-only resolution chart, consisting of sinusoids with a peak-to-peak phase excursion from $\pi/8$ to $3\pi/4$ and spatial frequencies $k_x/8$, $k_x/4$, $k_x/2$ and k_x, was combined with figure 47. Registration of the phase object with respect to the experimental amplitude data is shown in figure 48 and the profile of a section of the charts is shown in figure 49. This particular phase domain was chosen because it includes only principal values and therefore does not require the implementation of a phase unwrapping algorithm. In addition, it enables the study of phase errors at $\pi/2$ where the tangent function asymptotes to infinity. Spatial frequencies in the range $k_x/8$ to k_x were chosen to investigate the loss in phase measurement accuracy that occurs as the carrier frequency is approached.

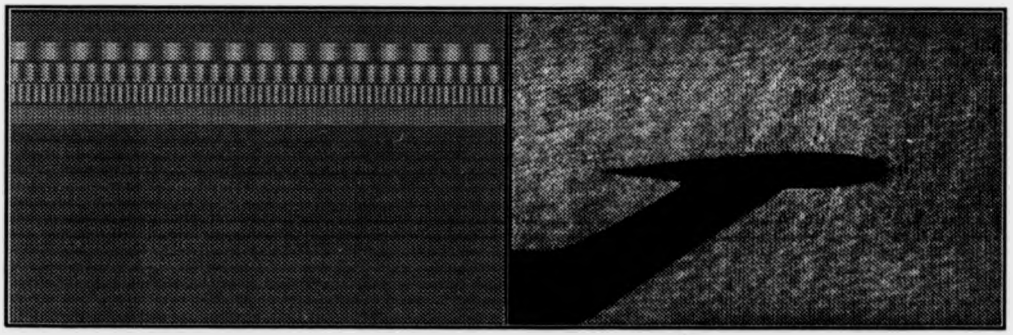


Figure 48: Phase and amplitude components of the synthetic fringe pattern

An FPC file, simulating the encoding of integrated flow information onto an object wavefront, was formed by combining the phase chart with the square root of the bias image $a(x,y)$. The real part of the array was assigned a value $\sqrt{a(x,y)}\cos\phi$ and the imaginary part $\sqrt{a(x,y)}\sin\phi$. A second FPC file describing a linear phase ramp $k_x x$ with amplitude $\sqrt{a(x,y)}$, was added to the first FPC file to simulate the formation

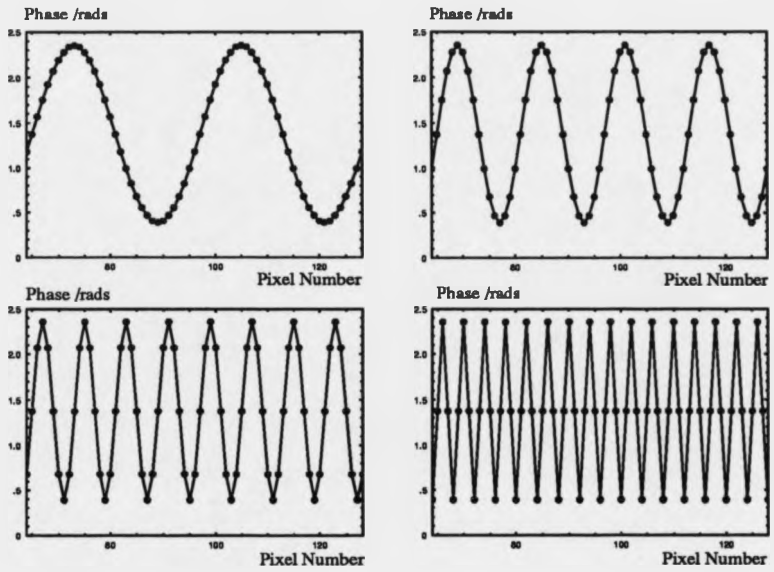


Figure 49: Sinusoids used to form the phase-only resolution chart

of a dual reference beam holographic interferogram. A synthetic finite fringe pattern was subsequently generated by multiplying the resultant FPC file by its complex conjugate to form an array of wholly real numbers (figure 50). This final calculation is analogous to the incoherent detection of a fringe pattern.

5.3.2 Aliasing of Encoded Phase Information

The phase errors introduced during Fourier transform fringe analysis are summarised in the flow chart of figure 51. The accuracy of Fourier transform fringe analysis was assessed at different spatial frequencies, by examining the restoration of the four sinusoidal phase functions (figure 49). To isolate aliasing errors from those produced by spatial filtering, the bias image $a(x, y)$ used to generate the interferogram was Fourier transformed and subtracted from the spectrum of the fringe pattern $I(u, v)$. The results from this study are applicable when the bias image is known to single bit accuracy. Subtraction of $A(u, v)$ from $I(u, v)$ introduces noise, because 8-bit raster images were used to represent both the interferogram $i(x, y)$ and the bias image $a(x, y)$.

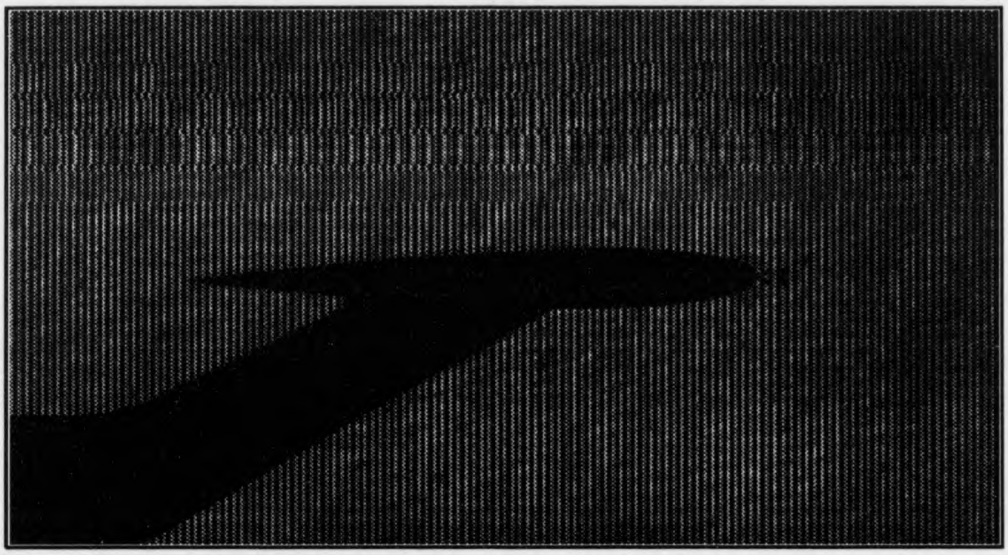


Figure 50: Synthetic finite fringe interferogram with phase encoded resolution chart

The dynamic range of the interferogram is larger than the bias image and therefore it contains larger quantisation errors after digitisation. The use of floating point numbers would have reduced these errors, but raster files had to be used to simulate the capture of experimental images.

The bias image was captured with the peak intensity value adjusted to a grey level of 255, so that the entire dynamic range of the framestore was used. Similarly, the synthetic interferogram was also scaled to have grey levels in this range. Consequently, the relative intensities of the interferogram and the estimate for the bias image were incorrect, and the Fourier transforms $I(u,v)$ and $A(u,v)$ had to be normalised prior to subtraction. The normalisation constant was found by calculating the median value of $I(u,v)/A(u,v)$ for the central 10×10 pixels.

The sequence of operations used to calculate the wrapped phase map are summarised below:

1. $A(u,v)$ was normalised with respect to $I(u,v)$ to produce $A_N(u,v)$.

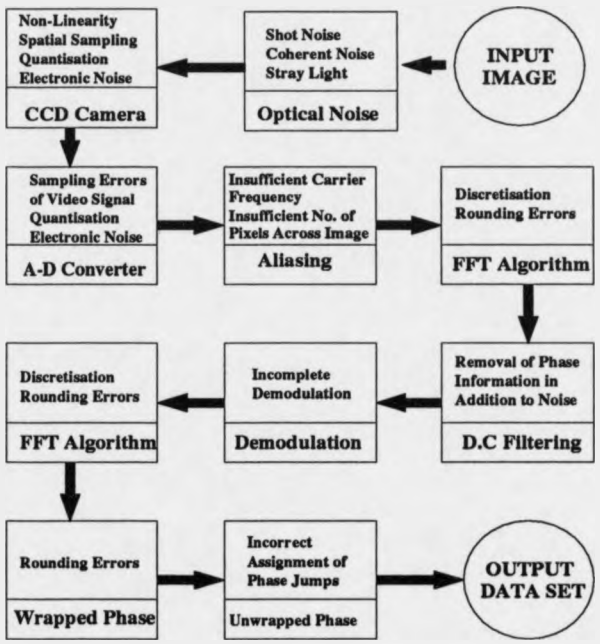


Figure 51: The introduction of phase errors during Fourier transform fringe analysis

2. $A_N(u,v)$ was subtracted from $I(u,v)$.

3. The central 3×3 pixels in the resulting spectrum were set to zero.

4. Half of the spectrum with $u < 0$ was discarded.

5. The flow information was demodulated by translating the first order back to the origin (shift of 128 pixels).

6. The inverse Fourier transform of the modified spectra was calculated.

7. The ratio of the imaginary and real components of the complex image were calculated to remove the noise term $b(x,y)$.

8. The inverse tangent of the data was calculated using the FORTRAN ATAN2 function.

Phase unwrapping was unnecessary because the resulting values lay in the range

$[\pi/8,3\pi/4]$. The accuracy of Fourier transform fringe analysis was subsequently assessed by taking a line-scan through each of the reconstructed sinusoids and comparing the data points to the corresponding input profiles. Phase values within 64 pixels of the image frame were discarded due to the increased noise in these areas. The accuracy of the phase retrieval process was assessed by calculating the rms error e_{rms} between 64 input and output phase values:

$$e_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\{[\Phi(ip)_i - \Phi(op)_i]^2\}} \tag{102}$$

Where,

$\Phi(ip)_i$ - input phase values
$\Phi(op)_i$ - output phase values
N - total number of phase values compared (64)

The line scans through the centres of the restored sinusoids are shown in figure 52 and the rms errors are presented in table 4. These results demonstrate that the accuracy of Fourier transform fringe analysis is a function of spatial frequency. From the limited set of spatial frequencies investigated, the highest reconstruction accuracy is 0.037 radians at a spatial frequency of $0.25f_x$. Errors increase for both lower and higher spatial frequencies, due to the noise term $A(u,v)$ at the origin and $B(u,v)$ at the first order. The low spatial frequencies, which are located close to f_x prior to demodulation, are reconstructed more accurately than high frequencies, because they are corrupted less during removal of $A(u,v)$. When the phase values at the edge of the field were processed (first 64 pixels), the rms error at 0.125 f_x was increased from 0.05 radians to 0.14 radians. This result illustrates the reduction in fringe analysis accuracy that occurs at the periphery of the phase map.

Wavelength /pixels	32	16	8	4
rms phase error /rads	0.050	0.037	0.096	0.669

Table 4: Phase extraction accuracy when an estimate for $A(u,v)$ has been subtracted

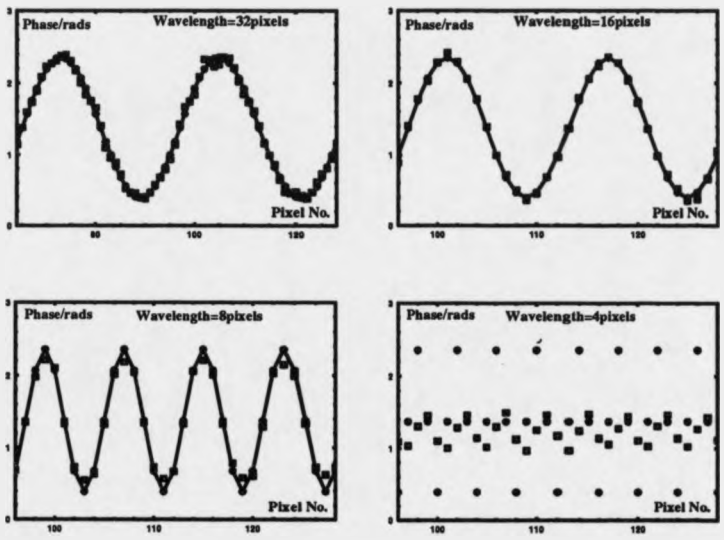


Figure 52: Comparison of the processed and original fringe profiles at different spatial frequencies

The extraction of phase information with spatial frequencies close to f_x, was investigated by repeating the calculations for sinusoids with wavelengths of 5,6 and 7 pixels. Fringe analysis accuracy is a function of the phase object's location in the field and therefore the three runs were done sequentially, with the sinusoid replacing the third chart in figure 49. The results (table 5) show that the rms phase error is approximately constant over the range $[0.57f_x, 0.8f_x]$, with a value of 0.130 ± 0.005 radians ($\sim$ 1/50th of a fringe). Hence, if the bias image $a(x,y)$ is known accurately, then it should be possible to use most of the bandwidth provided by the carrier fringes and the imaging sensor.

Wavelength /pixels	8	7	6	5
rms phase error /rads	0.096	0.129	0.124	0.136

Table 5: Phase extraction accuracy close to the carrier frequency

In the wind tunnel tests described in this thesis, a field-of-view measuring ap-

proximately 25cm×25cm was used and the images were digitised to a resolution of 512×512 pixels. The results from this study of Fourier processing suggest that if an accurate estimate for $A(u, v)$ is available, then flow features as small as 1.2mm may be calculated to an accuracy of 0.13 radians. A phase error of 0.13 radians represents a measurement accuracy of 0.1% when the flow field retards or advances the wavefronts by five wavelengths (typical result of the wind tunnel tests described in this thesis). In comparison, Green's method of processing would have provided a spatial resolution of only 39mm [199].

Unfortunately, in many experimental environments it may be difficult to obtain the bias image as accurately as in this computer simulation, due to camera non-linearity and coherent noise. A phase map resolution of ~1mm is desirable, because it is adequate for validating CFD predictions and is the approximate width of a shock wave. A resolution exceeding 1mm is unlikely to provide more flow information, because the phase measurements are path integrated and therefore localised features in the three-dimensional phase object are blurred by spatial averaging. In practice, the bias image cannot be determined to single bit accuracy, because of recording non-linearity and incomplete spatial filtering of the finite fringes. Consequently, to obtain phase maps with a resolution approaching 1mm, without recourse to piece-wise analysis of the field-of-view, it is necessary to investigate methods for optimal removal of the zero spectral order.

5.3.3 Optimal Removal of the Zero Order Bias Term

With Takeda's original Fourier transform fringe analysis algorithm [195] it is assumed that spatial frequency variations in the phase object are small compared to the carrier frequency. If this condition is satisfied, the zero order term can be removed by windowing, without the loss of flow information (figure 53). In practice this can only be achieved if the phase object is highly magnified, so that it contains spatial frequencies which are less than the carrier frequency. Efficient use of the space-bandwidth-product of the interferogram requires the extraction of spatial frequency information close to

the carrier frequency. However, these frequencies overlap the zero order and indiscriminate windowing is likely to introduce errors.

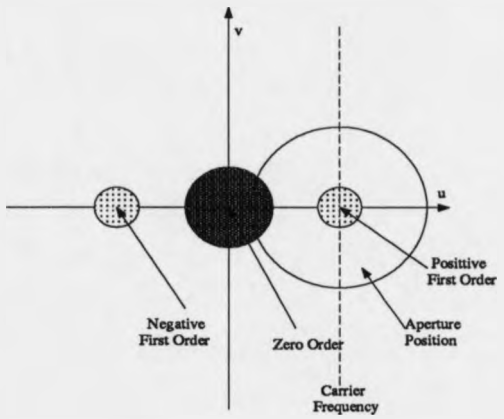


Figure 53: The use of windowing to isolate the first spectral order

The relative accuracy of removing $A(u, v)$ by windowing or subtractive techniques, was assessed using the synthetic finite fringe interferogram encoded with spatial frequencies in the range $[k_x/8, k_x]$. Figure 54 shows line scans through the processed phase map, when a circular aperture of radius 96 pixels was used to extract the first order in the Fourier domain. This window imposes a cut-off frequency at $0.75 f_x$ after the phase object has been demodulated and therefore all of the sinusoids, except the one at the Nyquist frequency, should be unaffected. The accuracy of the phase extraction process is summarised in table 6 as a function of spatial frequency.

Wavelength /pixels	32	16	8	4
rms phase error /rads	0.075	0.062	0.141	0.702

Table 6: Phase extraction accuracy when $A(u, v)$ has been removed by windowing

A comparison of table 6 and table 4 shows that the results are more accurate when an estimate for $A(u, v)$ has been subtracted during fringe analysis, than when windowing is used. The variation of accuracy with spatial frequency exhibits a similar trend, with a maximum value at $0.25 f_x$ and errors increasing more rapidly at high

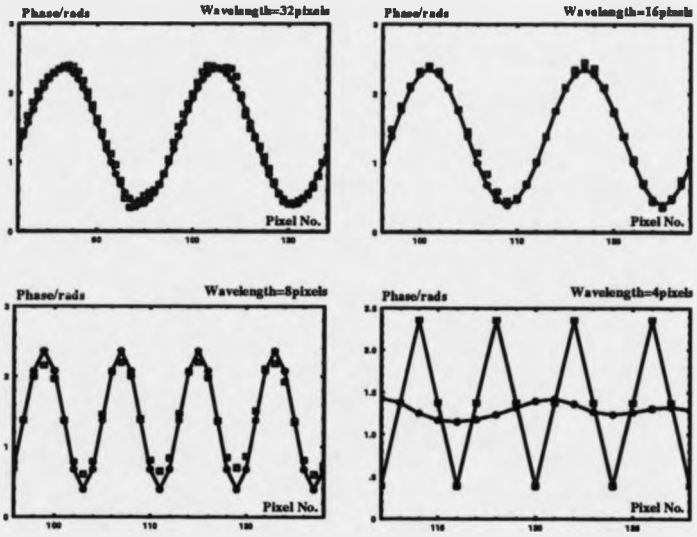


Figure 54: Phase extraction accuracy when the zero order has been removed by windowing

spatial frequencies. Visually, errors at low spatial frequencies appear as ripples in the phase map, whereas those at higher spatial frequencies are seen as areas of reduced phase contrast.

When windowing was used, the rms phase error at a spatial frequency of $0.125 f_x$ was 0.075 radians, compared to 0.050 radians when an estimate of the zero order image had been subtracted. Inspection of the respective line scans (figure 55) shows that the data is rippled in both cases. The noise in the phase map obtained by subtracting $A(u, v)$, was removed by positioning an aperture of radius 64 around the first order in the Fourier domain. The result (third graph of figure 55) is smoother than the original (first graph), but the rms phase error is slightly larger (0.051 radians) because of the reduced modulation of the filtered signal. A large increase in accuracy (0.020 radians) was subsequently obtained by relaxing the radius of the filter to 96 pixels (last graph of figure 55). This procedure exploits the benefits of both windowing and subtraction of an estimate for $A(u, v)$.

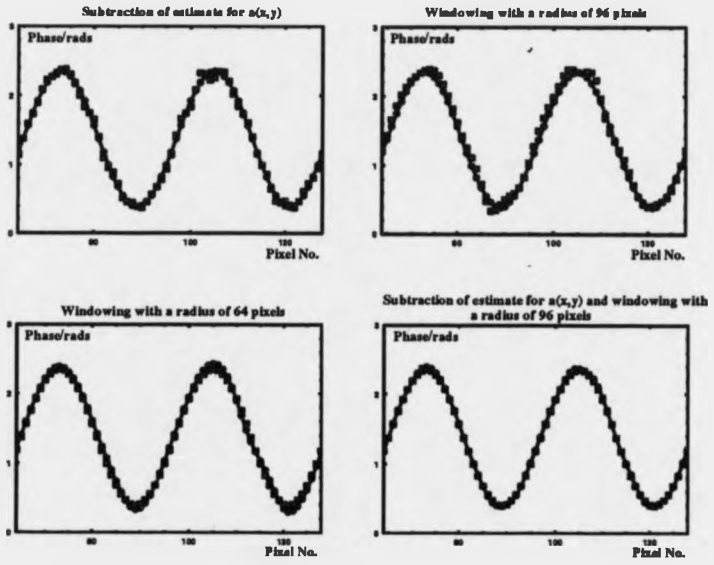


Figure 55: Comparison of strategies for removing the zero spectral order

5.3.4 The Importance of Recording Linearity

Section 5.3.3 has shown that Fourier transform fringe analysis has a maximum accuracy of approximately 0.02 radians (interpolation to 1/300th of a fringe) when the bias image is subtracted, and spatial frequencies exceeding $0.75 f_x$ are removed by centring a low-pass filter at the first order. These results were obtained under 'ideal' conditions, however, with minimal corruption of the bias image. In practice, fringe analysis accuracy is reduced by recording non-linearity and coherent noise. To estimate the effect of non-linearity, the synthetic fringe pattern was modified by multiplying each pixel in the image by a transfer curve.

The non-linear response of a CCD camera was modelled by a composite function consisting of three regions (figure 56). In the first region, at input intensities i_{ip} lower than i_1, the response was governed by a parabolically increasing function, with the minimum situated at the origin to simulate calibration of the framestore to a zero black level. The curvature of this function was set by a parameter k_1.

The middle portion of the piece-wise approximation, in the range $i_2 < i_{ip} < i_1$, was

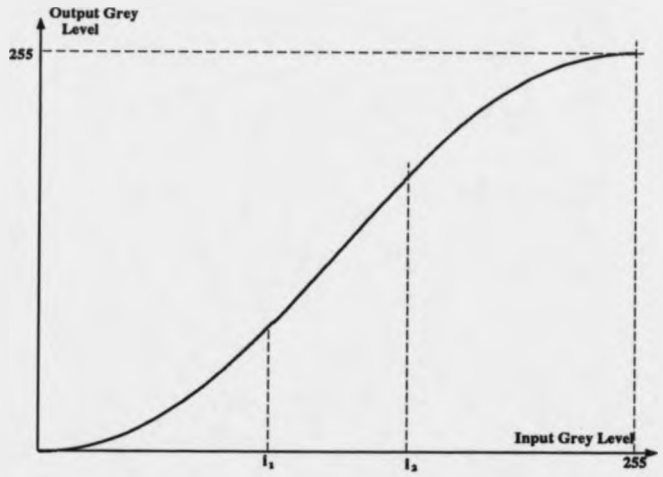


Figure 56: Piece-wise non-linear response curve

modelled by a linear function, with two variables k_2 and k_3 used to define the gradient and y-axis intercept respectively. The upper part of the response curve consisted of a 'nose-up' parabola, with a maximum centred on a grey level of 255 to simulate the asymptotic saturation of the camera. A parameter k_4 was used to control the gradient of this function.

Mathematically, the three response regions are defined by:

$$i_{op} = k_1 i_{ip}^2 \tag{103}$$

$$i_{op} = k_2 i_{ip} + k_3 \tag{104}$$

$$i_{op} = -k_4(i_{ip} - 255)^2 + 255 \tag{105}$$

The response curves contain four unknown constants k_1, k_2, k_3 and k_4, when the boundaries i_1 and i_2 are predefined. If these equations are to represent a realistic response curve, the composite function must be continuous and smoothly varying.

These conditions may be satisfied, and the unknown parameters assigned, if the values of the functions and their first order derivatives are equated at the boundaries i_1 and i_2. From this analysis, the constants are given as:

$$k_1 = \frac{255}{i_1(255 - i_1 + i_2)} \tag{106}$$

$$k_2 = \frac{510}{255 - i_1 + i_2} \tag{107}$$

$$k_3 = \frac{255 i_1}{i_1 - i_2 - 255} \tag{108}$$

$$k_4 = \frac{255}{(255 - i_1 + i_2)(255 - i_2)} \tag{109}$$

In this study, the response curve was taken to be symmetrical about the central input intensity value 127 and therefore the degree of non-linearity was definable by a single parameter i_1, where i_2 is equal to $(255 - i_1)$. Two values of i_1 were used for the study. Initially, i_1 was set to 100, so that the linear response occupied 22% of the dynamic range. The response was subsequently made more non-linear by increasing i_1 to 125 (2% linear region). The rms phase errors were then calculated at a spatial frequency of $0.125 f_x$ for the following cases:

- With the zero order removed by subtracting an estimate for $A(u, v)$.
- With a window of radius 96 pixels used to extract the first order.
- With an estimate for $A(u, v)$ subtracted and the noisy high frequencies removed by a window of radius 96 pixels.

The line scans of the phase map produced with $i_1 = 100$, and the zero order removed by subtraction, are shown in figure 57. The silhouette of the aerofoil in $a(x, y)$ is unaffected by the non-linear recording process, yet the phase data is badly

corrupted (rms error of 0.293). This result has important experimental implications, because it is often difficult to obtain a completely accurate estimate for the bias image. When windowing is used instead, the phase extraction accuracy is improved to 0.140 radians. Hence, it appears that the discrepancy between the estimate for $A(u,v)$ and the true zero order, is confined mainly to the high spatial frequencies. This is expected because the non-linear recording process causes saturation, which reduces the 'smoothness' of an image and results in the generation of spectral harmonics.

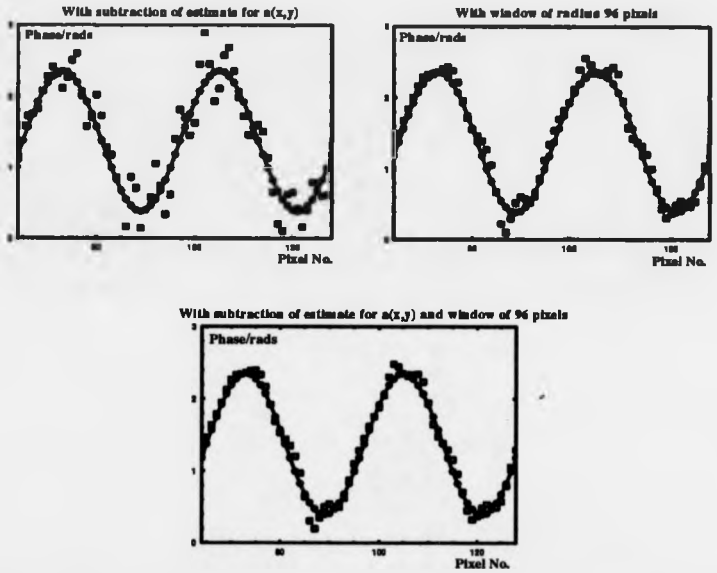


Figure 57: The effect of non-linearity on phase extraction accuracy

The phase data is restored most accurately (0.102 radians) when both an estimate for $A(u,v)$ is subtracted and the noisy high frequencies are removed by windowing with an enlarged aperture of radius 96 pixels. The resulting line scan (last graph of figure 57) shows that the restored phase profile still exhibits slight rippling. Unfortunately, if this noise is removed by using a smaller aperture, then the phase errors at higher spatial frequencies increase due to a loss of phase modulation (third graph of figure 52). The aperture must therefore be optimised to the mid-band of the phase object.

The results obtained when the response curve is less linear ($i_1 = 125$) are shown graphically in figure 58 and summarised in table 7.

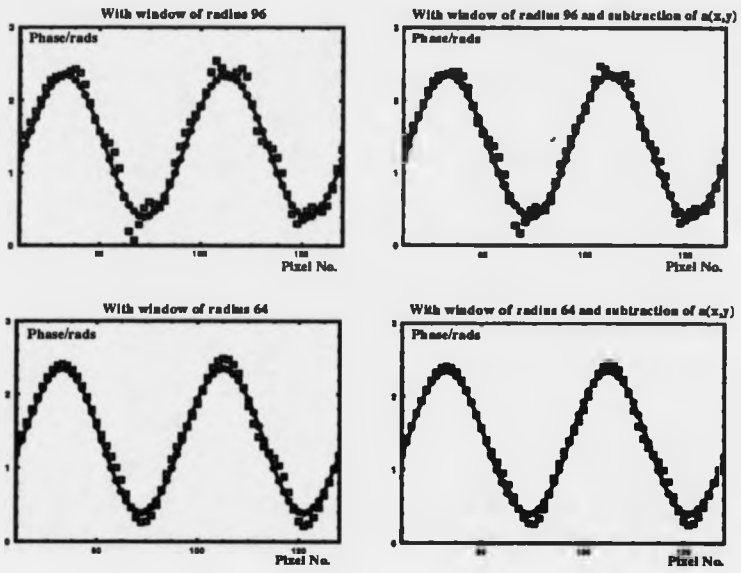


Figure 58: The effect of strong non-linearity on phase extraction accuracy

Aperture Radius /pixels	96	64
e_{rms} /rads ($A(u,v)$ not subtracted)	0.144	0.097
e_{rms} /rads ($A(u,v)$ subtracted)	0.107	0.081

Table 7: Phase extraction accuracy when the response curve is highly non-linear

The phase extraction accuracy obtained with $i_1 = 125$ is very similar to that obtained with $i_1 = 100$, even though the linear region is reduced from 22% to 2% of the full dynamic range. This result may be attributed to the fact that the non-linear process affects mainly the high spatial frequencies, which are removed during Fourier processing. The line scans in figure 58 show that the phase errors are largest at the turning points of the sinusoids. These areas of the phase map coincide with flow regions containing large density gradients. Refraction cannot be neglected in these regions and therefore fringe analysis errors are not the limiting factor in data extraction accuracy.

This study illustrates that phase errors are increased by non-linear recording of the fringe pattern. Fringe analysis accuracy may be improved if the cut-off frequency of the window is reduced, so that noise in the high frequencies is removed. Unfortunately, the penalty paid by increasing the accuracy of the measurements at low spatial frequencies, is that the resolution of the phase map is reduced. In practice, this limitation can be overcome by using an image acquisition system with a larger space-bandwidth-product, so that the sampling rate is increased and a higher carrier frequency can be employed. Alternatively, the interferogram can be magnified prior to processing, with the field-of-view analysed as several separate images.

5.3.5 Linearity of a Panasonic WV-CD50 CCD Camera

A Panasonic WV-CD50 CCD camera was used for capturing all of the interferograms presented in this thesis. In normal operation, an automatic gain control (AGC) is used to improve the dynamic range of the camera, by amplifying and attenuating the signal in extreme lighting conditions. This facility results in a highly non-linear response and therefore it was turned off during capture of the interferograms. The linearity of the camera was subsequently investigated by shining a collimated He-Ne laser beam onto the centre of the CCD array. The intensity of the input beam was adjusted for each successive reading by neutral density filters and measured using an independent photodetector, calibrated for use at 633nm. The output grey level was determined for each input intensity, by calculating the peak of the histogram obtained from a frame-grabbed 8-bit raster image of the gaussian laser beam. A graph of normalised intensity versus peak output grey level (figure 59) shows that the response is linear within experimental error. The variance of the normalised intensity values from the linear regression is only 3×10^{-4}.

The results from this study of camera non-linearity, show that a good estimate for the bias image $a(x, y)$ can be obtained. This image can be Fourier transformed and subtracted from the spectrum of the interferogram to remove the zero order. The accuracy of the phase data presented in this thesis should therefore exceed 0.1

radians, provided the positive and negative bandlimits of the phase object are carefully tailored to the carrier frequency and pixel spacing. In practice, the bandwidth of the encoded flow information may not be known accurately. Hence, the spacing of the finite fringes should not be allowed to deviate by more than 50% of their ambient value, to ensure that spatial frequencies are confined to the range $[-f_x/2, f_x/2]$. This condition prevents aliasing, and reduces the influence of recording non-linearity and coherent noise that mainly affect the high spatial frequencies.

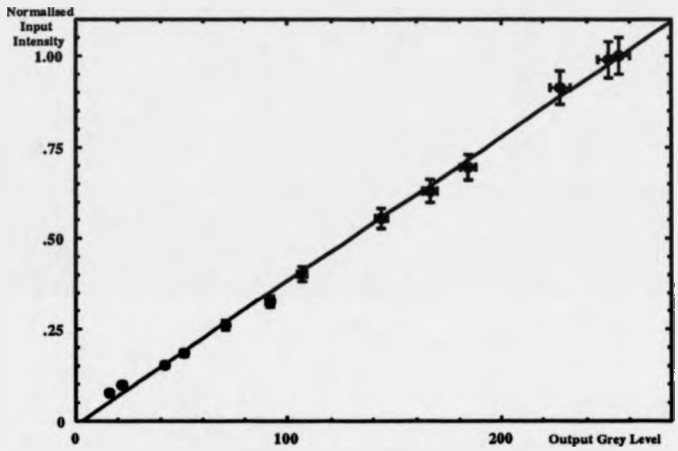


Figure 59: Linearity of the CCD camera used to capture the fringe patterns

5.3.6 Conclusions on Fourier Transform Fringe Analysis

When interferometric data is recorded and processed, the positive and negative bandlimits of the phase object must be matched to the carrier frequency of the finite fringes and the space-bandwidth-product of the imaging sensor. Aliasing may be prevented by ensuring that the fringe spacing does not deviate by more than 50% of the width of the unperturbed carrier fringes, so that the bandlimits of the phase object are constrained to $[-0.5f_x, 0.5f_x]$. The spatial frequency variations should not extend beyond this range, because fringe contrast is reduced as f_x is approached due to the finite sampling area of the pixels. Also, it is difficult to separate the zeroth and first

spectral orders if the phase object contains frequencies exceeding $0.5f_x$. The finite fringes should be sampled at twice the Nyquist frequency (four pixels/carrier fringe) to ensure that spatial frequencies as high as f_x may be encoded before the imaging sensor generates aliased moiré fringes.

Phase extraction accuracy is a function of spatial frequency, with an optimum value at approximately one quarter of the carrier frequency. When the bias image $a(x,y)$ is known accurately, and the image acquisition system is linear, the phase data may be extracted to an accuracy of approximately 0.02 radians (1/300th of a fringe). Lower frequencies are corrupted by the amplitude noise term $B(u,v)$ and high spatial frequencies are lost during removal of the zero order bias term $A(u,v)$. Phase extraction accuracy is highest when an estimate for the bias term has been subtracted and noisy high frequencies (typically $> 0.75f_x$) have been removed by windowing. The optimum filter radius is a function of the spatial frequency content of the phase object, and accuracy at low and high frequencies must be compromised to minimise the overall rms phase error. This may be achieved approximately, by optimising the filter radius at the mid-band frequency of the phase object. Visually, the phase map appears rippled if the aperture is too large, whereas phase modulation is reduced if the aperture is too small.

If the interferogram is recorded non-linearly then a smaller filter radius is required to suppress rippling in the phase map and phase extraction accuracy is reduced to approximately 0.1 radians (1/60th of a fringe). The use of a smaller cut-off frequency inevitably reduces the resolution of the data, although the accuracy at lower spatial frequencies is largely unaffected by the degree of non-linearity. Results from this study show that a highly linear response can be obtained if the camera's AGC is disabled. Although this procedure can be ineffective if the fringe pattern is noisy and has variable contrast, it provides phase maps of the highest resolution if the interferograms are correctly recorded.

5.4 Phase Unwrapping

Interferometric data is encoded in a 'wrapped' format because of the periodicity of the interfering object waves. Therefore, when one of the two object waves generated by a double-exposure hologram is retarded by an integer number of wavelengths, the resulting interferogram remains unchanged unless the coherence length of the radiation is exceeded. Interference fringes are sinusoidal in profile and therefore it is impossible to determine the sign of the phase change between adjacent fringes. Consequently, measurement ambiguities arise when the phase change exceeds 2π unless *a priori* knowledge is available.

Phase sensitive fringe analysis strategies resolve this problem, by detecting which direction the fringes move when the phase of one of the object waves is incremented in either the spatial or temporal domains. The result is a set of 'tan fringes', from which the 'unwrapped' phase map may be calculated. Unfortunately, only principal phase values are determined from the inverse tangent operation, because of the periodicity of trigonometric functions. If the phase is calculated by the FORTRAN ATAN2 function, or the signs of the numerator and denominator are used to determine the quadrant of the phase argument, the values are wrapped modulo 2π (compared to modulo π if sign information is not used).

Unlike a sinusoidal fringe pattern, the ambiguities in a wrapped phase map may be resolved by observing the sign of the phase gradient prior to each 2π discontinuity. A positive gradient reveals that the phase is increasing and that all values after the jump must be incremented by 2π. Conversely, a negative gradient indicates that the values must be decremented by 2π. Most interferograms are comprised of numerous fringes and therefore it is necessary to keep a running total of the increments with respect to a reference point in the field.

Experimental fringe patterns are degraded by noise and therefore the actual discontinuous phase jumps are less than the theoretical 2π. The discontinuities may be detected by assigning a threshold value (π is usually used), above which the phase

change is assumed to have arisen due to the wrapping process. If large phase gradients are present, then these may be misinterpreted as 'wrapping points' rather than continuous phase changes. These quantised errors are subsequently streaked along the rest of the unwrapping path and spatial inconsistencies are generated in the phase map. Errors may also occur at discontinuous phase jumps and in areas of high frequency large amplitude noise. These image features are common in flow field interferograms, where discontinuous shock waves occur and diffraction noise is generated around the silhouette of the wind tunnel model.

Phase errors may be reduced by preprocessing the image with a digital low-pass filter prior to phase unwrapping. This operation removes high spatial frequency noise, while leaving the low frequency undulations of the fringes unchanged. The median filter, which substitutes the value at the centre of the convolution mask by the median of the values constrained within the mask area, is often a good choice because it removes point noise without severely blurring the image. Alternatively, the incorrect assignment of phase increments can be minimised by devising a noise immune phase unwrapping algorithm. Error propagation can be prevented if the phase is unwrapped along a path which bypasses regions of high fringe density and noise. Global minimisation of errors during automated phase unwrapping, requires the use of an optimised unwrapping path.

The optimum unwrapping path may be calculated by using a knowledge based system to identify occluded regions and those with a low signal-to-noise ratio. These areas may then be removed from subsequent phase unwrapping processes by placing a binary mask over the image. The unwrapping algorithm should be capable of analysing an arbitrarily complicated structure in a convenient run-time and should interpolate correctly unwrapped data into noisy areas. This approach has been adopted by Brown for processing speckle pattern fringes of vibrating vehicle parts [204]. In his algorithm, areas in the interferogram are rejected if they exhibit the following features:

- Sheared fringes (adjacent phase discontinuities).

- Violation of the Nyquist sampling rate.
- Abrupt changes in the image boundary.

After the mask has been created, the individual regions are unwrapped along rectilinear paths from 'seed points'. Unwrapped areas which are completely enclosed by the mask, are iteratively corrected with respect to adjacent regions by adding multiples of 2π to ensure that phase changes across boundaries are consistent. This approach is computationally inefficient and manual intervention is often required to prevent error propagation if the phase unwrapping procedure fails. Nevertheless, Brown has routinely applied the algorithm in a factory environment and the software is constantly being refined to decrease the level of manual intervention.

Recent research into the phase unwrapping problem has concentrated on the development of simple mathematical processes that can operate on individual cells to produce a collectively complex behaviour. These mathematical systems are known as cellular automata and their history is reviewed by Cooper [205]. Cellular automata were first used for phase unwrapping by Ghiglia *et al* in 1987 [175]. Their results show that by performing local neighbour tests, cellular automata may produce consistent phase data without the need for complicated rules. In general, these techniques work by dividing the field into 2×2 groups of pixels (cells), that act as termination boundaries for erroneous unwrapping paths. The 'quality' of a cell may be assessed for inclusion in the unwrapping route, by checking that the net phase change is zero around its perimeter. An advantage of cellular automata over knowledge based strategies is that, because the algorithm is inherently simple and operates on individual cells, it can be implemented on a parallel computer architecture.

A limitation of cellular automata is that they are computationally inefficient, because many global iterations are required to obtain convergence. This problem has been recognised by Huntley, who has developed an algorithm that uses cellular automata in a single iteration to place 'cuts' across inconsistent paths prior to unwrapping [176]. A second limitation is that cellular automata are unable to recognise

large-scale image inconsistencies. Optimum data extraction from an interferogram requires the identification of errors on a hierarchy of scales. For example, phase inconsistencies may arise due to holes in the image created by regions of low object reflectivity, or because of aliasing at the imaging sensor. These problems have been addressed by Bone, who has improved Huntley's algorithm by taking account of phase curvature in the unwrapped phase map [206]. This information may be used to identify regions which have been unwrapped consistently, but are still erroneous on a larger scale. In some instances, however, second and higher order spatial derivatives must be used and these are prone to speckle noise. A review of the current status of phase unwrapping is given by Huntley *et al* [207]

The images in this thesis were phase unwrapped by a procedure that exploits the benefits of both cellular automata and a knowledge based system. The algorithm, originally developed by Judge [203], employs an unwrapping strategy that operates on a hierarchy of scales. On the lowest level, error propagation is prevented by choosing an unwrapping path whereby the sum of phase changes between pixels is minimised. This process provides a tolerance against speckle noise, which generates rapid fluctuations in the phase map. The total phase difference between connected pixels is minimised by calculating the minimum spanning tree configuration from an analysis based on graph theory [208].

Large scale phase inconsistencies, such as those generated by aliasing, image discontinuities and noise, are detected by dividing the wrapped phase map into an array of tiles. Unlike cellular automata, the size of these tiles may be customised so that regional errors lie within a single tile. Erroneous tiles may then be rejected from subsequent analysis and the missing data may be estimated by interpolation. The quality of a tile is assessed for inclusion in the unwrapping path by performing a series of tests:

- Detection of points of low modulation; i.e. when the numerator and denominator of equation 85 are small.
- Detection of discontinuities.

- Detection of aliasing errors (regions of high or low fringe density).

These criteria are used to assign a weight to the edge of each tile. The optimum unwrapping path is then determined by calculating the minimum spanning tree of these weights. Spatial inconsistencies in the phase map are minimised with this approach, because noisy tiles terminate the branches of the tree.

6 Analysis of the Two-Dimensional Flow Around a NACA 0012 Aerofoil

6.1 Principal Aims of the Test

The principal objective of this first wind tunnel test is to establish which method of holographic interferometry produces fringe patterns that are conducive to quantitative analysis. A two-dimensional model is tested, so that the path integrated measurements can be converted into density information without recourse to tomographic reconstruction techniques. The experiments are conducted in the continuous running transonic wind tunnel at Cranfield Institute of Technology. Diffuse and specular illumination geometries are compared, and an investigation of dual reference beam systems is made. In addition, the spatial resolution of the interferograms is assessed and techniques for eliminating fringe patterns generated by rigid body motion are explored.

The operation of a holographic interferometer in an industrial environment, poses numerous practical problems:

- Mechanical and optical design of the interferometer.
- Configuring the optical system to function within the available space.
- Stabilisation of the optical system, so that the object and reference beams remain phase-locked to within a single optical wavelength (0.7μm).
- Operation of a fragile pulsed laser in a dusty environment, while exposed to high levels of vibration.
- Isolation of optical components from contamination.
- Recording and developing of low noise holograms in adverse conditions.
- Building of a holographic reconstruction system within the test environment, to verify the performance of the interferometer.

- Safety management of a high power pulsed laser.
- Logistics of performing a trial away from base.

These practical skills must be acquired before more complicated optical configurations are implemented. To enable the extraction of quantitative information from the reconstructed interferograms, both narrow and wide angle dual reference beam systems were used. These configurations are designed to encode the two exposures uniquely and enable separate phase manipulation of the reconstructed wavefronts (section 3.4.4). When the angle between the reference beams is small ($\sim 1°$), the holograms are conducive to Fourier transform fringe analysis in a single beam reconstruction system (section 4.2.4). In contrast, a wide angle ($\sim 30°$) dual reference beam system is suitable for processing the interferograms by phase stepped fringe analysis (section 4.2.3). An image plane configuration was used to reduce diffraction noise and relax the coherence requirements of the reconstruction source.

6.2 Wind Tunnel and Model Description

The 9"×7.5" transonic wind tunnel at Cranfield Institute of Technology was used to investigate the two-dimensional flow around a standard aerofoil by holographic interferometry. The tunnel specification was as follows:

General

Drive motors	-	2×550h.p. (D.C)
Compressor	-	Single stage centrifugal
Running time	-	Two hours maximum
Stagnation pressure	-	2-15psi
Stagnation temperature	-	20-30°C
Mach number range	-	0.30-1.05
Model mounting	-	Sting attached to slotted liner
Schlieren quality windows	-	Rectangular: 875mm×256mm×25mm

Transonic Working Section Details

Width	-	9 inches
Height	-	7.5 inches
Length	-	48 inches
Number of slots	-	7 on floor and ceiling, none on sidewalls
Width of slots	-	0.225 inches
Open area ratio	-	10% (4 slots open)

The wind tunnel consisted of a closed circuit, with the air circulated continuously by a single stage compressor driven by two 550hp direct current motors. An electrical input power of typically 0.5MW was required to operate the facility, which was supplied by a dedicated substation. The electrical supply was subject to large fluctuations, because the power consumption of the campus rose significantly during tunnel operation. Consequently, to prevent 'surges' in the flow, the peak stagnation pressure was restricted to approximately 10psi. This stagnation pressure provides a maximum Reynolds number of typically 0.90×10^6 for models accommodated by the working section. Prior to commencing a run, the tunnel was evacuated to approximately 2psi so that the energy required to accelerate the flow to transonic speeds was reduced. Once the rarefied air had reached a transonic Mach number, the stagnation pressure was

gradually increased to the operating value over a duration of approximately twenty minutes.

Originally the wind tunnel was intended for supersonic use and the working section measured 229mm×229mm×1.07m. The tunnel was subsequently converted for transonic operation, by the insertion of a slotted liner on the floor and ceiling of the test section (figure 60). The purpose of the slots is to create a pressure drop through the walls. This increases the effective height of the working section and reduces the possibility of shocks reflecting back onto the model. The desired pressure drop is obtained by choosing the correct wall thickness and plenum chamber pressure.

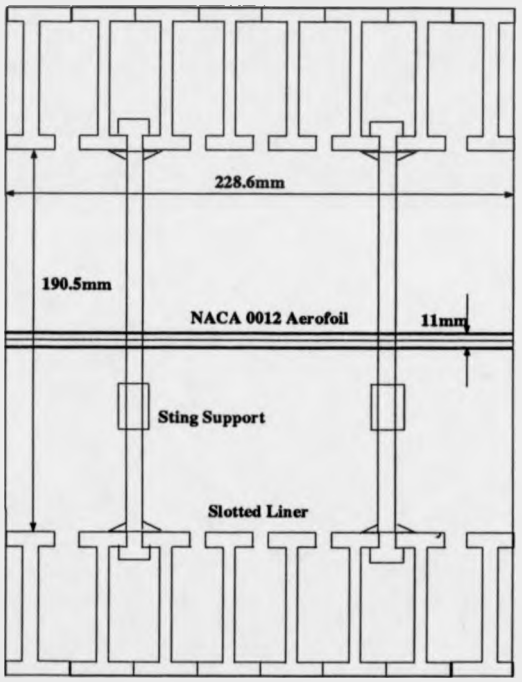


Figure 60: Scale diagram of the slotted liner and model used in the Cranfield transonic wind tunnel

The new liner was deeper than its predecessor, reducing the working section height to 191mm. The transonic working section was a 1:10.4 scale model of the European Transonic Wind Tunnel (ETW). Side wall slots were not used, as the tunnel was intended as a two-dimensional test facility. Further details of the design, construction

and operation of the tunnel are given by Gibb [209]. Gibb also shows that the empty tunnel noise levels are low, making the facility suitable for making unsteady pressure measurements.

A mechanically operated downstream throat was used to control the flow, so that holograms could be recorded with a range of Mach numbers. The throat was originally designed to provide Mach numbers in the range [0.30,1.05]. In practice, however, the range was reduced to [0.44,0.93] because of the wind tunnel model. Unfortunately, Mach 0.44 is too high to assume incompressible flow and therefore any quantitative analysis must consider the effect of using a perturbed density field for the quiescent state.

A two-dimensional NACA 0012 aerofoil (figure 61) was chosen as the aerodynamic body for the test. This model is one of the 'four-digit' series of aerofoils used by the National Advisory Committee for Aeronautics (NACA) in 1929, as part of a systematic study of aerofoil characteristics. The aerofoil is currently used for helicopter rotor blades and is an AGARD standard for testing numerical solutions [210]. Consequently, a wealth of both experimental and theoretical flow data exist for this model [211]. In addition, it has been the subject of an extensive holographic flow visualisation research programme at NASA Ames [212].

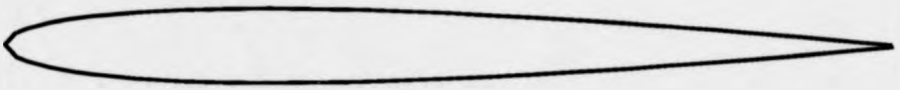


Figure 61: Section of the NACA 0012 aerofoil

To simulate free flight conditions in a wind tunnel, both the freestream Mach number M_∞ and Reynolds number Re of the flow should be matched. From equation 110 it can be seen that the Reynolds number of the flow is directly proportional to the linear scale of the model. Hence, because the model is of reduced scale, it is difficult to achieve a sufficiently large Reynolds number in wind tunnel tests.

$$Re = \frac{\rho u l}{\mu} \tag{110}$$

Where,

ρ - air density

u - flow speed

l_c - characteristic length of the object (e.g. chord)

μ - dynamic viscosity

One solution is to change the density or viscosity of the fluid by using gases other than air, or by operating the tunnel at cryogenic temperatures. In the Cranfield transonic tunnel this was not possible, and the only option available was to maximise the size of the model within the constraints of the working section. The use of a large model is also desirable optically, because some of the flow features are magnified and therefore lower resolution images may be tolerated.

Tunnel blockage imposes an upper limit on model size, and previous tests have shown [213] that steady tunnel operation is only possible when the combined frontal area of the model and sting, is less than 10% of the total cross-sectional area of the working section. The aerofoil was mounted in the tunnel using a pitch fork sting, which enabled adjustment of incidence and vertical height (figure 62). The largest model that could be used had a chord of 89mm (3.5"), when allowance was made for the sting. The shape of the support was designed to avoid tunnel interference and minimise additional blockage. In addition to affecting tunnel power requirements, 'solid blocking' results in modification of the flow velocity over the aerofoil. An empirical formula (equation 111) was used to calculate the effective freestream flow velocity u_e from the measured value u_m.

$$u_e = u_m \left[1 + \frac{\lambda_a}{\sqrt{12}} \left(\frac{\pi t_a}{h_w}\right)^2\right] \tag{111}$$

λ_a - geometrical factor dependent on the aerofoil section

t_a - aerofoil thickness

h_w - height of the working section

The geometrical factor is 4.5 ± 0.5 for a NACA 0012 aerofoil and therefore the effective flow velocity at the aerofoil is 1.2% higher than the freestream value calculated

from measurements of static and stagnation (total) pressure. The effective velocity is reduced by the use of a slotted liner because it increases the apparent height of the working section. However, this calculation shows that the Mach number measurements are accurate to approximately 1%.

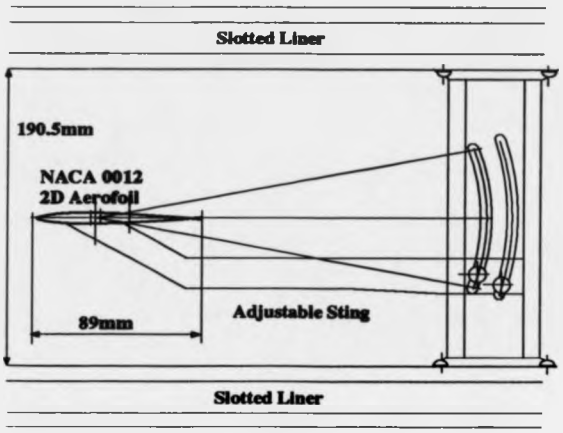


Figure 62: Scale diagram of the adjustable pitch fork sting and aerofoil model

A preliminary schlieren test, using a conventional mercury source, was made to ensure that the slender pitch fork sting and streamlined upright supports, were not generating tunnel interference. Rubber mounts were attached to the ends of the model where it contacted the test section windows, to prevent the glass from cracking when the windows bowed inwards during tunnel evacuation. Model incidence was set to zero by slip gauges, which enabled a vertical alignment accuracy of 100μm at the nose of the model, and corresponding angular precision of approximately 1mrad. This qualitative schlieren test showed that shock oscillations occurred at Mach numbers in the range [0.82,0.85] and at values exceeding approximately 0.93. However, there was no apparent interference between the model and sting.

Mohan has shown in a previous study [213] that the critical Mach number for the aerofoil is 0.725. His results also demonstrate that at Mach 0.744 the transition to supersonic flow occurs at a quarter chord from the leading edge, and that this region moves rearward with increasing Mach number. When Mach 0.80 is reached, wavelets are observed on the upper surface of the aerofoil, which coalesce as the flow velocity is increased to form a compression wave rather than an abrupt shock. A periodic flow regime is entered above Mach 0.820 and at Mach 0.835 the shock oscillation is strongest, with an amplitude of 15% chord and a frequency of 1.6kHz. Steady flow returns at Mach 0.844 and a weak shock is observed with a lambda foot. At Mach 0.93 periodic flow returns with amplitudes of approximately 30% chord, and lambda shocks form on both upper and lower surfaces. As freestream sonic speeds are approached, these shocks coalesce at the trailing edge and periodic flow ceases.

From the schlieren observations and Mohan's results, it was decided to produce holograms at Mach 0.80 and 0.90. At Mach 0.80 the critical Mach number is exceeded, and transonic flow is observed without the presence of strong shocks and periodic flow. This regime is suitable for holographic interferometry because fringe densities are optimal. Runs were also made in the steady flow window at Mach 0.90 to determine whether the fringe patterns are suitable for analysis at higher Mach numbers.

6.3 System Configurations for Specular and Diffuse Illumination

The double-exposure holograms were recorded using an actively Q-switched 3J ruby laser, with a pulse duration of 30ns and a separation of 6s. A pulsed laser was used so that flow unsteadiness and wind tunnel vibration were 'frozen' during the exposure (section 3.2.2). Two separate holographic configurations were constructed and tested, to compare the techniques of diffuse and specular illumination (section 3.4),

A schematic diagram of the specular illumination holographic interferometer is shown in figure 63. The optical arrangement is essentially a schlieren system with the knife edge removed and a reference beam added. The divergent object wave emerged from a negative lens mounted on the front of the laser and was collimated by an aluminised spherical schlieren mirror (305mm diameter, 3048mm focal length). These planar wavefronts were subsequently transmitted through the flow field normal to the test section windows. The beam was fractionally misaligned to prevent back reflections from re-entering the laser cavity. A second schlieren mirror was then used to reflect and converge the phase encoded object wave onto an achromatic doublet (50mm diameter, 500mm focal length). The centre of the working section was imaged by this lens onto a 102mm×127mm holographic plate (emulsion Agfa 8E75HD).

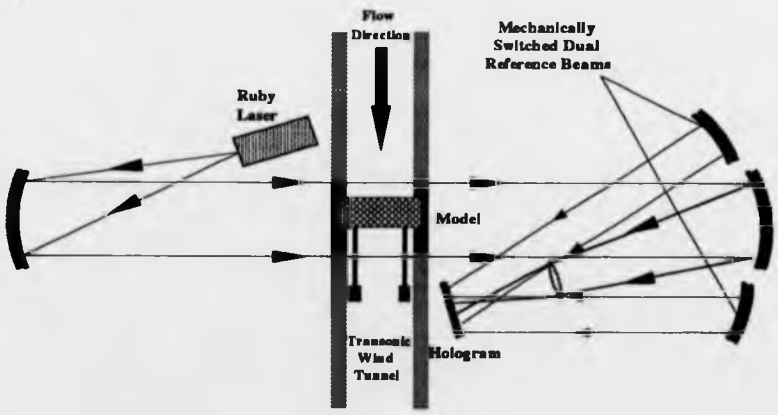


Figure 63: Dual reference wave holographic interferometer used at Cranfield

A diagram of an equivalent imaging system, that uses only refractive elements, is shown in figure 64. The magnification of the imaging system is given as 0.15 by thin lens theory (equation 112), which was found to be within 5% of the value measured from the scale of the holographic image. An argon ion laser was used to reconstruct the holograms because of its high power. The 26% mismatch in recording and replay wavelengths is unimportant when reconstructing specular illumination image plane holograms, because it only results in a scaling of the image.

$$f = \frac{uv}{u+v} \tag{112}$$

Where the magnification of the system m is given by v/u and:

- f - focal length of the optic
- u - object distance
- v - image distance

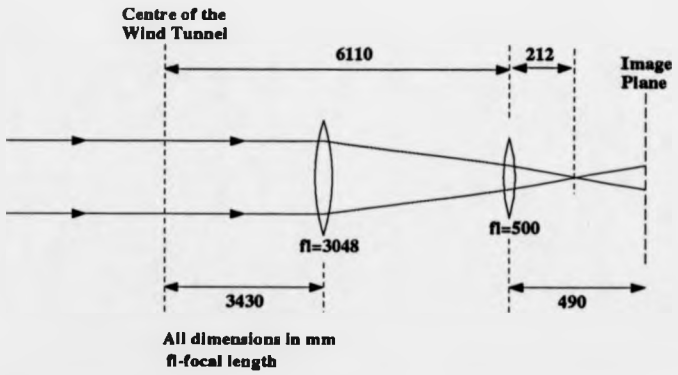


Figure 64: Model of the imaging system using only refractive elements

A replica of the original object wave, and its phase conjugate with wavefronts of opposite curvature, was diffracted from the hologram. With a reconstruction wavelength of 514nm, opposed to the recording wavelength of 694nm, the conjugate wavefront focussed 330mm from the hologram. From the Bragg law (equation 40) and thin lens equation, the expected focal position for this wavelength shift is 389mm from the recording plane. The 15% discrepancy in the calculated and measured positions is due

to shrinkage of the exposed silver halide during fixing. This could have been prevented by using a developer that tans the emulsion; however the additional complexity of this process was considered unnecessary.

The undiverged reference wave (not shown on the diagram for clarity) was generated within the laser head and reflected from a dielectric mirror mounted adjacent to the first schlieren mirror. This arrangement prevents relative motion between the two beams. Further phase locking was provided by keeping the two beams parallel for most of their length. A cubic beam splitter was used to generate two wide angle reference beams, which were expanded by negative lenses onto symmetrically disposed spherical schlieren mirrors (3048mm focal length). Collimated reference beams were produced by locating the source points at the foci of these mirrors. Parabolic mirrors would have provided a greater degree of collimation and subsequently fewer imaging aberrations; however these mirrors are prohibitively expensive and only function correctly at a single angle of incidence and reflection. Spherical reference beams could have been used, but this arrangement complicates alignment of the reconstruction system and inevitably results in wavefront errors. Also, it is harder to control the effects of emulsion shrinkage when the recorded fringe planes are parabolic in shape. The large separation of these collimating mirrors enabled manual switching of the two beams between consecutive exposures.

The normal of the holographic plate was arranged to bisect the two collimated reference beams. This geometry minimises the effects of out-of-plane emulsion shrinkage, because the interference planes forming the hologram lie approximately normal to the emulsion surface. Consequently, the orientation of the fringe planes remains largely unchanged during photographic processing and only the thickness of the grating is affected. The minimum angle between the reference beams was fixed by the proximity of the imaging lens to the holographic recording plane. An angle of 28° enabled separation of the reconstructed object beams, while maximising the spacing of the interference planes recorded on the hologram.

At a wavelength of 694nm the spatial frequency of the fringes given by equation 20

is approximately 700 lines/mm, whereas the recording capabilities of Agfa 8E75HD holographic emulsion extend to 4000lines/mm. A faster film (e.g. Agfa 10E75), with a lower resolution, could therefore have been used; however sufficient laser power was available and it was thought that the finer grained material would guarantee the recording of all spatial frequency information. The holographic plates were index matched to black float glass during the exposure to prevent multiple reflections. Practical details of the recording and photographic developing process are given in appendix 2. Fringe contrast of the holograms was maximised by using linearly polarised light, with the vectors of both beams aligned vertically, and by ensuring that the mean path lengths of the object and reference beams were equal.

The narrow angle dual reference beam holograms were recorded using the same optical system, but with one of the wide angle reference beams disabled. Finite fringes were generated by deflecting the beam incident to the collimating mirror, through an azimuthal angle of 1° between exposures. The angle was calibrated by mounting the steering mirror on a motorised rotation stage, with stops used to define the initial and final angles.

The diffuse illumination holograms were recorded using a different optical system. A schematic diagram of the interferometer is illustrated in figure 65 and a photograph of the apparatus is shown in figure 66. Unlike the specular illumination geometry, the diffuse system used an optically rough Lambertian diffuser to back-light the flow field which gave a azimuthal viewing angle of 75°. The diffuser consisted of a 1.8m×1.2m screen, manufactured from an aluminium honey-comb panel, with laminated aluminium face sheets. The screen was designed to be light and stiff, to prevent low frequency excitation by the wind tunnel. Also it simplified the construction of the supporting frame, which was required to suspend the top of the screen approximately 2.5m above floor level. The optical path difference between the object and reference beams, was stabilised to within 1μm for the 6s interval between exposures to prevent the appearance of vibration induced error fringes. This was accomplished by mounting the optics of both beams on common nodal points.

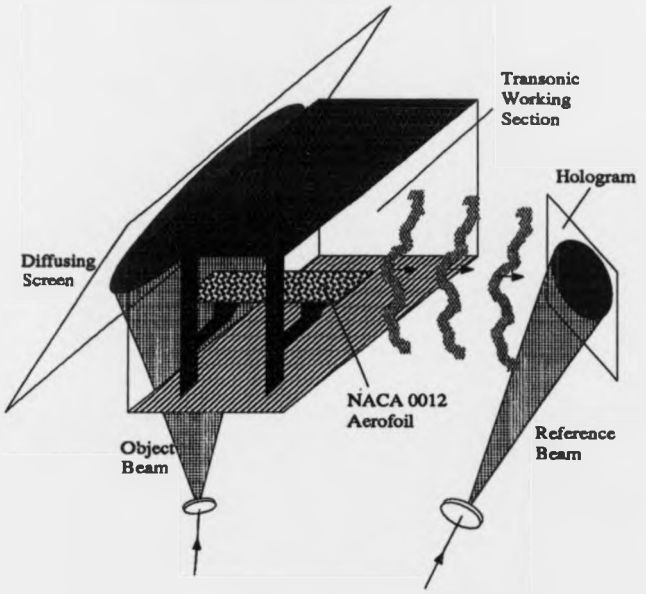


Figure 65: Diffuse illumination holographic interferometer used at Cranfield

Figure 66: Photograph of the diffuse illumination holographic interferometer

6.4 Experimental Method

A thorough appraisal was made of the resolution and stability of the optical systems prior to the wind tunnel tests. Image resolution is governed by the aberrations of the optical system and the signal-to-noise ratio of the holograms. The specular illumination interferometer was optimised by adopting a conventional 'Z' schlieren configuration, with the angles between the beams minimised. In this arrangement coma and astigmatism are reduced, and the beam paths are folded compactly. The diffuse illumination system did not include imaging optics, and therefore speckle and diffraction noise determined the maximum resolution that could be obtained from the reconstructed interferograms.

The signal-to-noise ratio of the recorded holograms was maximised by diligently controlling the following experimental conditions:

- Suppression of scattered light.
- Prevention of parasitic interference.
- Ratio of the object and reference beam intensities.
- Exposure density.
- Direction and type of polarisation.
- Cleanliness of the optical system and processing environment.

With a diffuse illumination holographic interferometer, the object beam is scattered randomly from the diffuser and a continuum of phase correlated point sources are produced. The wavefronts emitted from these points mutually interfere and form numerous component gratings during recording of the hologram. When the hologram is replayed these generate intermodulation noise, which reduces the contrast and resolution of the reconstructed image. Intermodulation noise may be suppressed by using a reference beam which is more intense than the object wave, so that the interference

6.4 Experimental Method

A thorough appraisal was made of the resolution and stability of the optical systems prior to the wind tunnel tests. Image resolution is governed by the aberrations of the optical system and the signal-to-noise ratio of the holograms. The specular illumination interferometer was optimised by adopting a conventional 'Z' schlieren configuration, with the angles between the beams minimised. In this arrangement coma and astigmatism are reduced, and the beam paths are folded compactly. The diffuse illumination system did not include imaging optics, and therefore speckle and diffraction noise determined the maximum resolution that could be obtained from the reconstructed interferograms.

The signal-to-noise ratio of the recorded holograms was maximised by diligently controlling the following experimental conditions:

- Suppression of scattered light.
- Prevention of parasitic interference.
- Ratio of the object and reference beam intensities.
- Exposure density.
- Direction and type of polarisation.
- Cleanliness of the optical system and processing environment.

With a diffuse illumination holographic interferometer, the object beam is scattered randomly from the diffuser and a continuum of phase correlated point sources are produced. The wavefronts emitted from these points mutually interfere and form numerous component gratings during recording of the hologram. When the hologram is replayed these generate intermodulation noise, which reduces the contrast and resolution of the reconstructed image. Intermodulation noise may be suppressed by using a reference beam which is more intense than the object wave, so that the interference

gratings formed between the object and reference beams dominate the reconstruction. However, as the intensity imbalance between the object and reference beams increases, the diffraction efficiency of the hologram decreases and Rayleigh scatter reduces the signal-to-noise ratio of the reconstructed image. The optimum intensity ratio depends on the mode of illumination and, for this particular diffuse system, the best results were obtained when the reference beam was four times as intense as the scattered object wave (measured in the plane of the hologram).

In contrast, the wavefronts in the specular illumination system propagate in predominantly the same direction from a common source point and therefore intermodulation noise is lower. A series of holograms were recorded of the working section using a range of beam ratios and the results showed that the signal-to-noise ratio was highest when the reference beam was 1.5 times more intense than the object beam. The holograms recorded in specular illumination gave a higher diffraction efficiency than those produced with diffuse illumination, because the object and reference beams were more closely balanced in intensity.

The resolution of the reconstructed holographic images from the specular illumination system was assessed by the use of an ASAF 1951 resolution slide (figure 67). The lowest spatial frequency represented on this chart is $1lp$/mm and from (figure 67) it can be seen that the limiting resolution is approximately 100μm. The resolution of the system may be higher than this value, but it was difficult to focus the camera when the image was photographed, due to speckle noise produced by the ground glass screen in the viewfinder. Image resolution is also reduced by interference of a weak secondary beam with the main image. Secondary reflections could have been prevented if the wind tunnel windows had been AR coated. Attempts at removing the unwanted beam by spatial filtering were unsuccessful because there was insufficient angular separation between the two images.

The resolution of diffuse illumination interferograms is limited by the depth-of-field of the localised fringes and speckle noise. The depth-of-field of a diffraction limited system, with a resolution of d_r, is given approximately by:

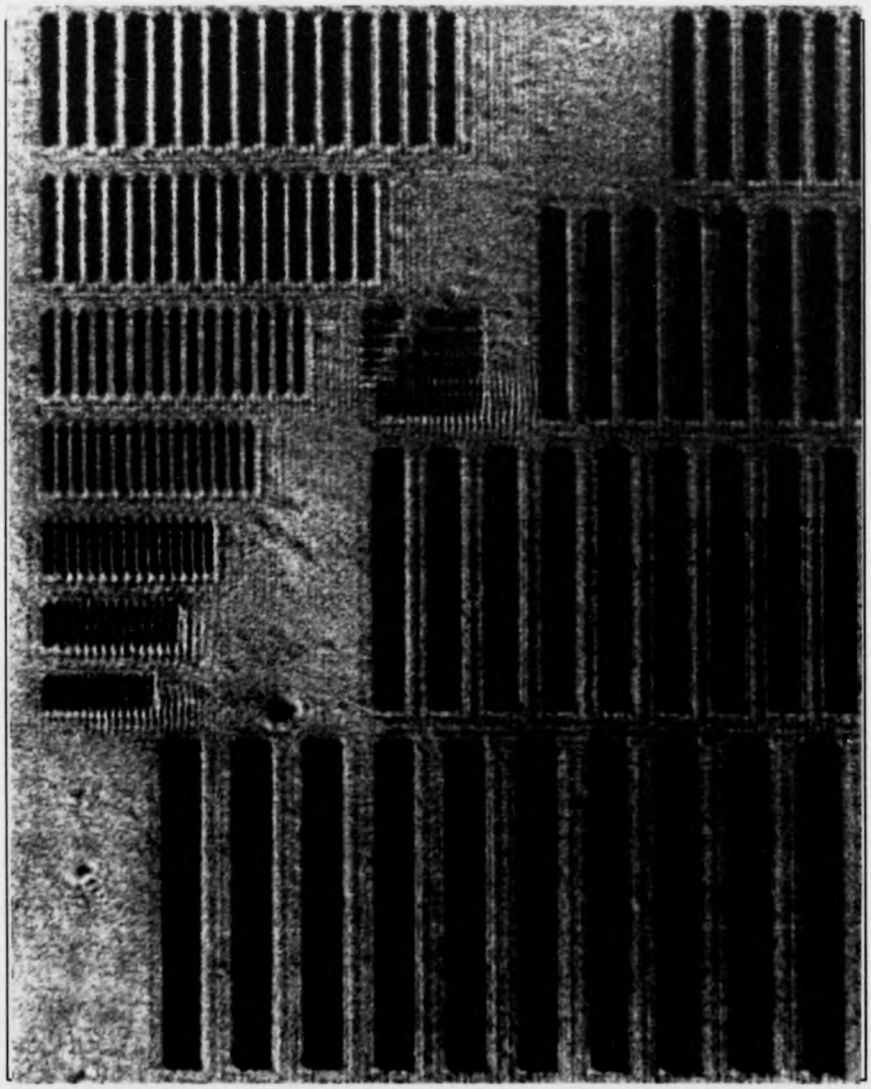


Figure 67: Holographic image of an ASAF resolution chart (coarse grid pitch of 1mm)

$$dof = \frac{d_r^2}{\lambda} \tag{113}$$

Hence, if a lens is used which ensures that all of the features across the working section are in focus (depth of field 229mm), then the maximum resolution that may be obtained is approximately 400μm. In practice, the resolution is reduced further by aberrations of the hologram, diffraction noise and speckle.

The stability of the specular and diffuse illumination holographic interferometers was assessed by recording a series of double-exposure holograms, with the delay between the exposures varied. Initially, exposures were made without the wind tunnel running, so that vibration was reduced. The results showed that error fringes could be eliminated from the reconstructed interferograms, when the delay was shorter than two minutes. To prevent air currents from generating spurious fringes, all heaters were turned off, and the laser was fired remotely.

Double-exposure holograms were then recorded with the tunnel running, but with no density changes occurring between the exposures, to assess whether the stability time is shortened by tunnel vibration. The results showed that the optical path difference between the object and reference beams changed by typically 5μm during the six second period taken to recharge the laser. The system was made more stable by adding weights to the optical mounts so that they became nodal points.

A major problem encountered when performing holographic flow visualisation in the Cranfield transonic wind tunnel, was that the facility could only be operated continuously and therefore it was difficult to record a hologram of the quiescent state prior to recording the perturbed density field. The tunnel took over twenty minutes to reach suitable operating pressures and this period was too long to obtain two phase correlated holograms. This problem was partially resolved by recording a hologram at the lowest possible Mach number by adjusting the downstream throat. If the flow field at this reference Mach number is incompressible then the interferogram will be identical to one taken with static air.

In practice, Mach 0.44 was the lowest speed that could be obtained, which was a little higher than that claimed by Gibb [209], and too high to assume a totally incompressible flow. The throat took approximately ten seconds to close, however, which was short enough to ensure stability of the optical system. To analyse the errors introduced by recording the quiescent density field at Mach 0.44 instead of at rest, sandwich holograms (see section 3.4.4) were recorded at this Mach number with two different test section pressures (2psi and 5psi). The front hologram of the pair recorded at 2psi was then combined with the rear hologram at 5psi. White spirit was used to index-match the holograms and the two respective object waves were reconstructed by placing the sandwich in a divergent replay beam from an argon ion laser.

When a sandwich hologram is reconstructed, linear fringes may be added or removed from the interferogram by sliding the plates over one another. Similarly, parabolic phase shifts may be introduced by changing the curvature of the wavefronts in the replay beam. These techniques were used to remove the error fringes, which occurred due to movement of the optical system in the ten minute period between recording the hologram at two different stagnation pressures. The images shown in figure 68 illustrate how both infinite and finite fringe interferograms may be extracted from a single sandwich hologram.

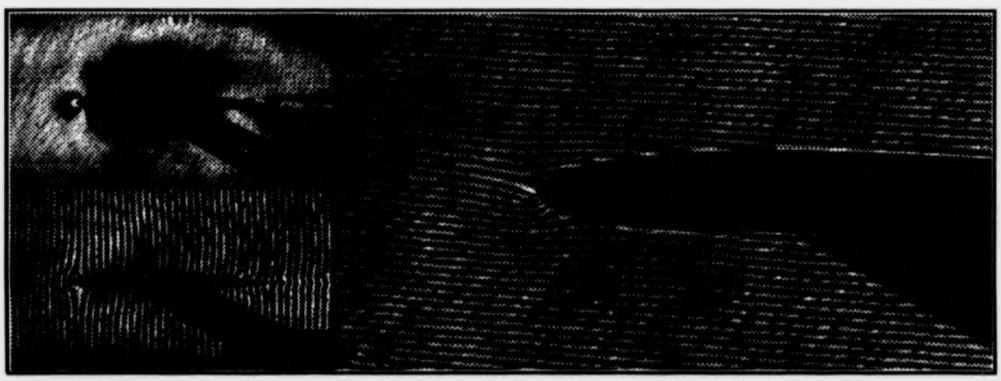


Figure 68: Multiple interferograms produced from a single sandwich hologram

The infinite fringe pattern of figure 68 indicates that there is approximately one error fringe over the field due to the compressibility of the air at Mach 0.44, for a differential pressure of 3psi. This result is shown more clearly in the wrapped phase map (figure 69) produced from one of the finite fringe patterns. The spurious fringe at the top of the image was produced by aberrations in the reconstruction system and the phase excursions over the silhouette of the model do not represent real flow information.

It was difficult to produce finite fringes with a high carrier frequency using sandwich holography, because a double-image of the model was produced when the object fields were sheared over one another. These results show that sandwich holography is a useful method for assessing the approximate magnitude of density changes, although it is insufficiently accurate, and too impractical, for general extraction of aerodynamic data.

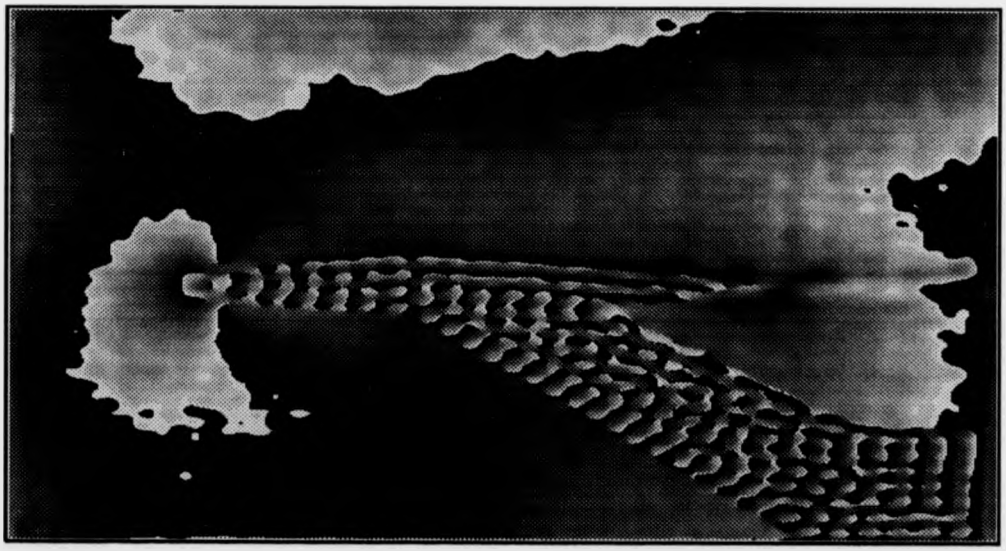


Figure 69: Wrapped phase map of the differential density field at Mach 0.44

The majority of the wind tunnel runs were made at 8psi, which was the highest stagnation pressure that could be used without the risk of 'surges' occurring in the

flow. At this stagnation pressure, there are approximately two error fringes over the entire field due to the use of a compressible reference density field. These errors are too large to be neglected and therefore when the experimental results were compared to those from a computational fluids dynamics (CFD) code, a numerical solution at Mach 0.44 was subtracted from the predictions at all other Mach numbers.

6.5 Results with Specular and Diffuse Illumination

6.5.1 Qualitative Assessment of Interferograms

Once the diffuse and specular illumination holographic interferometers had been stabilised, a number of infinite fringe holograms were recorded of the density field around the NACA 0012 two-dimensional aerofoil at freestream Mach numbers of 0.80 and 0.90. An infinite fringe interferogram at Mach 0.90, recorded in specular illumination and reconstructed in white light, is shown in figure 70. A comparison of this image, with one recorded in diffuse illumination (figure 71), shows that the resolution and fringe contrast has been improved by using specular wavefronts to record the hologram and light of reduced coherence during reconstruction. However, figure 72 shows how the viewing angle may be changed when capturing interferograms from a single diffuse hologram.

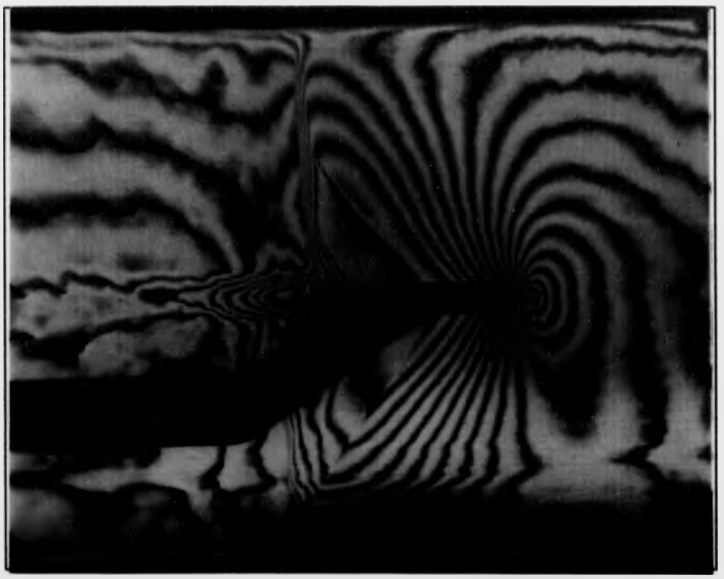


Figure 70: Specular illumination interferogram of the flow around a NACA 0012 aerofoil at Mach 0.90

In figure 70, the diagonal discontinuity in the fringe pattern, which extends back to the normal shock at the rear of the aerofoil, signifies the boundary between subsonic

Figure 71: Diffuse illumination interferogram of the flow around a NACA 0012 aerofoil at Mach 0.90

Figure 72: Diffuse illumination interferogram of the flow around a NACA 0012 aerofoil at Mach 0.90 from a lower viewing position

and supersonic flow. A lambda foot is formed at the base of the shock and separation occurs where it intercepts the boundary layer. From the approximate symmetry of the picture, the flow appears to be uniform; although a discontinuity in the fringe pattern near the floor of the tunnel indicates that there is a local anomaly in the flow.

An infinite fringe interferogram recorded in specular illumination at Mach 0.80, is shown in figure 73. In this picture there are no strong shock waves present, and the flow accelerates and decelerates isentropically over the aerofoil. A discontinuity in the fringe pattern on the top surface of the model, however, indicates that wavelets are forming. The stability of the flow at this Mach number was subsequently investigated by recording a sequence of differential interferograms with different pulse separations. Temporal instabilities were most clearly defined when a pulse separation of $5\mu s$ was used, and the reconstructed image (figure 74) exhibits wavelets near the surface of the aerofoil and bursting features in the boundary layer. The pulse separation and width of the wavelets, suggests that they are oscillating at 200kHz with an average speed of approximately $250ms^{-1}$.

Anomalous fringes close to the floor of the tunnel indicate that flow disturbances similar to those at Mach 0.90 are present. The fringe densities in figure 73 are not as high as those at Mach 0.90 and, because there are no strong shock waves present, this result is particularly suitable for quantitative analysis. Consequently, both wide and narrow angle dual reference beam holograms were recorded at this Mach number for analysis by phase stepped and Fourier transform fringe analysis.

6.5.2 Comparison of Specular and Diffuse Illumination Holographic Interferometry

The results in section 6.5.1 demonstrate that both diffuse and specular illumination holograms may be used to visualise transonic flow in a wind tunnel. However, the quality of the fringe patterns must be assessed to determine which method is appropriate for use as a routine flow diagnostic. The relative merits of the two modes of illumination are summarised in table 8.

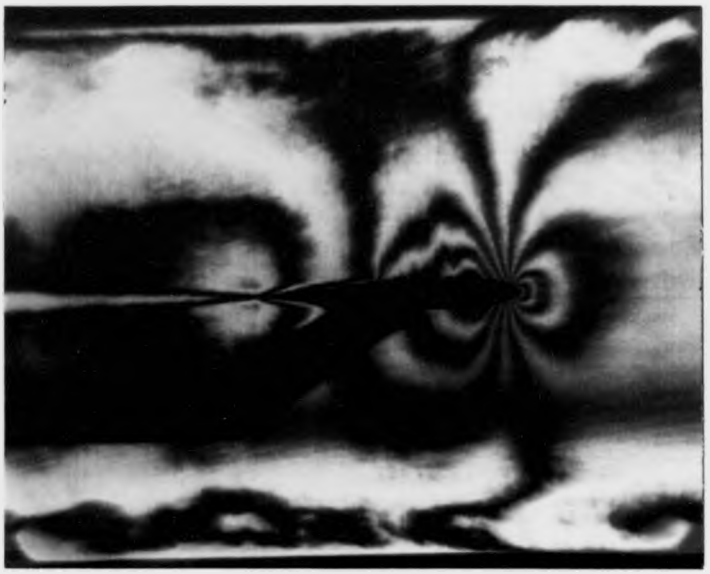


Figure 73: Specular illumination interferogram of the flow around a NACA 0012 aerofoil at Mach 0.80

	Diffuse Illumination	**Specular illumination**
Viewing Capability	3D for large holograms	2D
System Complexity	low	medium (imaging optics)
Optical Efficiency	low	high
Speckle Noise	high	low
Fringe Contrast	often low	high
Fringe Localisation	yes	no
Stability Requirements	high	medium
Laser Power Requirements	>200mJ/pulse	$\sim$ 1mJ/pulse
Diffraction Noise	low	low if image planed

Table 8: Comparison of diffuse and specular illumination holographic interferometry

Figure 74: Differential interferogram of flow around a NACA 0012 aerofoil at Mach 0.80

Diffuse illumination holograms are very spectacular because they provide a range of viewing angles common to pictorial holography. The projected views are misleading, however, because the fringes are the result of an integration through the flow and therefore cannot be considered as localised in space. When a camera is used to record the interferograms, it is frequently difficult to obtain high contrast images due to fringe localisation. This is illustrated in the photograph of figure 72, which was only presentable after the contrast of the image had been boosted during printing. The depth-of-field may be increased to improve fringe visibility, however this results in image degradation by speckle noise (section 3.4.1). Recording systems which use diffuse illumination are also prone to vibration because of the many ray directions propagating through the flow. Fringe errors appear superimposed on the interferogram and further localisation problems are produced.

In contrast, specular illumination holographic interferometry overcomes many of these problems by propagating light in a single direction through the flow. Although interferograms recorded using specular illumination are visually uninspiring, the digitised images are conducive to fringe analysis. Consequently, most of the holograms produced during the test were of this type, with an emphasis on dual reference beam recordings to enable accurate data extraction. The laser viewable master holograms from the diffuse illumination system were used to record white light copies. These display holograms demonstrate what happens when the viewing angle is changed and are useful for reference purposes during analysis of the specular illumination holograms.

6.5.3 Experimental Analysis of Phase Stepped Fringe Analysis

Holograms were recorded with a wide angle dual reference beam system (figure 63) to investigate the potential of phase stepped fringe analysis for processing flow field interferograms. The wave vector of the object beam was aligned to the surface normal of the holographic plate and the two reference beams were incident at azimuthal angles of $\pm 14°$. Exposures were made at Mach 0.44 and the 0.80, with the reference beam switched between recordings so that the individual object waves could be modulated

separately during reconstruction. Collimated reference beams were used to simplify reproduction of the recording geometry and the holograms were reconstructed using two replay beams (figure 75).

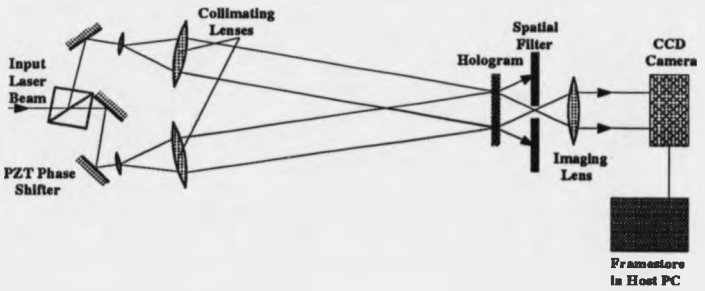


Figure 75: Optical system used to produce phase shifted flow field interferograms

Each replay beam generated two diffracted wavefronts in the direction of the original object wave: one of these converged to a focus and the other conjugate wave diverged. The conjugate beams were removed by spatial filtering to avoid parasitic interference. An interferogram was formed by imaging the remaining reconstructed object wavefronts onto a CCD camera. Initially an infinite fringe condition was not obtained, because of an angular discrepancy between the beams, and differences in wavefront curvature. These phase errors were removed by aligning one of the collimating lenses more accurately while viewing the 'live' fringe pattern. The lens was translated in a direction orthogonal to the wavevector to correct for linear (tilt) phase differences and in a longitudinal direction to compensate for parabolic errors.

The reconstructed interferogram was phase shifted by a PZT located in one of the replay beams. The PZT (PI Physical Instruments PZ 36E/37E series) was fitted with a strain gauge to compensate for hysteresis, non-linearity and drift. The effect of the strain gauge is illustrated by the two calibration curves in figure 76, which were produced before and after the actuator had been fitted. The phase stepping reconstruction system was calibrated by determining the applied voltage required to produce a 2π phase shift of the interferogram. This voltage was divided by the number of phase steps to determine the incremental voltage required for each step (the device

operated linearly with the strain gauge fitted). The accuracy of this process was subsequently confirmed by determining the angle of the PZT to the incident beam and calculating the expected voltage from figure 76.

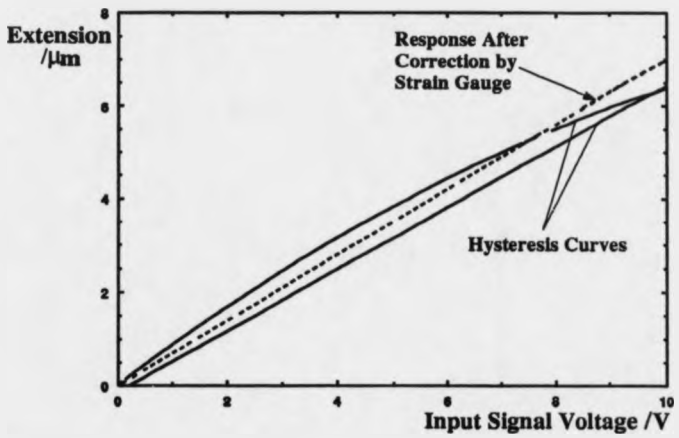


Figure 76: The control of PZT hysteresis by the use of a strain gauge

The reconstruction system was found to be highly prone to vibration, mechanical relaxation and air currents. These problems were severe due to the large separation of the expanded beams and they imposed a fundamental limitation on phase extraction accuracy. The fringe pattern was stabilised by mounting the optics on a pneumatically supported optical table, which was completely enclosed to prevent air disturbances. Even when these precautions had been taken, periodic phase shifts were observed which had a magnitude of approximately 1/10th of a fringe. These perturbations were mainly produced by oscillations of the optical table. When the pressure in the pneumatic supports was reduced, the coupling became less stiff and a lower oscillation frequency was observed. Unfortunately, the laboratory was located on the first floor and it was impossible to find a resonant frequency that was not excited by the building. The stability could have been improved if the path lengths of the divided beams were shortened. Alternatively, the images could have been captured in quick succession by rapidly stepping the voltage across the PZT.

To investigate the accuracy of phase stepped fringe analysis, four images were

recorded with the phase incremented by $\pi/2$ between successive frame captures. The resulting fringe patterns are shown in figure 77, where the order of the images is from left to right and top to bottom. The wrapped phase map was calculated from these images using equation 69 and the resulting image is shown in figure 78. The principal phase values were computed using the atan2 FORTRAN function and therefore the phase domain extends from $-\pi$ to π.

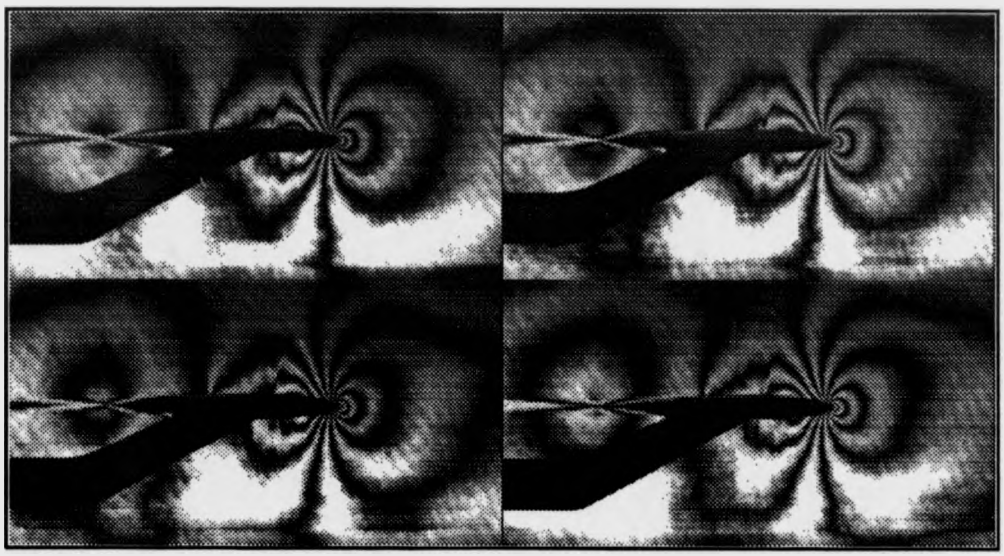


Figure 77: Series of $\pi/2$ phase-stepped interferograms

The accuracy of the wrapped phase map (figure 78) is questionable due to environmental disturbances, and an indication of the errors is provided by calculating the actual phase step using equation 76. The correct phase step should be $\pi/2$ over the entire field; however from the result shown in figure 79, it is apparent that the solution is also a function of the phase of the interferogram $\phi(x, y)$. This arises because the phase steps are unequal and therefore $\phi(x, y)$ is not eliminated from the equation. From these results it is apparent that even in an optics laboratory it is difficult to obtain accurate phase shifted interferograms, unless the successive frames are captured at video rates, or the apparatus is phase-locked.

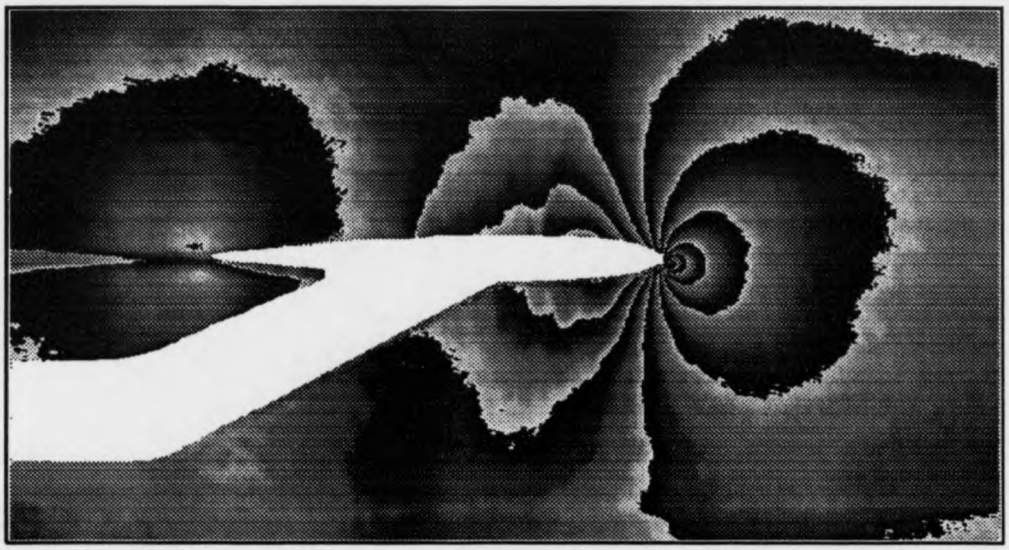


Figure 78: Wrapped phase map calculated from the four phase stepped interferograms

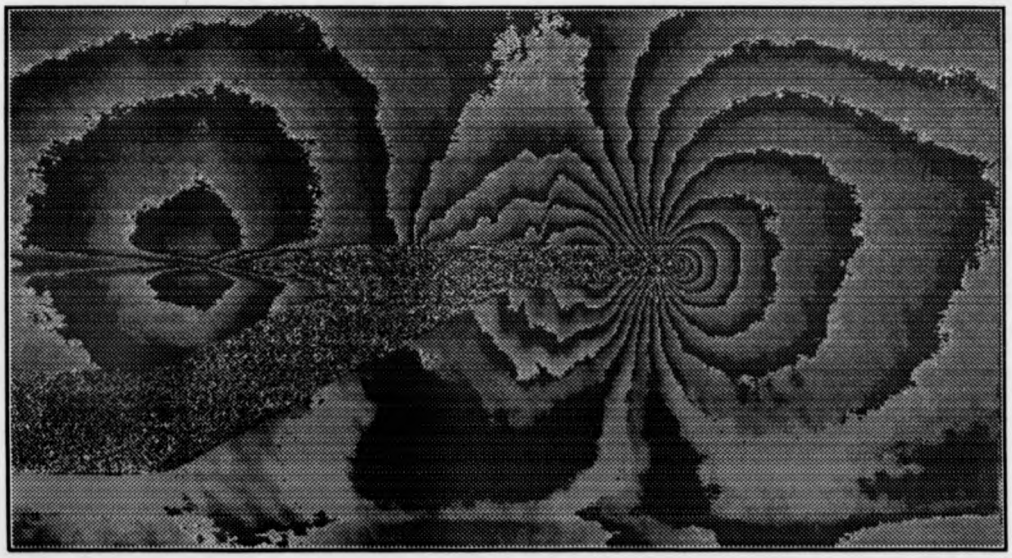


Figure 79: Estimation of phase stepping errors by calculation of the phase step

In summary, the major practical limitations of implementing quasi-heterodyne holographic interferometry in an industrial environment are:

- The holographic interferometer is bulky because two widely spaced reference beams are used.
- The reconstruction system requires interferometric stability and is difficult to maintain in accurate alignment.
- Laser illumination is required to reconstruct the holograms, thereby increasing cost and safety hazards.
- It is difficult to judge whether an infinite fringe condition has been obtained in the reconstructed interferograms due to linear and parabolic phase errors.
- The use of two widely separated replay waves leads to differential aberrations between the beams and corresponding phase errors in the processed data.
- The PZT phase shifter is expensive and requires accurate calibration.
- Phase errors are produced by intensity quantisation and recording non-linearity.

These practical problems must be addressed before phase stepped interferometry can be used routinely for extracting aerodynamic data. Many of these limitations have already been overcome by the optical shop testing industry and therefore it is worthwhile considering the potential accuracy of the technique.

The accuracy of holographically derived flow data is governed by the phase measurement accuracy of the interferogram. In figure 70 a phase excursion of approximately 70 radians occurs between the freestream region and the stagnation point at the leading edge of the aerofoil (not all of the original image is shown in the picture). The accepted accuracy of phase stepped fringe analysis is 0.1 radians (section 4.2.3), which constitutes an upper measurement accuracy of 0.14%. If the density field is converted into a Mach number distribution, by use of the isentropic flow relations, this

equates to a maximum velocity error of approximately $0.4ms^{-1}$. In the experiments presented in this thesis, the accuracy of the wrapped phase map was limited to approximately 0.6 radians by environmental disturbances and therefore the corresponding velocity error is increased to $2.4ms^{-1}$.

In comparison, PIV and LDA incur velocity errors due to the delayed response of the seeding particles. At transonic speeds these errors are in excess of $0.4ms^{-1}$ (section 2.1.1) and therefore holographic interferometry is competitive in terms of measurement accuracy. In addition, holographic interferometry does not require particle seeding, which is a considerable advantage at transonic and supersonic speeds where mass flow rates are high. A disadvantage of holographic interferometry is that data analysis is difficult for highly three-dimensional flows and therefore methods which use particle seeding may be advantageous in these circumstances.

This study has shown that wide angle dual reference beam holographic interferometry, with phase stepped fringe analysis, is a competitive method of obtaining quantitative flow information at transonic and supersonic speeds. The practical difficulties of operating the data acquisition system in an industrial environment are not unsurmountable; however it is desirable to develop a technique which requires a simpler reconstruction system that is less prone to vibration and misalignment. Consequently, a narrow angle dual reference beam system was tested, with the resulting finite fringe interferograms analysed by Fourier transform fringe analysis.

6.5.4 Extraction of Density Data using Fourier Transform Fringe Analysis

The object waves recorded on a double-exposure hologram may be encoded uniquely when the reference beams differ in angle by only 1°. A major advantage of this narrow angle geometry is that the two reference beams may be generated from the same collimating mirror, by deflecting the incident beam between exposures. The resulting optical system is simpler and more compact than a wide angle configuration.

If an image plane configuration is used to record the holograms (see section 6.3),

and a single reference beam is used in replay, the two reconstructed object waves emanate from two laterally displaced virtual source points. For analysis purposes, these points may be taken to lie on a virtual object located behind the hologram (figure 80).

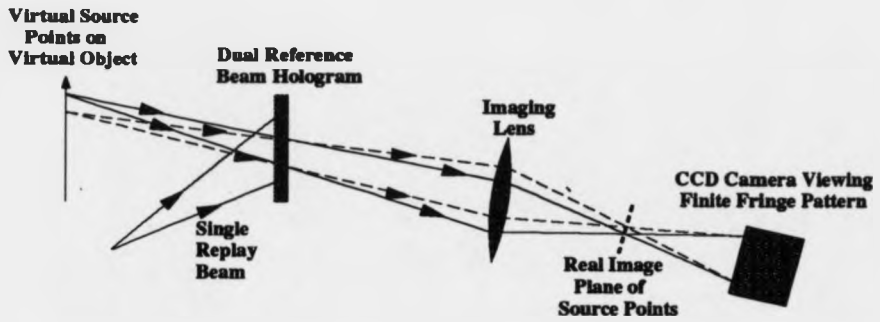


Figure 80: The reconstruction of a finite fringe pattern from a dual reference beam hologram

A positive lens may be used to image these virtual source points, so that the object waves subsequently diverge and overlap. When the angle θ_r between the reference waves used to record the holograms is small ($\sim 1°$), the object waves cross one another and an image of the wind tunnel working section is formed where the beams intersect. A finite fringe pattern is produced in this plane because of the inclination of these wavefronts and, if emulsion shrinkage of the hologram is neglected, the mean spatial frequency $\overline{f}$ of the fringes is given by:

$$\overline{f} = \frac{2}{\lambda_2} \sin\left(\frac{\theta_r \lambda_2 m}{\lambda_1}\right) \tag{114}$$

Which for small θ_r may be approximated by:

$$\overline{f} \approx \frac{2\theta_r m}{\lambda_1} \tag{115}$$

Where,

m - magnification of the reconstruction system

λ_1; λ_2 - recording and replay wavelengths

From equation 114 it is apparent that when θ_r is small, the spatial frequency of the finite fringes is independent of the reconstruction wavelength.

The carrier frequency can only be modified by either recording fresh holograms with a different angle between the reference beams, or by changing the magnification of the reconstruction system. The magnification is easily adjusted and therefore it appears that the carrier frequency may be controlled during replay, to facilitate the analysis of flow information on different scales. This assumption is incorrect, however, because the image size increases as the carrier frequency decreases, ensuring that the spatial frequency of a particular flow feature is always the same percentage of the carrier frequency. Separate control of the carrier frequency and encoded phase information is possible, however, if two reconstruction waves are used (section 6.5.3). Unfortunately, this system requires far greater stability due to the larger beam separation, as shown in section 6.5.3 on phase stepped fringe analysis. Hence, the carrier frequency should be encoded at the correct frequency during the wind tunnel test, because it is difficult to rectify an error during reconstruction.

A finite fringe pattern at Mach 0.80 is shown in figure 36. In this picture, the phase increases monotonically across the interferogram because the deviation in finite fringe spacing is small. This image was captured by a CCD camera and framestore, and processed by Fourier transform fringe analysis (chapter 5). The two-dimensional Fourier transform of the interferogram is shown pictorially in figure 81. The sharply peaked zero order term $A(u,v)$ is centred on the origin and contains the largest values in the transform, although this is not apparent in figure 81 because of saturation. The flow information is encoded in the two outer lobes, which are broadened by convolution with the amplitude noise term $B(u,v)$. The zero and first orders are spatially separated and therefore it should be possible to calculate the density field without data corruption.

The Fourier transform was processed by the method outlined in section 4.2.4, with the bias image $a(x,y)$ derived by optical spatial filtering. The wrapped phase map is shown in figure 82, with the cosinusoidal fringes replaced by a set of discontinuous

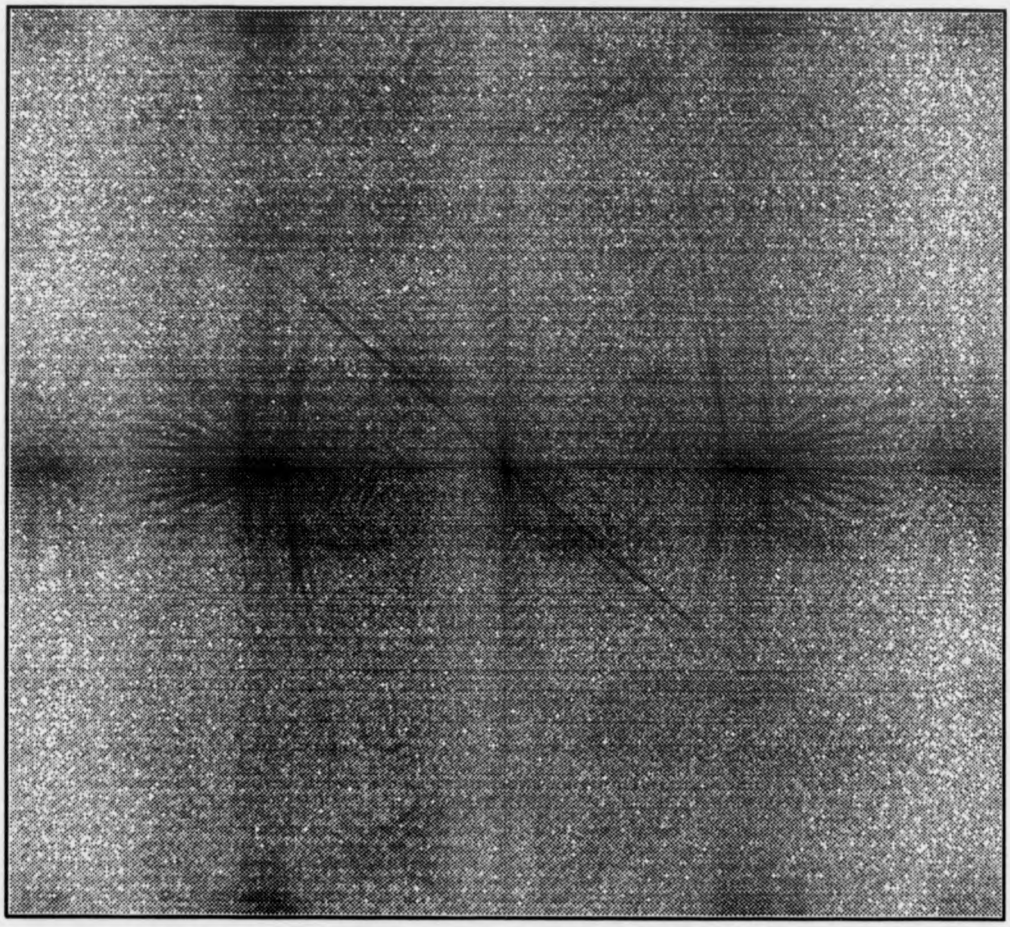


Figure 81: Logarithmic grey level representation of the 2D Fourier transform of a finite fringe interferogram

phase ramps. After Fourier processing, the image was phase unwrapped by software written by Judge [203] (chapter 5). The phase map was converted into a whole-field density distribution by calibrating the data at a single known density value and by calculating the constant differential $d\phi/d\rho$ from physical parameters (equation 56).

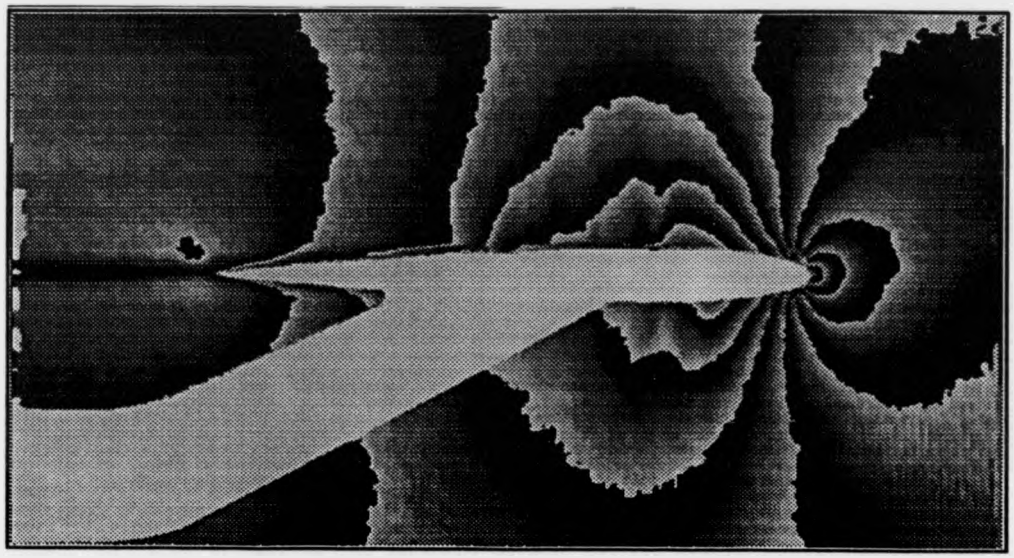


Figure 82: Wrapped phase map

The density field was calibrated by measuring the stagnation pressure in the settling chamber with a pitot probe. This pressure measurement was converted into density by use of the ideal gas law (equation 116) and equated to the phase value at the leading edge of the aerofoil.

$$\rho = \frac{p}{RT} \tag{116}$$

R - ideal gas constant

T - absolute temperature

Figure 83 shows the processed density field as a pseudo-colour image, where dark blue signifies a density of 0.51kgm^{-3} and red a density of 0.24kgm^{-3}. This image

is not simply a qualitative picture of the flow, but represents an array of 512×512 density values (actual resolution is limited to approximately three pixels by aliasing, as shown in chapter 5). Static pressure measurements were recorded along the floor of the tunnel using piezo-electric transducers (Druck model PDCR10). These devices had been pre-calibrated by performing a linear regression of pressure versus output voltage. The pressure measurements were used to assess the Mach number uniformity along the length of the tunnel and to provide additional verification of the calculated density field.

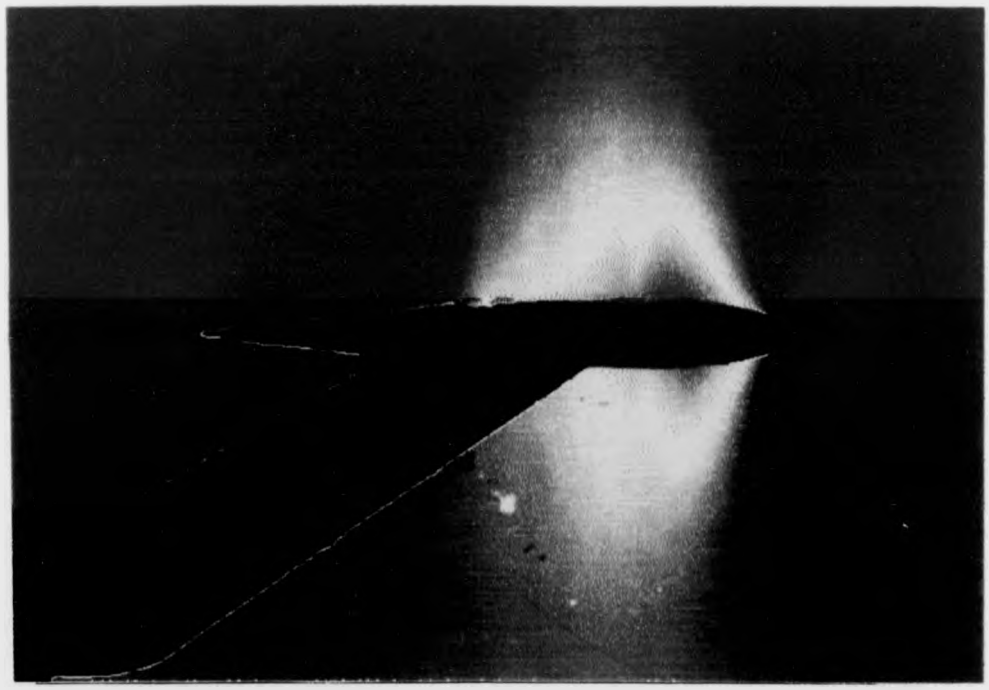


Figure 83: Density field calculated from the unwrapped phase map

In a typical run, the results showed that the Mach number decreased from Mach 0.82, at a point 268mm ahead of the leading edge, to Mach 0.79 directly below the leading edge and finally to Mach 0.77 at a distance 510mm downstream from the leading edge. The reduction in Mach number down the length of the tunnel indicates that the flow is not recovering due to 'solid blocking' and this result is confirmed in the experimental phase map by the blue region behind the model. It would have been

desirable to obtain static pressure measurements on the surface of the model, as these values would have provided a more accurate assessment of the experimental results. Unfortunately this was not possible due to the expense of pressure tapping the aerofoil. Additional validation of the experimental data was provided by comparing the results to a Navier-Stokes CFD prediction (figure 84). A laminar boundary layer was assumed in the theoretical prediction, whereas it was suspected that the experimental boundary layer may be turbulent.

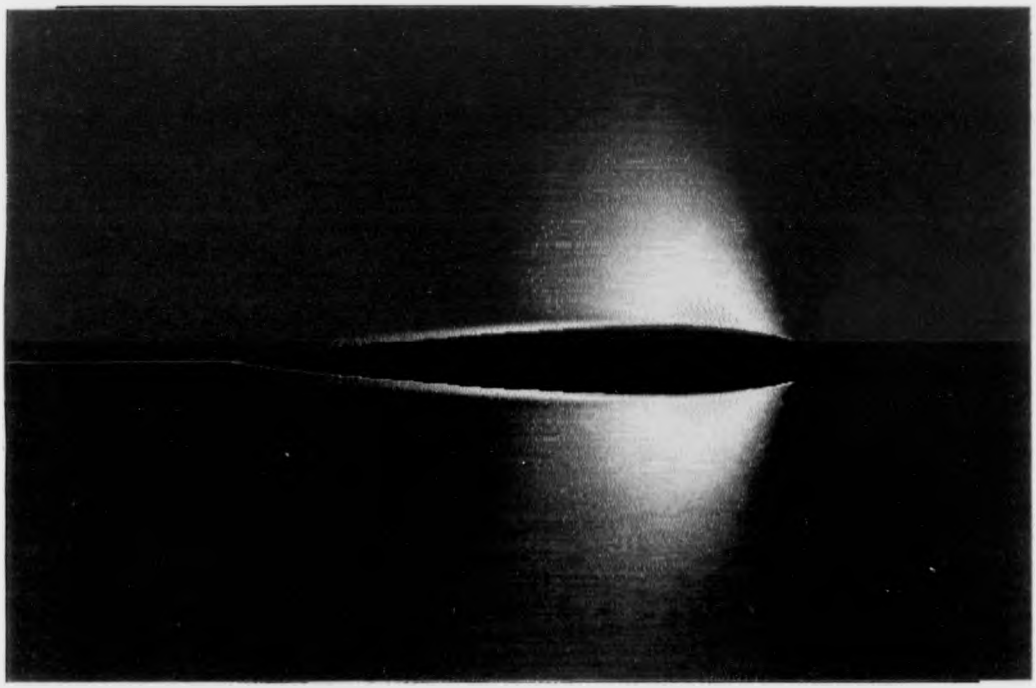


Figure 84: Navier-Stokes CFD prediction of the experimental density field

The experimental results at Mach 0.80 and 0.90 have associated Reynolds numbers of 7×10^5 and 8×10^5 respectively. With these flow parameters it is ambiguous whether the boundary layer on the model is laminar or turbulent. Insight into which flow prevails may be gained by calculating the width of a trip wire d_w required to induce turbulence (equation 117).

$$d_w > \frac{36\nu}{u}\left(\frac{ux}{\nu}\right)^{1/4} \tag{117}$$

Where,

x - distance of the wire from the leading edge of the aerofoil
ν - kinematic viscosity

Substitution of the test parameters into equation 117 indicates that a trip wire thickness of approximately 80μm is required to induce turbulence. The machining accuracy of the aerofoil is typically $50 - 100\mu$m, with a surface smoothness exceeding 1μm. These values suggest that a laminar boundary layer will be observed; although it is possible that Mach waves, shocks, and unsteadiness in the flow, will result in turbulent transition. A comparative study of the experimental and computational results should reveal the nature of the experimental boundary layer.

The experimental and computational results were compared quantitatively by calculating a whole-field difference map. The discrepancies were subsequently expressed as a percentage of the difference in stagnation densities at the two Mach numbers:

$$\epsilon_\rho = \frac{[\rho_{EXP}(0.80) - \rho_{EXP}(0.44)] - [\rho_{CFD}(0.80) - \rho_{CFD}(0.44)]}{\rho_0(0.80) - \rho_0(0.44)} \times 100\% \tag{118}$$

Where,

$\rho_{EXP}(M)$; $\rho_{CFD}(M)$ - experimental and CFD density fields at Mach M
$\rho_0(M)$ - stagnation density at Mach M

This definition was chosen because the errors are referenced to a typical density excursion in the flow. If the differences had been expressed as a percentage of absolute density then the errors would have been much smaller. However, this would have involved a less valid comparison of differential and absolute quantities. The whole-field error map is shown in figure 85 and the percentage differences in the field are defined as follows:

black - 0-1%
red - 1-5%
yellow - 5-10%
green - 10-15%
blue - 15-20%
white - > 20%

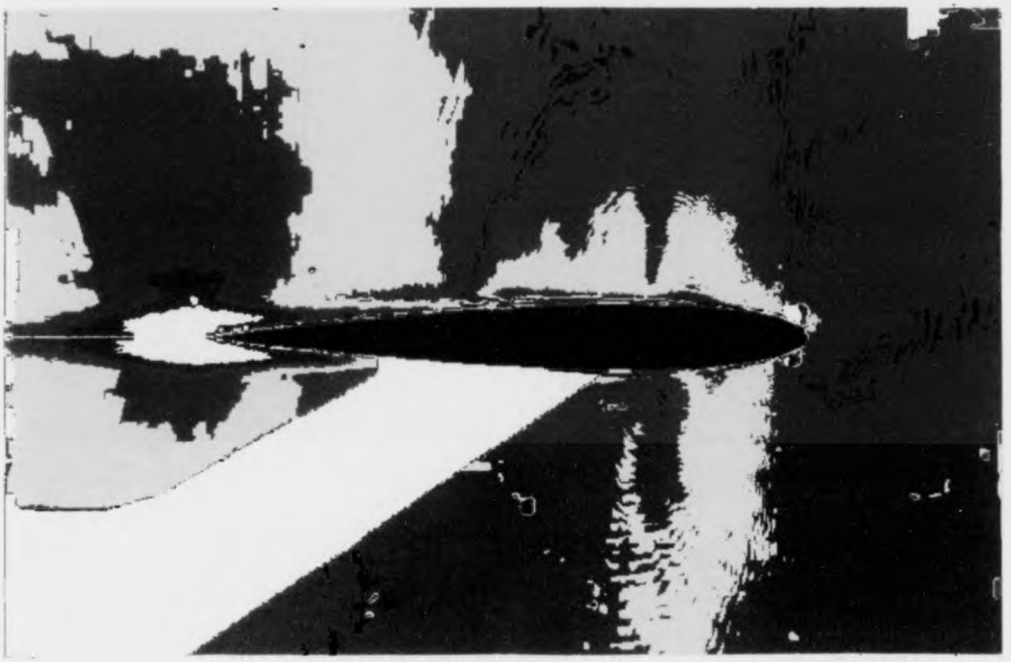


Figure 85: Pseudo-colour difference map of the experimental and computational results

The largest differences between the computational and experimental results (~25%) occur in the separated flow and wake, which suggests that the flow undergoes a turbulent transition. Differences in the regions around the aerofoil are primarily due to unsteady transonic flow. These transient features would probably be ameliorated by temporal averaging of the experimental data, although this would inevitably result in a loss of understanding of the fluid dynamics.

Confinement of the flow by the floor and ceiling of the wind tunnel was not modelled by the CFD code, but the effect on the experimental data will be small as slotted liners were used. Errors due to tunnel blockage are approximately 1% and were ac-

counted for in the Mach number measurement. The side walls were not perforated and therefore boundary layer growth over the windows, amounting to 5-10% of the tunnel width, introduced further modelling deficiencies. In practice, these were reduced by using mildly compressible reference density field for the first exposure, and by calibrating a single experimental measurement against a known value in the flow. The presence of the pitch fork sting, errors in flow incidence angle, and the presence of weak oscillatory shocks, will cause flow asymmetry. A primary source of optical errors are due to refraction, but these were neglected at Mach 0.80 due to the excellent superposition of the two reconstructed images of the model.

In conclusion, the agreement between the experimental and computational density fields is excellent in the majority of the field, with significant discrepancies only in the separated flow and wake. It is extremely difficult to model transonic flows computationally and these results demonstrate how important it is to investigate the flow experimentally. This is particularly true at transonic speeds where the streamlines are distorted a long way from the model and surface measurements provide only a partial description of the flow.

In this example it is valid to compare the theoretical CFD prediction with the experimental results because the flow is two-dimensional. When the flow is three-dimensional it is necessary to integrate through the CFD solution to produce compatible images. Although this provides some insight into the similarity of the flows, caution must be exercised in making quantitative comparisons because the integrated CFD data does not possess a unique solution. Also, there may be a non-linear relationship between any differences in the flow and integrated data, so that the experimental and theoretical images may look similar even when the flow fields are very different.

Although useful for validating CFD predictions, the scalar density field provides an incomplete understanding of compressible flow over a body. In aircraft design it is also advantageous to determine the scalar pressure field and vectorial velocity field. Pressure information is required in the calculation of aerodynamic forces and moments on a body; whereas velocity data is important in understanding fluid mo-

tion. The pressure field could be determined using holographic interferometry if two double-exposure holograms are recorded simultaneously at different wavelengths. The Gladstone-Dale constant is a function of temperature and therefore the fringe patterns could be used to deduce the pressure field from the ideal gas law (equation 116).

The local Mach number may be calculated using the isentropic relation for the density ratio:

$$\frac{\rho_0}{\rho} = \left[1 + \frac{(\gamma - 1)}{2} M^2\right]^{\gamma/(\gamma-1)} \tag{119}$$

Where,

ρ_0 - stagnation (total) density
ρ - static density
γ - ratio of specific heats (1.4 for air)

The experimental phase map (figure 83) shows that the static density reaches a minimum value of $0.24 kgm^{-3}$ at a point close to the aerofoil surface (above the boundary layer) and approximately one fifth chord from the leading edge. The stagnation density was determined as $0.51 kgm^{-3}$ from a pitot tube located in the settling chamber and therefore the highest local Mach number in the flow is 1.33. This example illustrates how a subsonic flow can become locally supersonic by acceleration over a body. The critical Mach number of the flow is 0.725 [213] and the associated sonic density is $0.32 kgm^{-3}$. The schematic diagram in figure 86 shows the different Mach number regions.

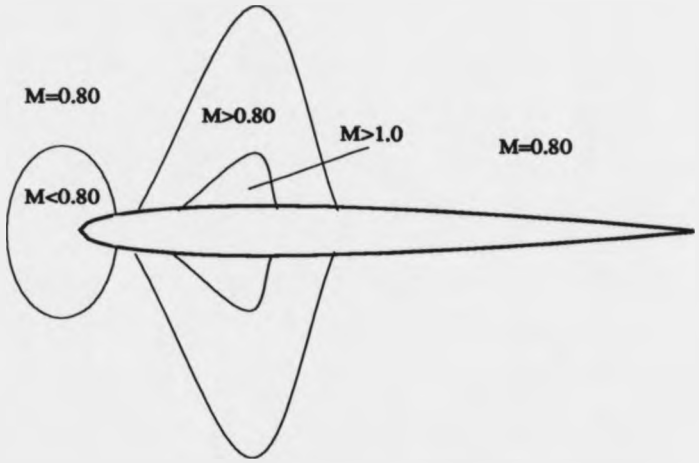


Figure 86: Mach number regions within the flow

6.6 Preliminary Study of Axisymmetric Flow Around a Cone Model

The results in this chapter demonstrate how holographic interferometry can be used to analyse two-dimensional density fields. Flows of practical interest are rarely two-dimensional and therefore it is necessary to consider how holography can be extended to fully three-dimensional phase objects. The acquisition of three-dimensional interference data is complicated and involves large optical systems. Also, computer aided tomography is required to reconstruct the projections. Consequently, it is desirable to consider an intermediate test case, which may be analysed from a single projection as well as from multiple views. An axisymmetric flow fulfills these criteria because it can be reconstructed from a single phase map by use of the Abel transform [214].

A second wind tunnel test, in a transonic wind tunnel at City University, was planned for the investigation of axisymmetric flows. This facility, which is described in chapter 7, operates intermittently and therefore enables a more efficient use of run-time. Prior to this second test, the diffuse illumination holographic interferometer installed in the Cranfield tunnel was used to identify any problems that may arise. Two axisymmetric models were chosen for this experiment:

- Cone of semi-vertex angle 8.0° and base diameter 26mm.
- Cone of semi-vertex angle 16.6° and base diameter 38.25mm.

The models were suspended by a rear cylindrical sting, which was attached to the uprights used to support the aerofoil. A diagram of the model and the sting configuration is shown in figure 87 in relation to the position of the pressure tappings on the floor of the tunnel. It was anticipated that the larger cone would produce greater density perturbations and that one these models would provide optimum fringe densities for analysis purposes. Initially, the cone of semi-vertex angle 8° was tested at both subsonic and supersonic freestream Mach numbers. Infinite fringe interferograms at Mach 1.13 and 0.99 are shown in figure 88 from an orthogonal viewing direction.

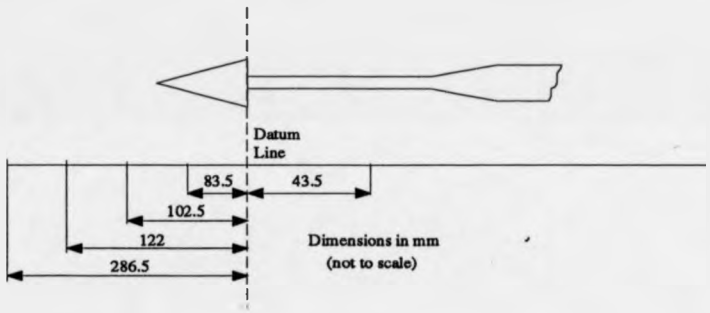


Figure 87: Diagram of the cone model and sting

When the freestream Mach number is supersonic, the flow approaching the model decelerates to subsonic speeds through a weak attached shock wave (top image of figure 88). A shock wave also appears a small distance behind the nose of the model, which is not actually distinct from the first, but rather a result of integrating through a three-dimensional shock surface. The shock surface is approximately parallel to the line of observation at the tunnel side-walls and where it attaches to the model, and therefore in these regions the fringe densities are largest. If the surface is viewed obliquely instead of orthogonally, the apparent width of the shock increases and the feature is less pronounced. At the rear of the model, the flow accelerates through an expansion fan and returns to a supersonic freestream Mach number.

Figure 88: Transonic flow over a cone with freestream Mach numbers of 1.13 and 0.99

M_∞	1.13				
Tapping No.	1	2	3	4	5
p/p_0	2.195	2.124	2.111	2.110	2.239
M	1.122	1.096	1.091	1.090	1.138

Table 9: Longitudinal static pressure measurements and calculated Mach number at a freestream Mach number of 1.13

The static pressure measurements obtained at Mach 1.13 support this general flow description, although the precise Mach numbers appear to be in conflict with the image (table 9). The readings show that the flow decelerates to approximately Mach 1.09 as it passes over the model and then accelerates back to the freestream value without going subsonic. This anomaly may be attributed to the fact that the sensors are flush with the tunnel floor and therefore measure boundary layer values. This effect is shown more clearly in an absolute interferogram of the larger cone model at Mach 1.13 (figure 89). In this picture the shock wave is distorted by the boundary layer at the floor of the tunnel and the Mach number at the top of the boundary layer approaches freestream values.

The absolute interferogram at Mach 0.99 is markedly different from those produced at supersonic speeds (lower image of figure 88). In this subsonic flow the air decelerates isentropically over the model and no shock waves are produced. The fringes are deviated most severely mid-stream, which suggests that the Mach numbers are lowest near the model. The static pressure measurements produced during this run (table 10) also show a monotonic deceleration down the length of the tunnel. Subsonic speeds are less interesting aerodynamically than supersonic Mach numbers and the associated flows are less stable. Consequently, low supersonic freestream Mach numbers were chosen for the second wind tunnel test, so that transonic flow fields are produced around the model. The results obtained with the larger cone exhibit suitable fringe densities (figure 89) and therefore cones with semi-vertex angles of $\sim 15°$ were investigated.

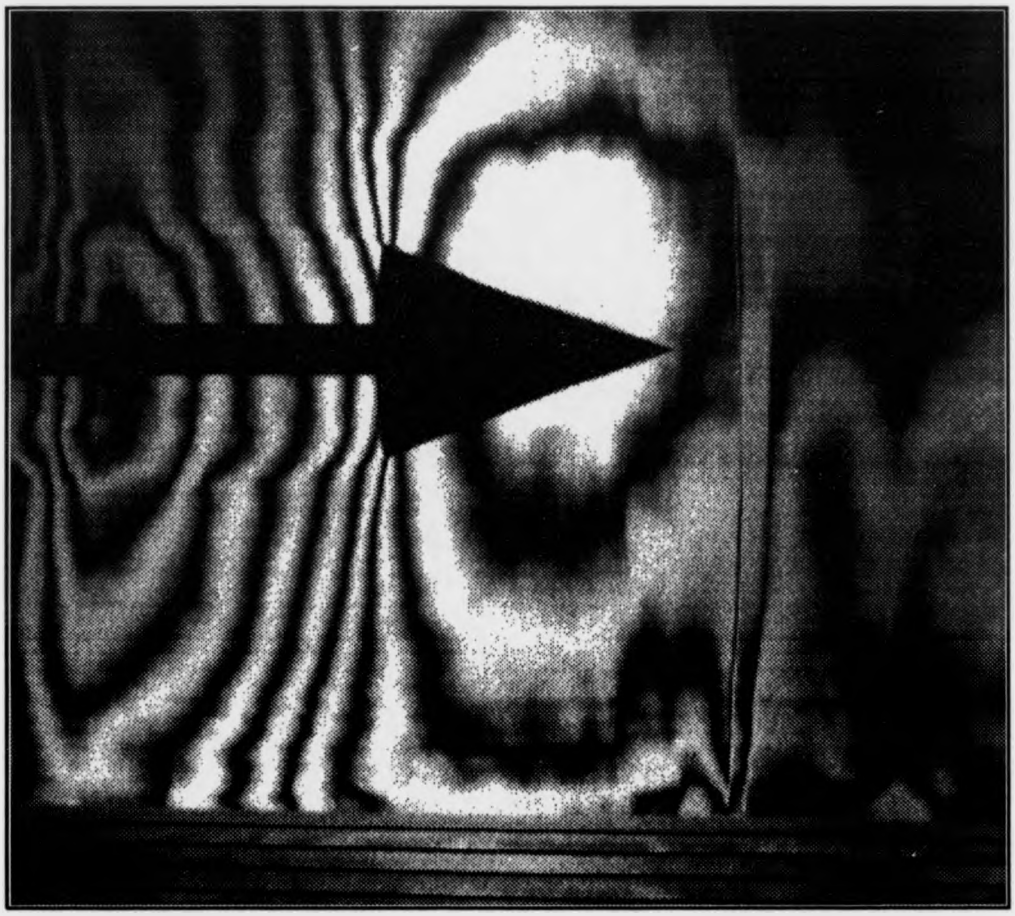


Figure 89: Transonic flow over a cone of semi-vertex angle 16.6° at a freestream Mach number of 1.13

M_∞	0.99				
Tapping No.	1	2	3	4	5
p/p_0	1.866	1.791	1.776	1.767	1.767
M	0.988	0.952	0.944	0.940	0.940

Table 10: Longitudinal static pressure measurements and calculated Mach numbers at a freestream Mach number of 0.99

7 Analysis of the 3D Flow Field Around a Cone-Cylinder Model

7.1 Introduction

The Cranfield transonic wind tunnel test demonstrated that the accuracy of holographic measurements, processed by Fourier analysis, is comparable to that obtained using PIV and LDA. Holography has the additional advantage that particle seeding is not required, thereby reducing experimental inconvenience. The ability to study unseeded flows is particularly important for transonic and supersonic testing, because at these speeds prohibitively large quantities of seeding agent are required to obtain sufficient particle number densities.

Specular and diffuse illumination holographic interferometry were compared in chapter 6 and the results showed that fringe contrast is higher for specular systems. Also, speckle noise is reduced, because the wavefronts are not scattered from optically rough surfaces. Unfortunately, specular illumination provides only a single viewing direction for each double-exposure hologram and therefore three-dimensional flows cannot be studied, unless duplicate optical systems are arranged around the working section. This configuration is highly restrictive as most flows of practical significance are three-dimensional and the test section is usually inaccessible from large viewing angles.

This chapter describes the design and implementation of a novel holographic interferometer, which overcomes the limitations of conventional specular illumination systems, by using holographic optical elements (HOE's) to generate multiple optical channels. The resulting design is exceptionally compact, because the transmission HOE's perform both collimation and steering operations. Consequently, the system is easily isolated from air currents and the optics may be located on common masses to increase mechanical stability.

7.2 Wind Tunnel Description

The multi-channel holographic interferometer was installed in an intermittent transonic wind tunnel at London's City University. The tunnel was designed by Armstrong-Whitworth Aircraft Ltd in 1952 and used at Whitley, Coventry, before its transfer to the Engineering Department in 1968. A closed circuit design is used to circulate the air (figure 90), with an induced flow produced by a steel ejector box located downstream of the working section.

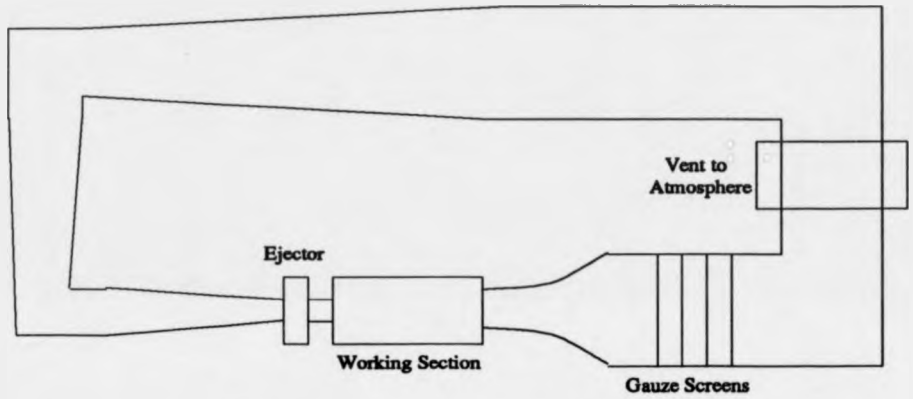


Figure 90: City University transonic wind tunnel

The closed circuit consists of a series of diffusers and steel cascades set between wooden spacers. Excess air is collected by a parallel sided square duct in the final leg of the circuit and vented through a louvre in the laboratory wall. The settling chamber, prior to the working section, consists of four gauze screens and is fitted with a pitot tube to measure total pressure. Four ejectors are used to drive the tunnel, which are fed from a manually gated supply pipe (200mm diameter). The reservoir of compressed air is provided by two $9.6m^3$ storage tanks working in parallel at a maximum operating pressure of $0.7MNm^{-2}$.

The working section configuration (figure 91) is based on an original design by Wright and Ritchie [215]. The internal width is 254mm and the height increases linearly from 192mm to 206mm over its working length of 610mm. Four slots are incorporated in the liners, to reduce boundary constraints on the flow and minimise

the possibility of reflected shocks. The slots open to a width of 7.62mm over the first 150mm of the throat, to accommodate boundary layer growth, and remain at this spacing for the remaining 450mm.

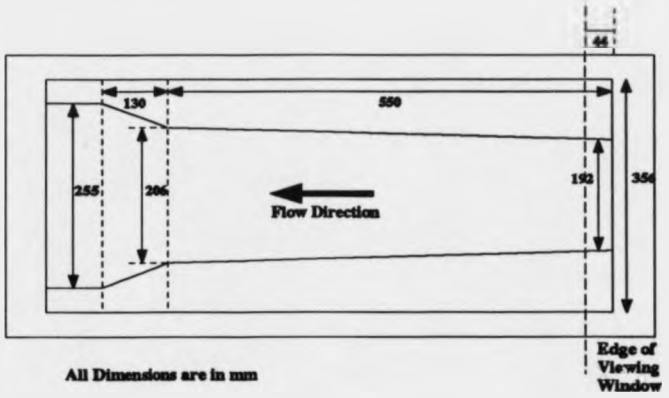


Figure 91: Working section of the City University transonic wind tunnel

In the region where air is drawn from the plenum chamber into the main flow, the height of the working section increases to 255mm, with a wall divergence of 11°. The plenum chambers are both 76mm deep and pressure tapped to enable measurement of static pressure. The freestream Mach number M_∞ was calculated from an average of these static pressure measurements, and total stagnation pressure in the settling chamber p_0, using the relationship between local and stagnation conditions for isentropic compressible flow:

$$\frac{p_0}{p} = \left(1 + \frac{\gamma - 1}{2} M_\infty^2\right)^{\frac{\gamma}{\gamma-1}} \tag{120}$$

Freestream velocities from rest to Mach 1.2 were obtained by controlling the mass flow rate through the ejector supply valve. The results from the Cranfield test (section 6.6) show that speeds in the range Mach 1.0 to 1.2 are suitable for generating transonic flow over conical bodies. The pressure tapped wind tunnel models were chosen from a selection used for undergraduate teaching. Instrumentation of the models enabled measurement of the static pressure field and simplified calibration of the holographic phase maps.

Flow quality in the working section was investigated by Sykes shortly after installation of the tunnel [216]. This study, which is outlined in appendix 3, shows that the Mach number is uniform throughout the working section and that turbulence levels are acceptable for comparing experimental measurements against CFD predictions. Unlike the Cranfield wind tunnel, the City University facility operates intermittently. The advantages of intermittent testing for holographic data capture are summarised below:

- Lower energy consumption
- Fewer low frequency vibrations
- Rapid flow termination for recording of the flow-off condition

The accompanying disadvantages of intermittent testing are:

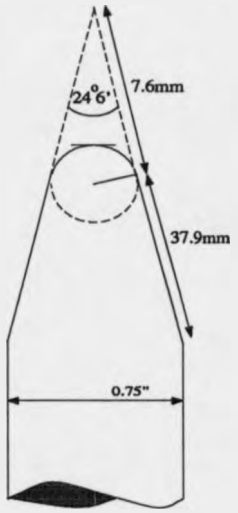
- Increased possibility of observing spurious transient flow features
- Larger pressure changes are experienced in the working section

The main advantage of intermittent testing is that energy consumption is significantly reduced. This is particularly relevant when using holographic interferometry, because the whole of the flow information is recorded in a few seconds. A continuous facility is only strictly necessary if unsteady features are studied over a long period using a real-time interferometer. The City University intermittent tunnel employs a 150kW compressor that operates for approximately five minutes per run. In contrast, the Cranfield continuous running tunnel consumes an average 500kW, with a hologram taken every twenty minutes. Hence, the intermittent tunnel provides an order of magnitude energy saving, while supporting an equally sized working section at comparable speeds and Reynolds numbers.

An intermittent tunnel is passively driven and therefore mechanical parts are stationary during blow down. Consequently, fewer low frequency vibrations are introduced than with a continuous tunnel and the optical system can be made interferometrically stable by loading the components to shift their resonance to lower frequencies. The Cranfield transonic wind tunnel suffered large amplitude low frequency vibrations from its two 550hp motors and the optics had to be loaded excessively to achieve stable operation. With the intermittent tunnel, however, sufficient stability was provided by the intrinsic mass and stiffness of the magnetic bases used to hold the optics.

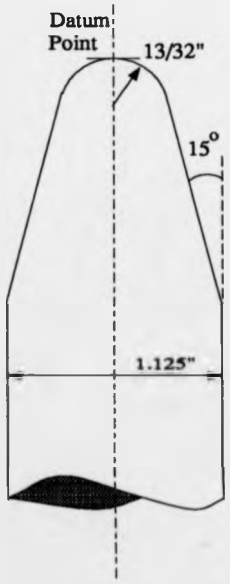
Holographic interferometry requires the recording of two correlated wavefronts, with the flow activated between the exposures. The interval between the two exposures should be minimised to prevent production of erroneous vibration fringes. With a continuous running tunnel, it takes several minutes to accelerate the air from rest to the operating Mach number and pressure. This is too long to maintain stability and it is necessary to use dual plate sandwich holography to remove the error fringes [152]. Alternatively, a throat can be adjusted to reduce the flow to low subsonic incompressible speeds for recording of the 'flow-off' exposure (chapter 6). However, this approach complicates data analysis if the flow is not entirely incompressible at the lower Mach number. In contrast, the intermittent tunnel can initiate the flow within the charging cycle of the pulsed laser.

Three pressure tapped models were tested during the wind tunnel runs. Initially, a small blunt nosed cone-cylinder model, with a semi-vertex angle of 12° and 19mm base diameter, was studied to avoid tunnel blockage (2.3% of the working section occupied by the model). The positions of the pressure tappings are tabulated in figure 92, alongside a diagram of the model. Two larger models, with base diameters of 28.6mm, were subsequently tested to increase the density changes in the flow. The first of these larger models had a blunt nose, with a half angle of 15° (figure 93). It was anticipated that a bluff body would produce a strong detached shock at supersonic freestream Mach numbers. The second large model had an identical internal angle and a pointed nose (figure 94).



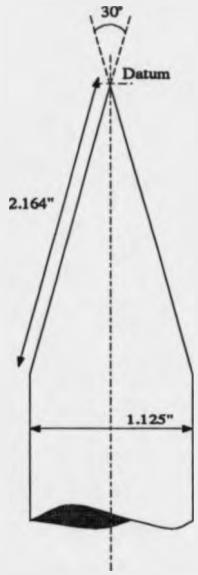
No	Distance/mm	Angle
1	1.016	39.5°
2	45.827	24°
3	36.830	23.5°
4	39.177	37°
5	1.864	11.5°
6	66.145	36.5°
7	3.348	23°
8	5.842	19.5°
9	9.398	29.5°
10	13.980	27.5°
11	19.096	25°
12	24.679	30.5°
13	30.785	15°
14	Not used	18°

Figure 92: Geometry and location of the pressure tappings on the small bluff body



No	Distance/mm	Angle
1	2.870	0°
2	6.240	22°
3	7.535	41°
4	8.895	42°
5	10.275	40°
6	15.165	20°
7	21.900	79°
8	24.770	21°
9	27.165	20°
10	29.635	20°
11	42.400	20°
12	55.035	19°

Figure 93: Geometry and location of the pressure tappings on the large bluff body



No	Distance/mm	Angle
1	9.535	0°
2	17.660	18°
3	25.930	19°
4	34.385	21°
5	42.390	20°
6	50.745	19°
7	53.345	20°
8	59.735	21°
9	58.305	21°
10	71.015	20°
11	83.890	20°

Figure 94: Geometry and location of the pressure tappings on the large pointed model

The wind tunnel models were aligned in the working section by a rear mounted sting and calibrated quadrant. Zero incidence was confirmed by measuring the static pressure at diametrically opposed points on the cylindrical section of the model. This calibration was performed over a number of runs to average flow unsteadiness.

The flow volume was viewed by installing custom-made rectangular windows with dimensions of 720mm×255mm. Schlieren quality glass was used to minimise optical aberrations when imaging through the test section. The vertical dimension of the window was designed to match the length of standard format holographic plates, to simplify the recording and fitting of HOE's. A maximum viewing angle of 127° was provided by the full aperture of the window. In practice, an angle of 116° was used, to prevent light from propagating through the steeply divergent region of the working section. The window was made as thick as possible (25mm) to prevent the glass from cracking when the test section pressure dropped during tunnel operation.

7.3 Design of the Optical System

7.3.1 Preliminary Design Considerations

The aim of this study is to design an optical system, which combines the versatility of diffractive holographic optics, with the efficiency and simplicity of conventional refractive optics. In particular, the system must be capable of producing specular image plane holograms over a wide viewing angle; and should be compact, to reduce differential movements of the optics and minimise air currents. Also, the point spread function of each optical channel should be comparable to the grid dimensions of contemporary CFD predictions.

HOE's are analogous to Fresnel zone plates and image by diffraction instead of refraction. A collimated beam is divided into three components by a zone plate: a directly transmitted zero order and two diffracted waves. The positive first order focusses at the focal length of the element, whereas the negative first order generates a divergent conjugate beam. An on-axis element is unsuitable for imaging air flows, unless the unwanted diffraction orders can be spatially filtered. A simple solution is to use an off-axis configuration so that the diffracted beams are spatially separated.

Optical aberrations affect both diffractive and refractive components. If HOE's are designed and combined incorrectly, then the resolution of the reconstructed image is insufficient to provide accurate flow information. Chromatic aberrations are particularly severe for diffractive systems, due to the dispersive properties of grating structures. Also, aberrations are produced when imaging extended objects, because the recording and replay geometries are different. Specular holographic interferometry does not suffer from these limitations, because the illumination is monochromatic and all of the object rays emanate from a single point when refraction through the flow is negligible.

The optical system designed to produce three angularly spaced views of a cone-cylinder model is shown schematically in figure 95. The source and image points are

positioned higher than the wind tunnel axis to ensure that the central HOE's are off-axis and generate separated diffraction orders. In addition, this design simplifies the introduction of an overhead reference beam. A more detailed plan view and side elevation of the system is shown in figure 96.

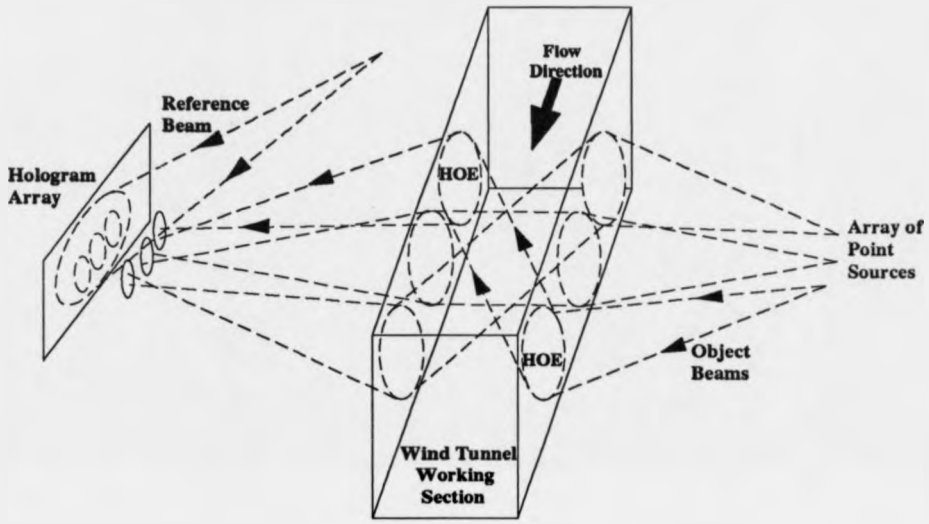


Figure 95: Multi-directional holographic interferometer

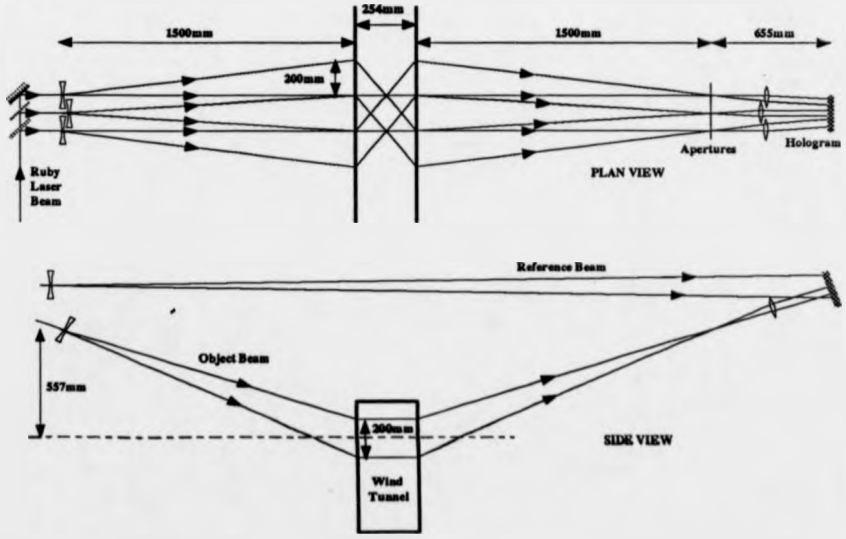


Figure 96: Plan view and side elevation of the multi-directional holographic interferometer

The first array of HOE's collimates the three divergent object waves and steers them at the correct angles across the working section. A second set of identical HOE's, but rotated through 180°, perform the inverse operation. Conventional refractive optics are used to diverge the three object waves and image the beams emerging from the tunnel onto the holographic plate. In practice, the three images were of unequal intensity, due to the varying diffraction efficiencies of the HOE's, and therefore neutral density filters were used to correct this disparity.

An overhead reference beam is introduced at an angle of $\sim 20°$ to the hologram normal and attenuated so that its intensity is 1.5 times that of the object wave. Holograms recorded with this beam ratio during the Cranfield test (chapter 6), had high diffraction efficiency and low intermodulation noise. Finite fringe interferograms were recorded by rotating the reference beam about a vertical axis between exposures, using a steering mirror mounted on a motorised rotation stage. The angle of rotation was calibrated to constrain the spatial frequencies in the phase object to the range $[-0.5f_x, 0.5f_x]$. In chapter 5 it was shown that frequencies in this waveband may be extracted to an accuracy of between 0.02 and 0.1 radians, depending on the linearity of the image acquisition system.

In most laboratory based interferometers, the object and reference beams are confined to a horizontal plane and therefore interference is optimised if the wavefronts are plane polarised in the vertical direction. The multi-directional interferometer is, however, a three-dimensional system, with the object and reference beams arranged in a predominantly vertical plane. The ruby laser produces a vertically polarised beam and therefore a half wave plate was used to rotate the plane of polarisation by 90°.

7.3.2 Imaging Considerations and Aberration Theory

HOE's do not generate an exact replica of the original object wave if the recording and replay geometries differ. In particular, aberrations are introduced if there is a shift in wavelength, change in reference or object wave source position, or a modification of the recorded fringe structure during photographic processing. The aberration

the recorded fringe structure during photographic processing. In the proposed configuration, the aberrations are small because the recording and replay geometries are identical; and the same wavelength is used to record and reconstruct the holograms. The aberration coefficients of the HOE's used in this study are derived in appendix 4.

In practice, the phase object and silhouette of the wind tunnel model are encoded onto the aberrated wavefront between the two identical HOE's. These aberrations should be small, when refraction through the flow is neglected, because all of the rays emanate from a common object point (virtual focus of the negative lens used to expand the beam during recording and replay). However, some aberrations are generated by the refractive optics used for beam expansion and imaging. Also, the collimated wave used during manufacture of the HOE's is aberrated due to the use of spherical schlieren mirrors instead of more expensive off-axis paraboloids.

The aberration coefficients may be determined by an optical design program (e.g. Genii or Code V); however, this exercise is futile if large aberrations are generated by emulsion shrinkage during photographic processing of the exposed HOE's. When amplitude holograms are fixed, unexposed silver halide is removed from the emulsion to prevent subsequent fogging and the interference planes collapse. Emulsion shrinkage is most detrimental to the imaging properties of a HOE when the interference planes lie parallel to the surface of the emulsion. Fringe orientation is parallel to the vectorial difference of the wave vectors forming the interference fringes. Consequently, when the normal of the holographic plate bisects the ray directions of the object and reference beams, the fringes lie perpendicular to the emulsion surface and the effects of shrinkage are minimised. Hence, the central grating pair of the wind tunnel system should suffer fewer aberrations than the two off-axis pairs which have slanted fringes.

7.3.3 Production of the Holographic Optical Elements

The HOE's were recorded on 254mm×203mm glass plates coated with Agfa 8E75HD emulsion. Silver halide was chosen because it has sufficient speed to enable production

of large aperture elements, with modest laser powers and exposure times (exposure requirement of $10\mu Jcm^{-2}$). Developed silver grains produce Rayleigh scatter that is detrimental to final image quality, but it was considered that other 'continuous' materials would complicate production of the prototype system. If the optical configuration proves successful, but is limited in performance by scatter, then photoresist, photopolymer and dichromated gelatin (DCG) [142], will be considered.

In this preliminary design study of the multi-directional holographic interferometer, it is necessary to calculate the required diffraction efficiency η of the HOE's, defined as the percentage of light diffracted into the desired order from the incident beam. The laser energy E required to expose the holographic plate located in the image plane of the holographic interferometer is:

$$E \geq \frac{nAs}{\eta^2 \epsilon^2 N} \tag{121}$$

Where,

- N - number of superimposed exposures
- n - number of optical channels
- A - area of exposed region on holographic plate
- s - exposure requirements of the holographic emulsion
- ϵ - proportion of beam diameter intercepted by the HOE

This calculation assumes that the beam profile is of constant intensity, which is a reasonable approximation for a ruby laser beam that does not have a true gaussian profile due to the multi-stage amplification process. Hence, the required diffraction efficiency is:

$$\eta \geq \frac{1}{\epsilon}\sqrt{\frac{nAs}{E}} \tag{122}$$

The three holographic views (n=3) were recorded on Agfa 8E75HD emulsion, which has a sensitivity of $10^{-5}\mu Jcm^{-2}$ for an exposure density of 0.5D when used with Agfa G282c developer. A ruby laser, with a 3J maximum pulse energy, was used

for both recording and reconstruction to eliminate chromatic aberrations. The area of each image was 12cm^2 with two sequential exposures used to record the interferogram (N=2). To ensure beam uniformity, the input beams to the HOE's were over-expanded by a factor of three (ϵ =1/3). Hence, from equation 122, the required diffraction efficiency is 2.3%.

The diffraction efficiency of a hologram is influenced by several factors:

- Whether the hologram is thick or thin
- Whether the hologram is amplitude or phase
- Modulation depth of the emulsion if the hologram is bleached
- Levels of scatter
- Reconstruction geometry (matching of the Bragg condition)
- Fringe profile (linearity of recording)
- Surface relief (blazing)

Hologram 'thickness' is quantified by the Q-factor, which is a function of physical emulsion thickness and other parameters:

$$Q = \frac{2\pi\lambda d_e}{n\Lambda^2} \tag{123}$$

Where,

d_e - physical thickness of the emulsion
Λ - fringe spacing
n - bulk refractive index

When Q>10 the grating is considered thick and volume properties dominate. The unswollen emulsion thickness of Agfa 8E75HD emulsion prior to chemical processing is $\sim 7\mu$m, with a mean refractive index of approximately 1.5. When the angle between

the object and reference wavefronts is $\sim 20°$, and the normal of the recording plane bisects the beams, the mean fringe spacing is 2μm. This hologram has a Q-factor of 5 (thin grating) and will only exhibit weak volume properties such as angular and wavelength selectivity. The maximum theoretical diffraction efficiency for a thin hologram with sinusoidal modulation, is 6.25% for an amplitude grating and 25% for a phase grating.

The diffraction efficiency of an amplitude hologram is modified by volume effects and Kogelnik's coupled wave theory [217] shows that the peak diffraction efficiency for a thick transmission hologram is reduced to 3.7%. With a Q-factor of 5, and a degree of exposure latitude provided by using faster developers for recording the flow hologram, this study indicates that sufficient light is available if amplitude holograms are used. Phase holograms, although more efficient, exhibit significantly higher scatter; require more complicated photographic processing for high quality results [218]; and have poor archival properties due to print-out. If future systems are developed that use more optical channels, then bleaching chemistries could be studied to minimise scatter and emulsion damage. The use of phase holograms, and a faster 10E75 emulsion, could in principle supply a forty fold increase in apparent laser energy for the implementation of more views.

Three different optical configurations were used to produce the six HOE's. The optical systems were constructed in the laboratory from plans of the wind tunnel. Care was taken to ensure that the recording and replay geometries were equal, so that aberrations were minimised. The optics were positioned using translation mounts, rotation stages and a plumb line. Fine adjustments were made using a He-Ne laser aligned co-axially with the ruby beam.

During recording of the HOE's, black float glass was index matched to the back of the holographic plates to suppress multiple reflections within the glass substrate. White spirit was used as the index matching fluid because it has a refractive index close to that of the gelatin substrate (n=1.5). Closer index matching is possible with xylene and other organic liquids, but these fluids are hazardous and were considered

unnecessary for the limited gain in recording uniformity.

Once exposed, the holograms were pre-washed in Ilfotol wetting agent. This process serves three purposes: the white spirit is removed into solution, emulsion speed is increased by hypersensitisation, and the scatter of the final hologram is reduced. The holograms were processed in undiluted Neofin Blue proprietary developer for two minutes at room temperature, followed by an intermediate wash for two minutes, and a final fix for two minutes in Kodak Amfix. Amplitude holograms developed in Neofin Blue exhibit low noise and high diffraction efficiency [219]. Also, acceptable recording densities may be obtained with this developer, when the hologram is severely under-exposed. Drying marks were eliminated from the holograms by washing them in running water and Ilfotol for ten minutes, with a final rinse in de-ionized water. Emulsion damage was avoided by limiting the duration of the fixing step and by drying the holograms slowly at room temperature. A degree of emulsion shrinkage was expected because Neofin Blue does not tan the emulsion prior to removal of unexposed silver halide during the fixing stage.

After the HOE's had been recorded, their diffraction efficiency and angular selectivity were determined to assess the optical throughput and alignment tolerances of the proposed wind tunnel system. The HOE's were mounted on a graduated rotation stage and reconstructed by an unexpanded He-Ne laser. The first diffraction order was monitored with a calibrated photodetector and intensity measurements were taken at varying angles of incidence. A graph of diffraction efficiency versus angle, for the central HOE, is shown in figure 97. The angle is defined to the surface normal of the hologram and is therefore zero when the incident beam is retro-reflected.

The results show that the peak diffraction efficiency of the central HOE is $(2.31 \pm 0.01)\%$ and the full-width-half-maximum angular tolerance is $\pm 34°$. The maximum theoretical diffraction efficiency depends on the Q-factor (equation 123) and is in the range [3.7%,6.25%]. When the HOE is recorded in silver halide, the diffraction efficiency is limited by several effects:

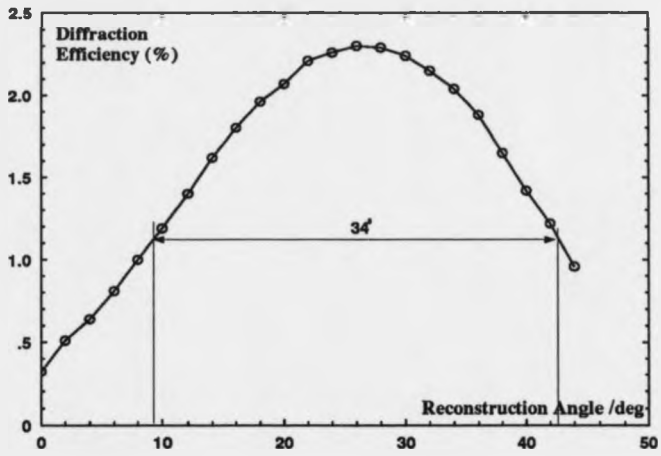


Figure 97: Diffraction efficiency versus angle for the central HOE

- Reflection from the front and back faces of the hologram
- Presence of higher diffraction orders because of recording non-linearity
- Rayleigh scatter from the developed silver grains
- Presence of parasitic gratings
- Presence of unmodulated exposure (absorption)
- Surface relief
- Failure to meet Bragg condition during reconstruction

In practice, it is difficult to exceed a diffraction efficiency of 2.31%, unless the developing chemistry, emulsion hardness and surface relief are carefully controlled. This value is equal to the minimum required for the proposed system (equation 122), although a degree of latitude is provided by using faster developers or holographic plates. The large acceptance angle of the HOE is typical of a thin hologram and simplifies realignment of the optics. The peak diffraction efficiency occurs at 22.05°, which is 1.71° larger than the recording angle. This implies that the emulsion has shrunk by 15%, which is consistent with the experimental results presented in section 6.3.

The results for the outer HOE's (figure 98) are more difficult to interpret. The diffraction efficiency is lower than the central channel at 1.61%, which is probably due to the greater mismatch in the recording and replay geometries at large beam angles. Also, the interference planes in the outer HOE's are slanted and therefore liable to greater distortion during emulsion shrinkage. The larger angle between the recording wavefronts should, however, produce a thicker hologram with a greater angular selectivity. This is not observed and the full-width-half-maximum angular tolerance is larger than the central channel at 52°. In addition, the peak diffraction efficiency occurs at approximately 17° which suggests that the emulsion has swollen. Apparent swelling of the emulsion may be explained by the fact that the fringe planes are slanted and therefore appear broader when viewed obliquely. The large acceptance angle may be attributed to viewing the hologram away from the Bragg condition, so that a sharp response is not observed. The reduction in peak diffraction efficiency is further evidence of this effect.

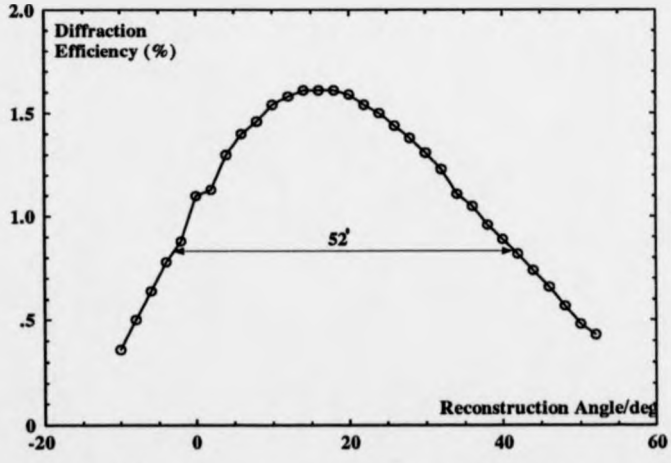


Figure 98: Diffraction efficiency versus angle for the outer HOE

In summary, amplitude transmission HOE's were recorded that generate three collimated beams which cross the working section at designated angles. Amplitude gratings were produced because they generate less Rayleigh scatter than bleached elements. Although the diffraction efficiency of these elements was only $\sim$ 2%, the

3J/pulse ruby laser provided adequate output energy during reconstruction. The HOE's had a large acceptance angle and therefore re-alignment in the wind tunnel system was simplified.

7.4 Experimental Results

7.4.1 Distortion of the Wind Tunnel Windows

The transmitting and receiving optics were mounted on two tripods. Figure 99 shows a view of the transmitting stage, which was used to generate three divergent object waves and a variable angle reference beam. The HOE's were attached to the wind tunnel windows using microscope immersion oil. White spirit was not used, because it is too inviscid to remain trapped between the glass plates for periods longer than a few hours. A photograph of the HOE's attached to the wind tunnel window is shown in figure 100 (each HOE measures 254mm×203mm).

The resolution of each optical channel was assessed by recording holograms of an ASAF 1951 test target positioned on the wind tunnel centre line. Diffraction noise in the reconstructed image was minimised by sharply focussing the wind tunnel model onto the recording plane. The correct focal position for the holographic plate was determined to an accuracy of ±5mm using the thin lens equation (equation 112). Finer adjustments were then made by recording a series of exposures, with the plate holder translated in increments of 0.5mm using a micro-positioner. The plane of best focus was judged by viewing the resolution chart exposures with an eye piece.

The bottom left hand chart in figure 101, which has a 500μm pitch, is clearly resolved by the central optical channel. At this magnification, image resolution is limited by the sampling rate of the camera, rather than the aberration characteristics of the imaging system. The diagonal wavy line across the two coarse grids is due to a scratch on the glass resolution slide.

At frequencies higher than the Nyquist sampling rate, the visibility of successive bar charts fluctuates periodically as moiré fringes are generated (explained in section 5.2.2). The resolution chart was subsequently viewed with an eye-piece to prevent camera aliasing and the narrowest grid spacing that could be resolved was 100μm. This resolution is relatively poor optically but is sufficient for analysing flow

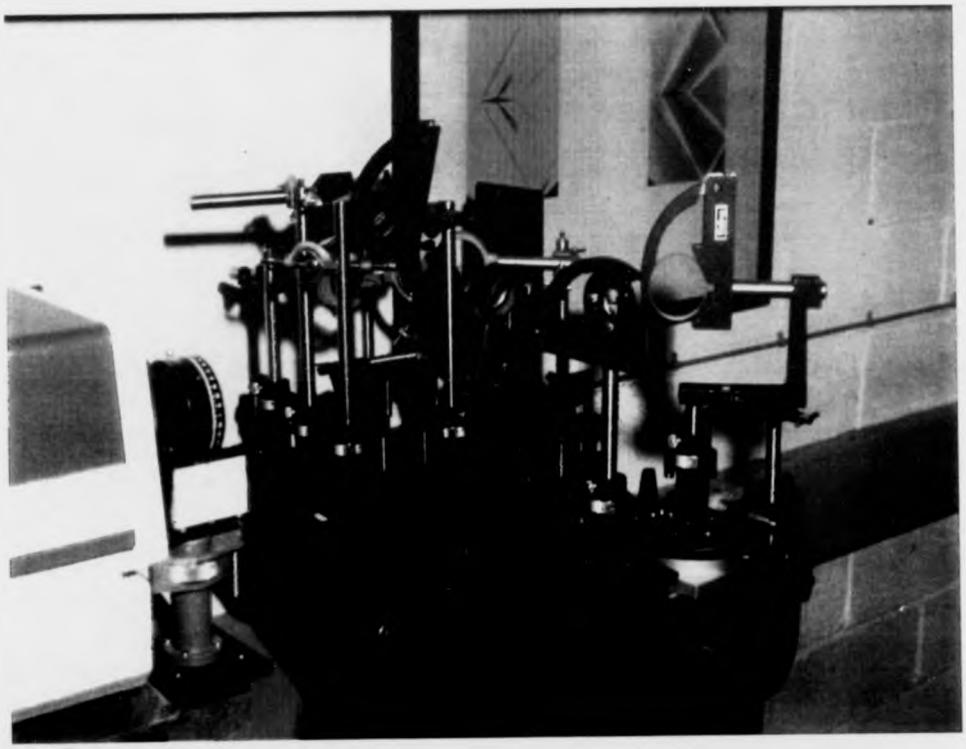


Figure 99: Transmitting optics stage

Figure 100: View of the HOE's on the wind tunnel window

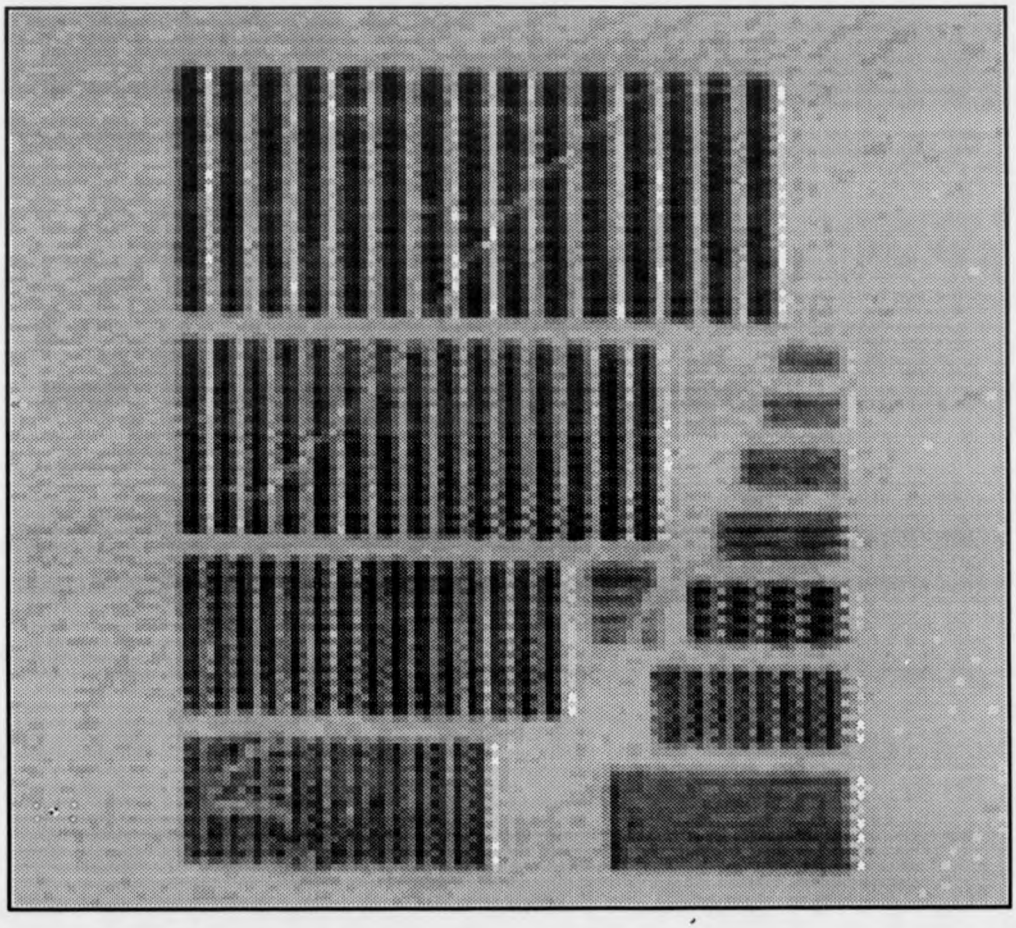


Figure 101: Imaging of a test chart by the central optical channel

fields. Duston has shown [220] that if the developing process is stringently controlled to avoid emulsion shrinkage, then a resolving power of 4μm can be achieved with HOE's recorded in silver halide. Hence, these results demonstrate that it is unnecessary to use more exotic recording materials (DCG, photopolymer and photoresist).

An image of the resolution chart produced by one of the outer optical channels is shown in figure 102. The central region of the picture is magnified to illustrate that bars can still be resolved with a pitch of four pixels (half the Nyquist frequency). This imaging system had a lower resolution of 250μm due to astigmatism: it was possible to image vertical or horizontal features to a much higher resolution, but not simultaneously. Astigmatism is produced by an anisotropic collapse of the interference planes in the emulsion, leading to different focal powers in the two orthogonal axes.

Figure 102: Imaging of a test chart by one of the outer optical channels

The magnification of the reconstruction system was designed to provide a test section field-of-view measuring 125×125mm. The interferograms were digitised to an 8-bit accuracy and a resolution of 512×512 pixels. With this magnification, the pixel size (244μm) is approximately equal to the resolution limit of the imagery and

maximum information is extracted. However, it should be noted that the resolution of the final unwrapped phase map is approximately the spacing of the unperturbed finite fringes, rather than the width of the pixels. The carrier fringes were arranged to have a period of four pixels, based on the results of chapter 5, and therefore the resolution of the phase map is approximately 1mm.

In addition to the capture of resolution chart images, a number of holograms were recorded of square grille graph paper, copied onto transparent acetate sheet. The purpose of this exercise was to optimise the alignment of the reconstruction system so that astigmatic blurring was equal in both orthogonal directions. Figure 103 illustrates that 2mm squares are clearly resolved, with equal definition of the horizontal and vertical lines.

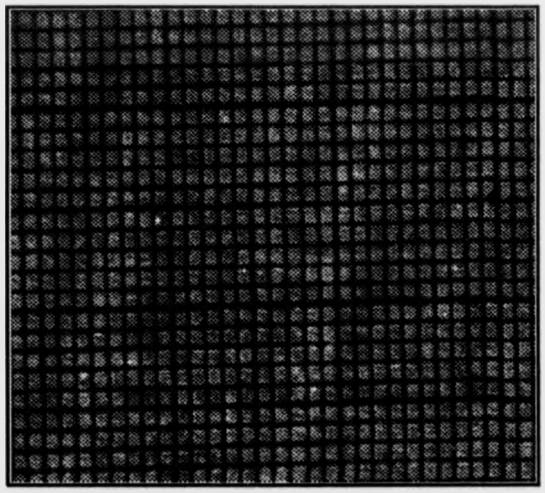


Figure 103: Holographic reconstruction of a grid of 2mm squares

It was noticed that the holograms recorded from the three optical channels exhibited markedly different diffraction efficiencies. The intensity and relative path lengths of the object and reference beams were correctly balanced, and therefore it was suspected that the polarisation vector was altered by propagation through the system. This was confirmed by monitoring the fluence of each channel, as a linearly polarising

filter was rotated in the beam. In practice, it was impossible to balance these efficiencies using a common input polarisation state, and therefore the half wave plate was replaced by a quarter wave plate. This produced circularly polarised light which gave equal diffraction efficiencies for holograms recorded by each channel.

After this preliminary study of reconstructed image quality, a number of runs were made with an empty wind tunnel to assess the performance of the holographic interferometer with a nominally uniform flow present. Initially, double-exposure holograms were produced with both recordings made at the same Mach number. This test determined the stability of the optical system and provided a quantitative measure of density fluctuations in the flow. The left image in figure 104, recorded at Mach 1.14 with a pulse separation of 300μs, is a typical result at supersonic speeds. The interferogram exhibits a circular fringe which is due to the Mach number changing slightly between exposures (explained below). Apart from this fringe, however, there are no abrupt phase changes that characterise unsteady flow. In contrast, interferograms recorded at high subsonic speeds, as shown by the right image of figure 104 produced at Mach 0.81, exhibit large unsteady regions. Transonic flow over the cone models was generated using a supersonic freestream Mach number and therefore these results provide confidence in the flow quality at these speeds.

After the dynamic behaviour of the flow had been investigated, double-exposure holograms were recorded of the quiescent and transonic flow conditions with a model present. This experiment determined whether any static deformations of the optical system occurred between exposures. Figure 105 shows an absolute interferogram of transonic flow over the large bluff model at Mach 1.21. At supersonic speeds, disturbances cannot travel upstream of the bow shock (neglecting the tunnel boundary layers) and therefore this region should be devoid of fringes. It is evident from figure 105 that there is either a hill or valley superimposed on the data. However, when the Mach number was constant for both of the exposures, the error fringes did not occur. Hence, this effect is attributable to the reduction in static pressure that occurs in the test section during blow-down. The pressure drop causes the windows to bow

Figure 104: Flow unsteadiness at supersonic and subsonic Mach numbers

inwards and therefore the rays travel a greater distance through air at atmospheric pressure.

The distortion amplitude was estimated by measuring the angular deflection of a He-Ne laser beam reflected from the edge of the window. A deviation of 3mrads occurred for an internal static pressure drop of 0.5atm, which from simple trigonometry implies a deflection amplitude of 0.6mm when the distortion is assumed to be a circular arc. From equations 27 and 116, the increase in optical path length produced by the concave deflection of both windows is 5.5μm (eight fringes), which is consistent with experimental observation.

In an attempt to eliminate the window deflection, the HOE's were suspended in front of the glass using metal bars. This was partially successful and the fringe numbers were reduced by a factor of approximately three. An unacceptably high residual error was still present, however, due to the non-uniform refraction of the rays through the distorted window. Attempts at restraining the outside of the glass and the frame/gasket combination were unsuccessful, and considered dangerous due to the creation of highly localised stresses in the glass.

Figure 105: Interferogram exhibiting the parabolic error surface

The distortion of a glass plate clamped around its periphery under an even pressure loading is described mathematically by:

$$w(x,y) = \frac{4pa^4}{\pi^5 D} \sum_{m=1,3,5} \frac{1}{m^5} \left[1 - \frac{\alpha_m \tanh\alpha_m + 2}{2\cosh\alpha_m} \cosh\beta_m + \frac{\alpha_m 2yb^{-1}}{2\cosh\alpha_m} \sinh\beta_m \right] sin\gamma_m \tag{124}$$

Where,

$$\alpha_m = m\pi b/2a \qquad \beta_m = 2\alpha_m y/b \quad \gamma_m = m\pi x/a$$
$$D = Et_w^3[12(1-\nu^2)]^{-1}$$

and,

- $w(x,y)$ - window deformation
- p - pressure loading
- a, b - length and breadth of window
- E - Young's modulus (5.5×10^5Pa for glass)
- t_w - window thickness
- ν - Poisson's ratio (0.20 for glass)

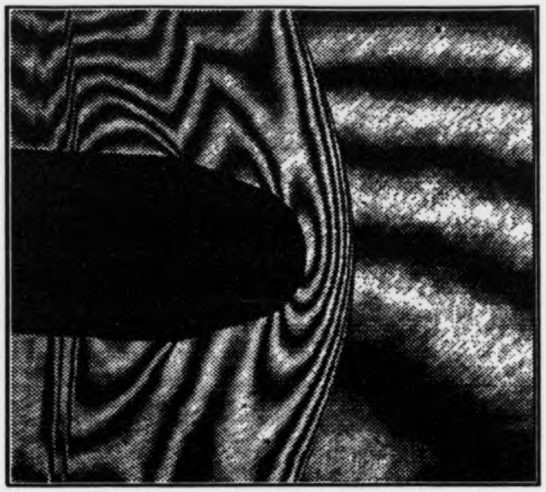


Figure 105: Interferogram exhibiting the parabolic error surface

The distortion of a glass plate clamped around its periphery under an even pressure loading is described mathematically by:

$$w(x,y) = \frac{4pa^4}{\pi^5 D} \sum_{m=1,3,5} \frac{1}{m^5} \left[1 - \frac{\alpha_m \tanh\alpha_m + 2}{2\cosh\alpha_m} \cosh\beta_m + \frac{\alpha_m 2yb^{-1}}{2\cosh\alpha_m} \sinh\beta_m \right] sin\gamma_m \tag{124}$$

Where,

$$\alpha_m = m\pi b/2a \qquad \beta_m = 2\alpha_m y/b \quad \gamma_m = m\pi x/a$$
$$D = Et_w^3[12(1-\nu^2)]^{-1}$$

and,

- $w(x,y)$ - window deformation
- p - pressure loading
- a, b - length and breadth of window
- E - Young's modulus (5.5×10^5Pa for glass)
- t_w - window thickness
- ν - Poisson's ratio (0.20 for glass)

Figure 106 shows a graph of the normalised window deflection along a horizontal line passing through the centre of a window with an aspect ratio of 2.82:1. The second line on the graph is a parabolic least squares fit to the more complicated analytical solution. The agreement is reasonable, with a maximum deviation of 7.6% and a standard deviation of 3.4%. Figure 106 illustrates that a parabolic trial function is a good approximation to equation 124. Figure 107 shows the correlation between experimental data and the parabolic curve fit. It is arguable whether additional accuracy could be obtained by using the correct analytical expression, because larger errors are probably introduced when separating the surface distortion from the flow data.

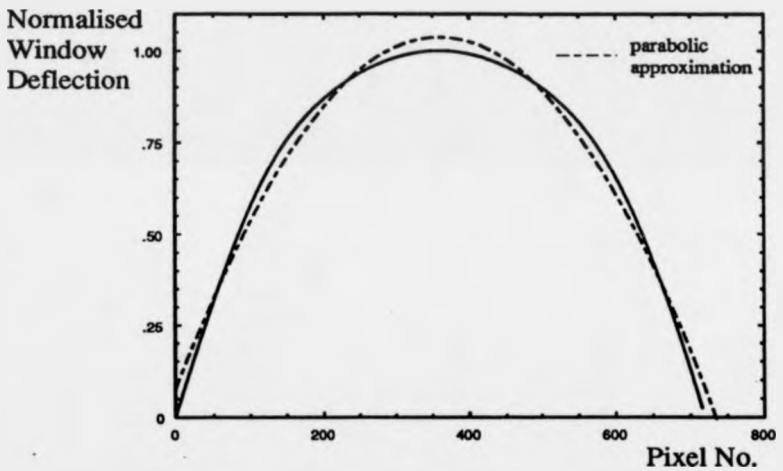


Figure 106: Comparison of analytical solution of window deflection and parabolic approximation

When the aspect ratio is reduced, so that the window becomes more square, the parabolic approximation improves and in the case of a circular window the analytical solution is purely parabolic. Figure 108 shows the analytical solution and parabolic approximation for a window with an aspect ratio of two. The agreement in this case is excellent, with a standard deviation of only 1.3%.

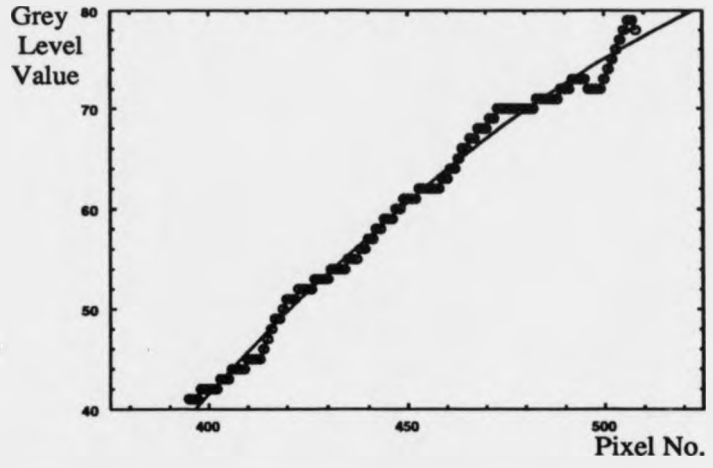


Figure 107: Parabolic curve-fit to the experimentally measured error surface

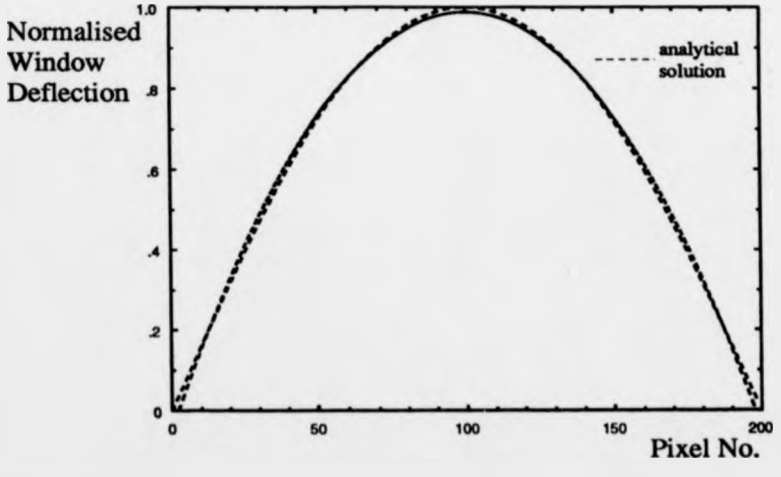


Figure 108: Comparison of analytical solution of window deflection and parabolic approximation for an aspect ratio of 2

7.4.2 Image Processing of the Phase Maps

A disadvantage of one-dimensional curve fitting is that the line scans result in horizontal streaking of the image. This may be ameliorated by fitting a two-dimensional parabolic surface to the data, with the unknown coefficients ascertained by a least squares minimisation. A more fundamental limitation of curve and surface fitting is that it can only be applied when there is a large region of uniform density in the flow, so that the error surface can be separated from the aerodynamic data. In supersonic flows the density field is uniform ahead of the shock, whereas with many transonic flows the freestream conditions may be disturbed a full model length upstream.

Holographic interferometry is most useful at transonic speeds, because optimum fringe densities are generated and shocks are weak. However, in this regime it is unlikely that the surface distortion can be accurately separated from the continuous flow data. Hence, curve fitting was abandoned in favour of a technique that calculates the surface distortion from an initial calibration of the empty tunnel. When the freestream flow was subsonic, the fringe patterns of the empty tunnel revealed unsteady features (figure 104) and therefore this technique was only applied at supersonic freestream Mach numbers. It was assumed that because the distortion amplitude is proportional to the fourth power of the linear window dimensions (equation 124), the deformation would be isomorphic even if the pressure distribution over the glass is altered by the model. Higher order distortion modes were considered unlikely, because these would require the glass to bend with a small radius of curvature.

When the calibrated error surface was subtracted from the data, it was found that the carrier fringes added to the two interferograms were of slightly different frequency. This occurred because the motorised rotation stage driving the reference beam mirror between consecutive exposures was not indexed. A computer controlled mirror, driven by stepper motors, was not considered worthwhile because vibration levels were too high to ensure reproducible results. A result of the rotational misalignment was that a linear phase difference existed between the two phase unwrapped images. Consequently, a ramp was left on the data when the phase maps of the flow around

the model and through an empty tunnel were subtracted. This problem was resolved by cross-correlating the Fourier transforms of the two interferograms, with their zero order component removed. The correlation function was sharply peaked, with the co-ordinates of the peak giving the relative shift in frequency space between the two Fourier transforms. Correction of relative frequency shifts, also provided a means for removing linear fringes produced by rigid body motion of the optical system.

In addition to correcting for tilt, the amplitude of the surface distortion had to be calculated. A least squares optimisation of the surface fit was unsuitable because the positive flow data would have biased the result. In practice, the amplitude was determined heuristically, by visually examining lines in the image after subtraction of the estimated error surface. Although partially subjective, this technique was found to be satisfactory as only one variable had to be adjusted and it was relatively simple to assess whether a parabolic error remained. If the data extraction process was to be automated then this approach would be unacceptable and the window distortion would have to be removed at source.

An interferogram of transonic flow over the large cone-cylinder model was recorded at Mach 1.20 (figure 109), and processed by Fourier transform fringe analysis and phase unwrapping. Static pressure measurements were also recorded during this wind tunnel run to complement the holographic results. Figure 110 shows the unwrapped phase map prior to removal of the window distortion and qualitatively it appears that the underlying surface is parabolic. The integrated density field ahead of the model in the differenced phase map (figure 111) is approximately uniform and flow symmetry has improved.

The unwrapped phase map (figure 111) is proportional to the line integrated density field and is known as the 90° forward projection of the flow. The pseudo-colour transition from dark blue through to yellow, denotes a reduction in the integrated density values. An Euler CFD code, based on the Jameson algorithm [221], was used to provide an inviscid computational prediction for the flow. Initially the CFD data was compared to the experimental static pressure measurements and shock tables, for

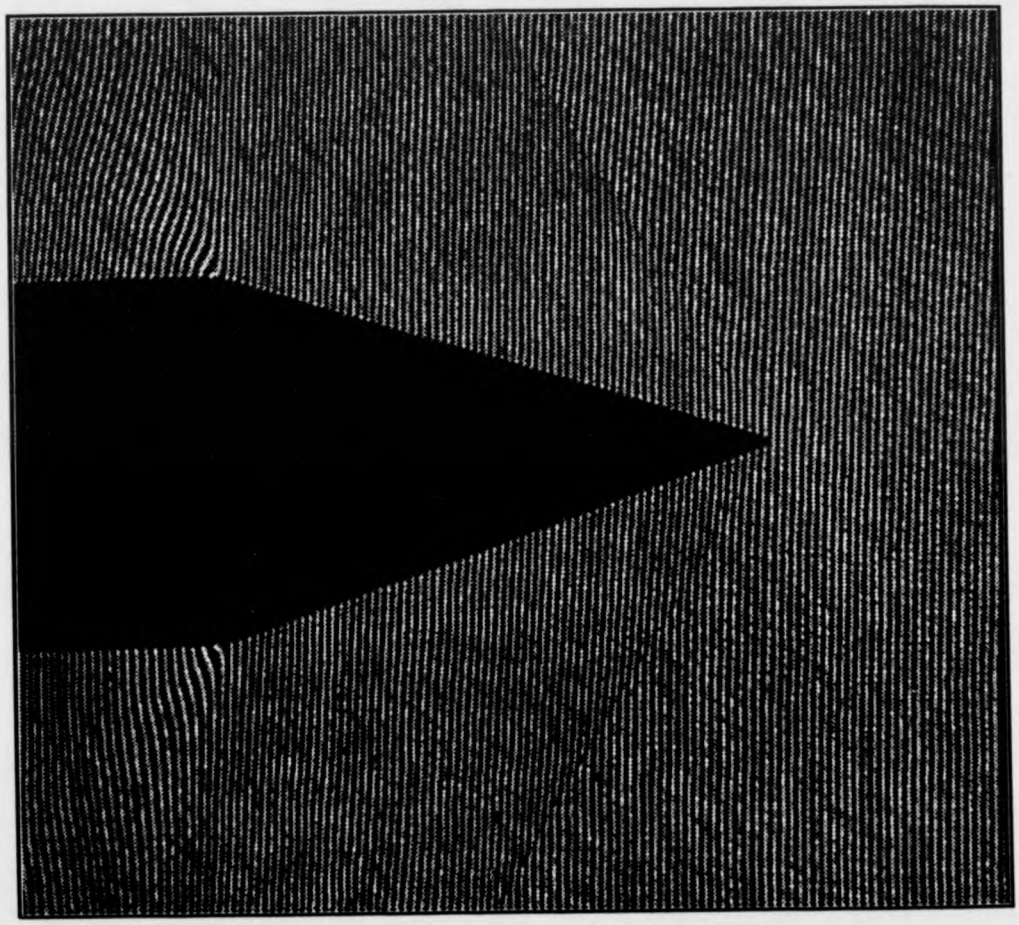


Figure 109: Finite fringe interferogram of transonic flow over a cone-cylinder model at Mach 1.20

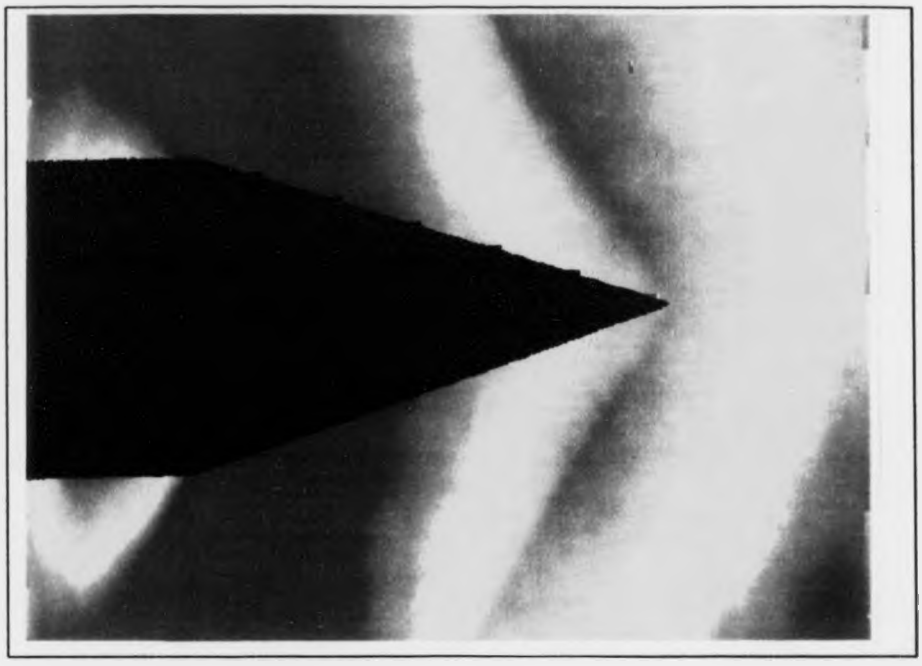


Figure 110: Unwrapped phase map with a parabolic error surface

a longitudinal plane in the flow. Comparison with the holographic measurements involves either integrating through the CFD data field, or inverting the interferometric data tomographically. These methods are discussed after the initial validation of the CFD code.

Figure 113 shows a pseudo-colour image of a longitudinal plane through the CFD generated density field. The shock angle measured from this image is $62° \pm 1°$, which is in agreement with the value given by shock tables (figure 112). The CFD density values along the surface of the cone were subsequently converted into static pressures by the adiabatic flow relation:

$$p/p_0 = (\rho/\rho_0)^\gamma \tag{125}$$

Where the subscript '0' denotes stagnation conditions.

The graph in figure 114 compares the computational static pressure profile along

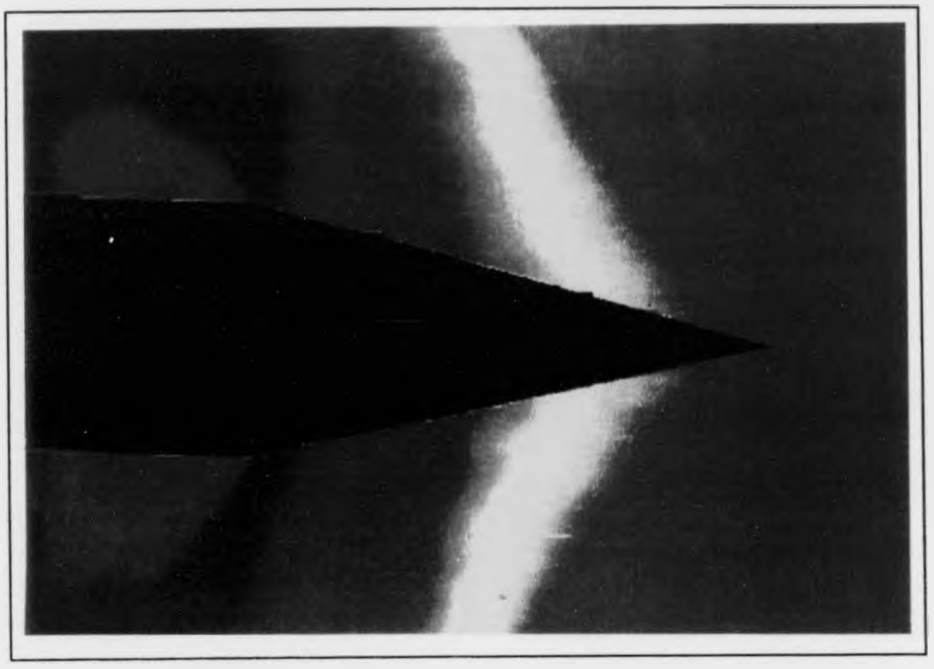


Figure 111: Unwrapped phase map with the error surface removed

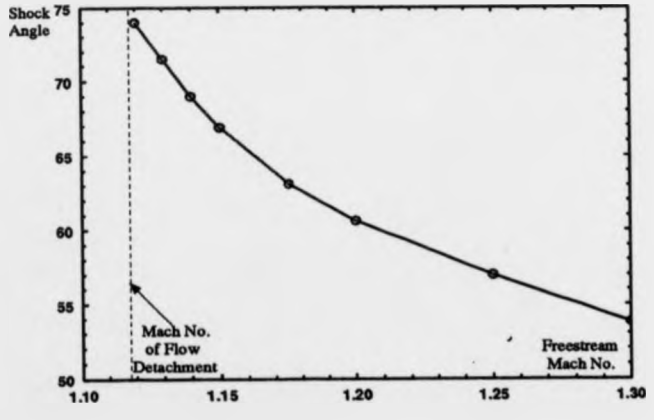


Figure 112: The variation of shock angle with Mach number for a cone of 15° semi-vertex angle

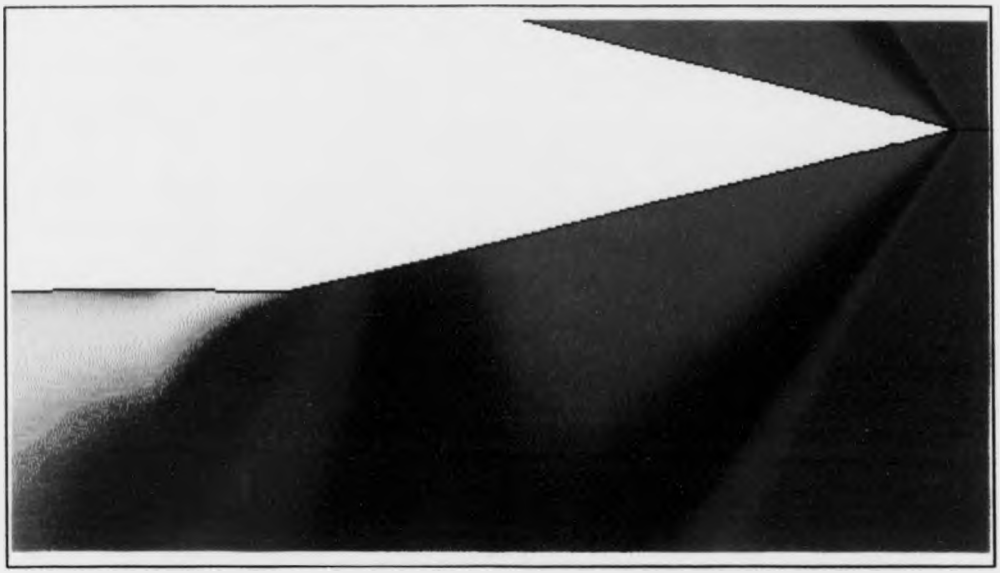


Figure 113: CFD density prediction for a longitudinal plane in the flow

the length of the model, against the experimental measurements. The longitudinal distance from the nose is plotted along the x-axis in calibres (cone diameters) and the static pressure is plotted along the y-axis. The CFD prediction and experimental measurements agree to within 3% in the region approaching the cone-cylinder junction and on the cylindrical sting, but the discrepancy exceeds 10% close to the model's nose. The 3% error can be attributed to flow unsteadiness and boundary layer effects, whereas the larger errors at the nose warrant closer investigation.

Insight into the increasing discrepancy in the measurements approaching the nose, is provided by examining the finite fringe reconstruction (figure 109). The shock angle in this picture is $46.5° \pm 1°$ above the model and $44.5° \pm 1°$ below it. The discrepancy in these two angles is probably due to errors in the model's incidence and the effect of a slight transverse velocity gradient in the tunnel (possibly indicated in figure 104). These angles disagree with both the Euler CFD prediction and cone theory, and cannot be attributed to an error in the Mach number measurement, because the wind

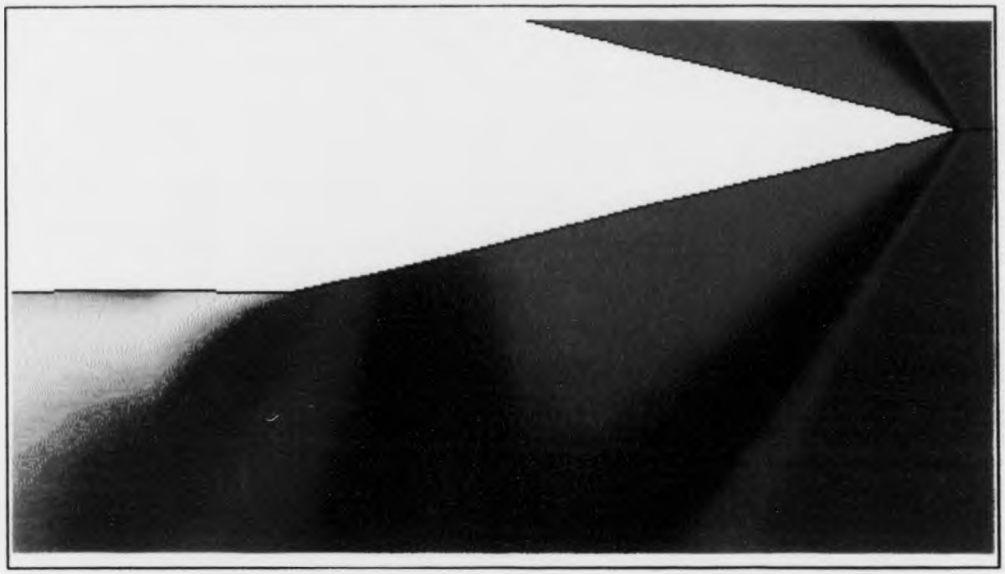


Figure 113: CFD density prediction for a longitudinal plane in the flow

the length of the model, against the experimental measurements. The longitudinal distance from the nose is plotted along the x-axis in calibres (cone diameters) and the static pressure is plotted along the y-axis. The CFD prediction and experimental measurements agree to within 3% in the region approaching the cone-cylinder junction and on the cylindrical sting, but the discrepancy exceeds 10% close to the model's nose. The 3% error can be attributed to flow unsteadiness and boundary layer effects, whereas the larger errors at the nose warrant closer investigation.

Insight into the increasing discrepancy in the measurements approaching the nose, is provided by examining the finite fringe reconstruction (figure 109). The shock angle in this picture is $46.5° \pm 1°$ above the model and $44.5° \pm 1°$ below it. The discrepancy in these two angles is probably due to errors in the model's incidence and the effect of a slight transverse velocity gradient in the tunnel (possibly indicated in figure 104). These angles disagree with both the Euler CFD prediction and cone theory, and cannot be attributed to an error in the Mach number measurement, because the wind

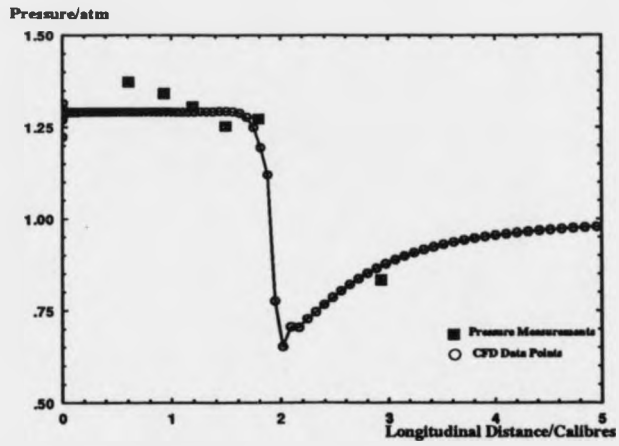


Figure 114: Comparison of the CFD predicted pressure distribution along the model and experimental static pressure measurements

tunnel had a maximum speed of Mach 1.20 (shock angle decreases with increasing Mach number). A holographic recording at Mach 1.12 (figure 115) was also found to be in error, with a shock angle of 55.5° ± 1° opposed to the theoretical value of 73.5°. The ratio of the measured and theoretical shock angles is 1.30±0.05 for both of these Mach numbers, which indicates that the error is systematic.

The region of closely packed fringes, approximately 60% of the way along the conical section of the model (figure 115), provides some insight into the discrepancy of the experimental and theoretical shock angles. This large density gradient occurs where the shock wave impinges on the test section windows. The width of the tunnel is 254mm and therefore the shock angle towards the observer is 74.5° ± 2°, which is in agreement with the theoretical shock angle (73.5°). The disparity between the vertical and out-of-plane shock angles, is shown more clearly when a lower part of the field-of-view is examined (figure 116). In this interferogram, the shock envelope on the window should not be confused with the shock reflection from the tunnel floor, which is seen to cross the picture diagonally.

Figure 116 indicates that the conical shock front is actually flattened in the vertical direction, with an elliptical rather then circular cross-section. True conical flow exists

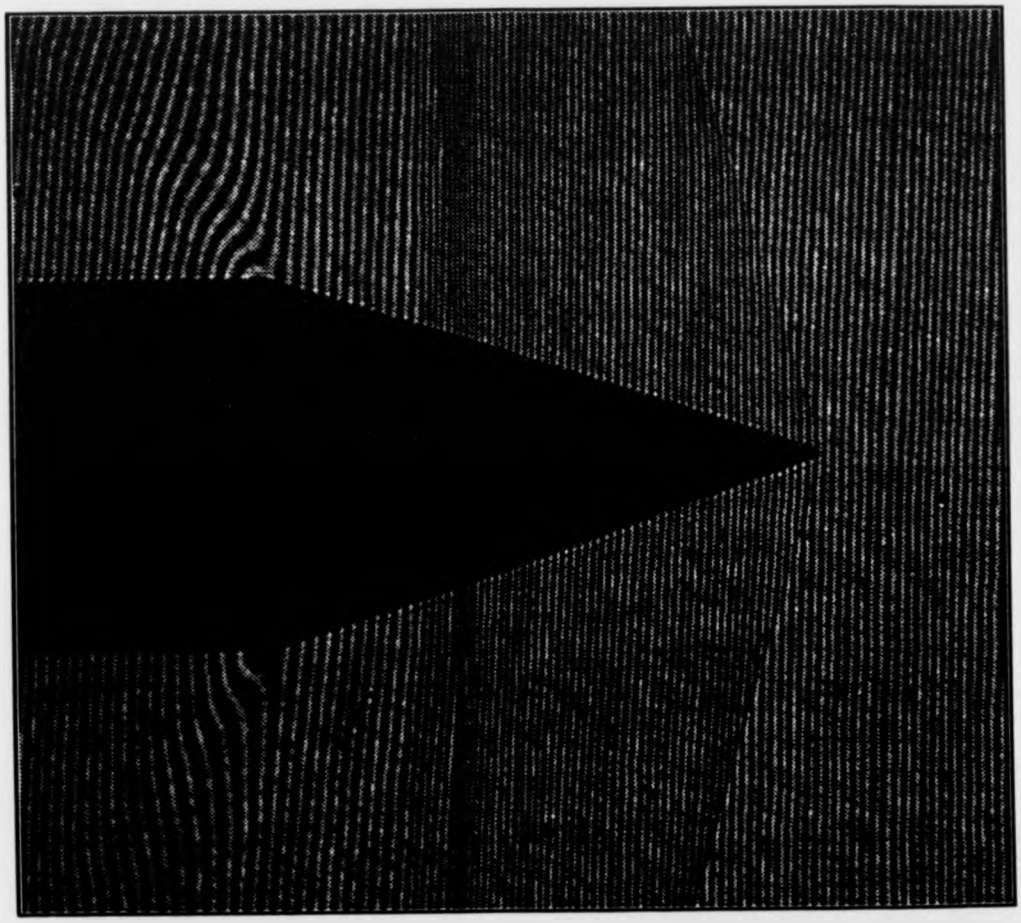


Figure 115: Finite fringe interferogram of transonic flow over a cone-cylinder model at Mach 1.12

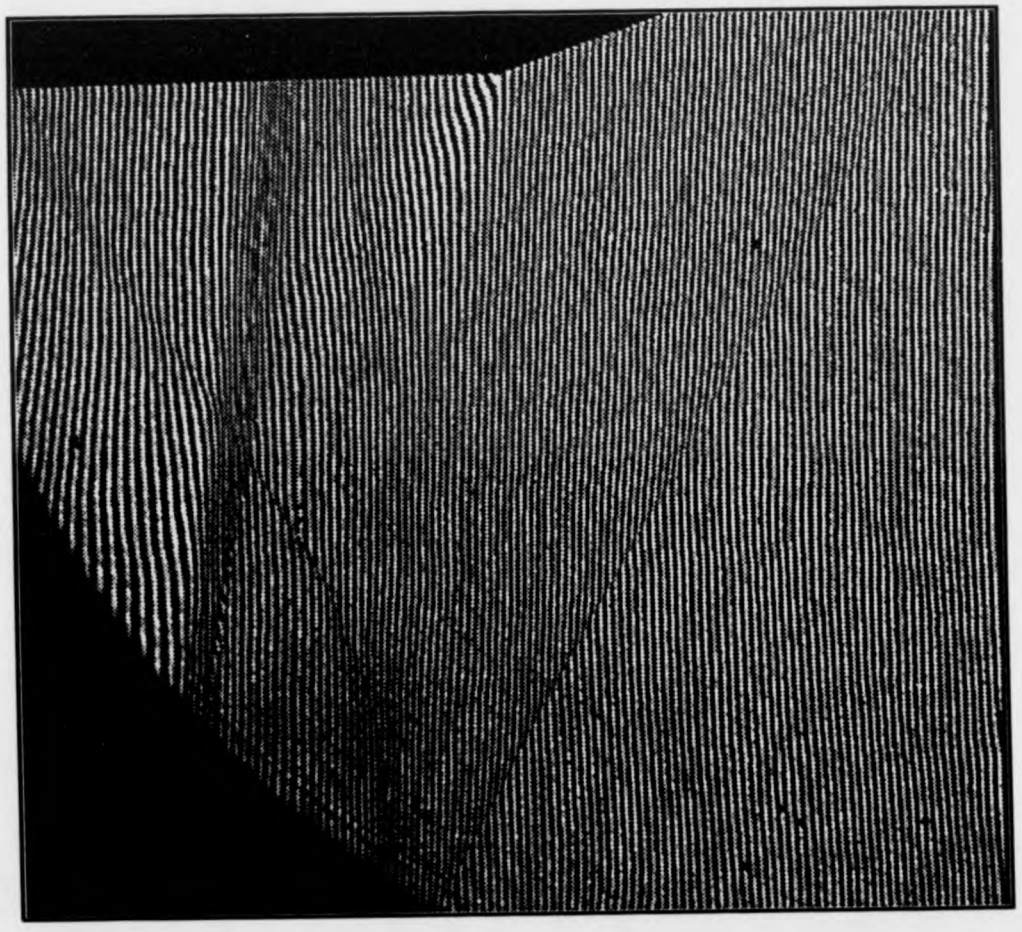


Figure 116: Interferogram showing shock position on the wind tunnel window

when all flow properties are constant along lines drawn from the vertex of the cone. This is clearly not observed experimentally, because the static pressure varies along the conical section of the model (figure 114). In contrast, the surface static pressure data derived from the CFD code is uniform along the length of the model, except in the vicinity of the cone-cylinder junction. This discrepancy between the experimental pressure measurements and theoretical prediction, supports the evidence that the shock front is distorted.

When the freestream Mach number is lower than the critical value, a subsonic region is formed between the model and the shock cone (figure 117). A shock wave, defining the regions of supersonic and subsonic flow, is not observed because the air is compressed isentropically. This is a rare occurrence in transonic flows and it makes the conical test case particularly suitable for analysis by holographic interferometry. At transonic speeds, the cylindrical section of the model perturbs the whole density field, because information can propagate upstream from the cone-cylinder shoulder. In practice departure from conical flow is observed as a bending of the shock wave, as shown in a finite fringe reconstruction at Mach 1.08 (figure 118). If the Mach number is decreased further, or the cone angle is increased, the shock wave becomes detached as with flow around a bluff body. Shock detachment is illustrated by the finite fringe reconstruction recorded of the blunt nosed cone-cylinder model at Mach 1.22 (figure 119).

Although departure from conical flow suggests why the experimental shock cone is distorted, insight into the magnitude of the effect may be gained by considering flow confinement. The working section had a width of 254mm, whereas the height was reduced to 200mm by the slotted liner (figure 91). This aspect ratio (1.27) is equal, within the bounds of experimental accuracy, to the discrepancy in the horizontal and vertical shock angles at both Mach 1.20 and 1.12 (1.30 ± 0.05).

Distortion of the shock cone prevents an accurate comparison of the holographic measurements with the CFD prediction. Also, a rigorous analysis requires the use of ray tracing to model the effect of refraction through the flow [80]. However, to demon-

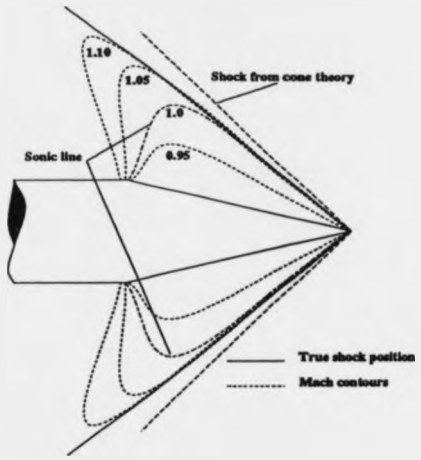


Figure 117: Mach number distribution in a transonic cone-cylinder flow

strate the principles involved, software was written to compute the line integrated density values (forward projection) from the CFD code. The CFD code employed a non-cartesian C-grid and therefore interpolation was used to determine the values in the summation. The forward projected CFD prediction (figure 120) is qualitatively similar to the experimental image, with an attached shock wave in both pictures. The shock wave is weakly defined because it is not viewed orthogonally. The largest discrepancies occur at the cone-cylinder junction, due to the inviscid assumption used in the CFD code.

After investigating the flow from an orthogonal viewing direction (0°), the process was repeated for one of the two outer optical channels (58°). A shock wave does not appear in the finite fringe reconstruction (figure 121) because it is viewed obliquely and the linear fringes are curved by the window distortion. The interferogram was processed by Fourier transform fringe analysis and phase unwrapped by the techniques described in section 4.2.1. The error surface was subsequently calculated by producing the corresponding phase map of the empty tunnel.

It was difficult to calculate the respective amplitudes of the error surface and flow induced phase data for the outer optical channel, because the projection did not contain a uniform flow region. Data separation was achieved by using the a

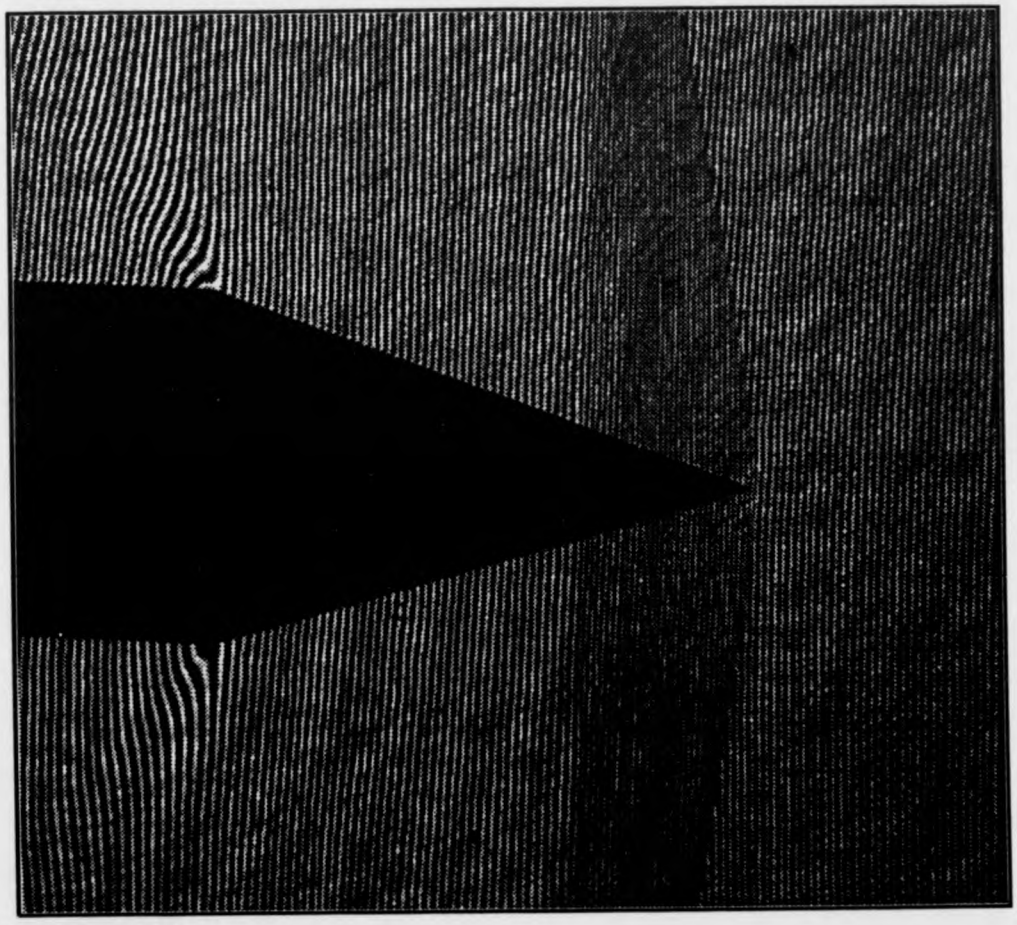


Figure 118: Shock curvature at a sub-critical freestream Mach number

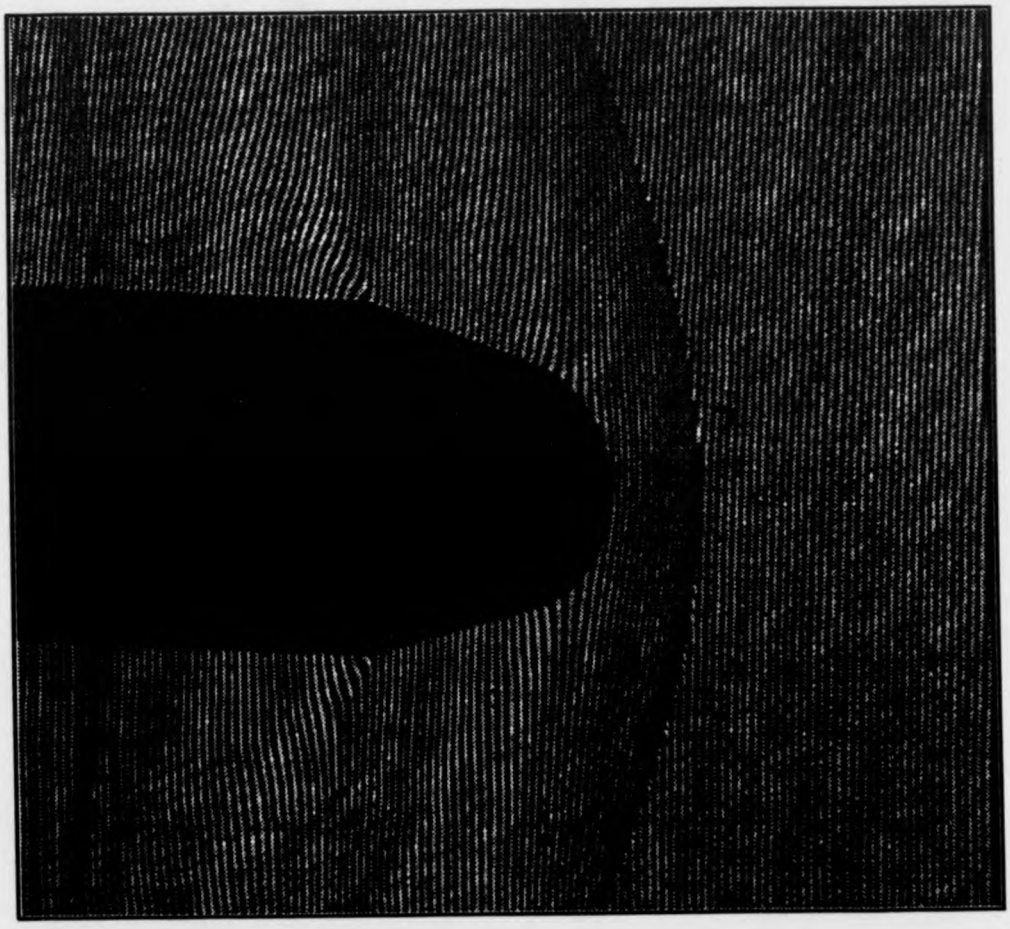


Figure 119: Shock detachment on a blunt nosed cone-cylinder model at Mach 1.22

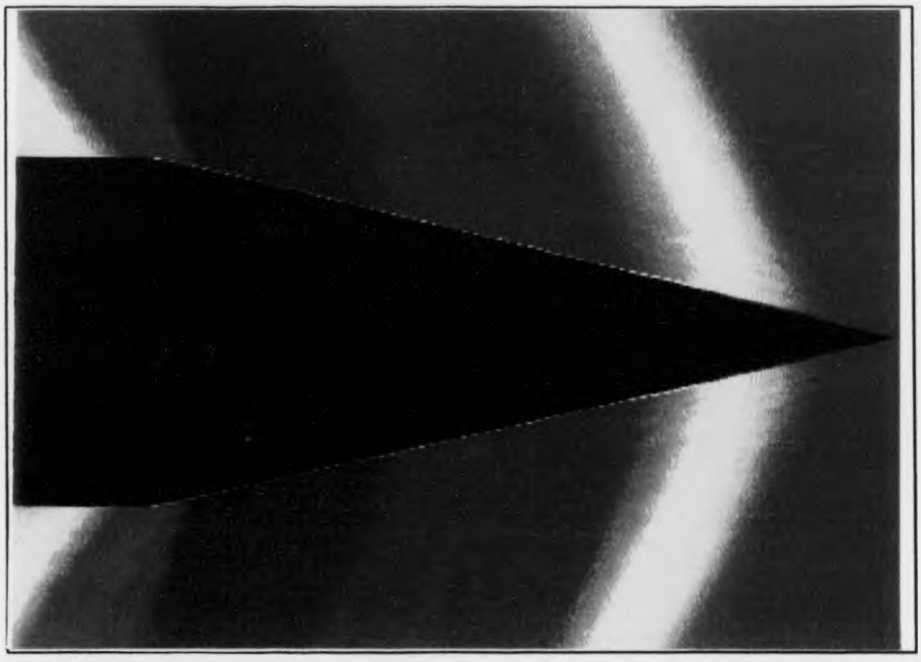


Figure 120: Euler CFD prediction

priori knowledge that the error surface is parabolic, whereas the flow data contains more rapid undulations. The difference of the normalised phase maps (figure 122) is relatively bland, because line integration through the flow tends to average features of high and low density. A view at 58° was also generated from the CFD data (figure 123). This image is qualitatively similar, with the largest integrated values at the shoulder of the model and smallest values at the nose.

Although the integrated views enable a quantitative comparison of the holographic measurements and CFD prediction, they provide less physical insight than if the actual density fields had been compared. Density data may be extracted from integrated measurements by computer aided tomography. These techniques are examined in detail in chapter 8 using projection data from a CFD code.

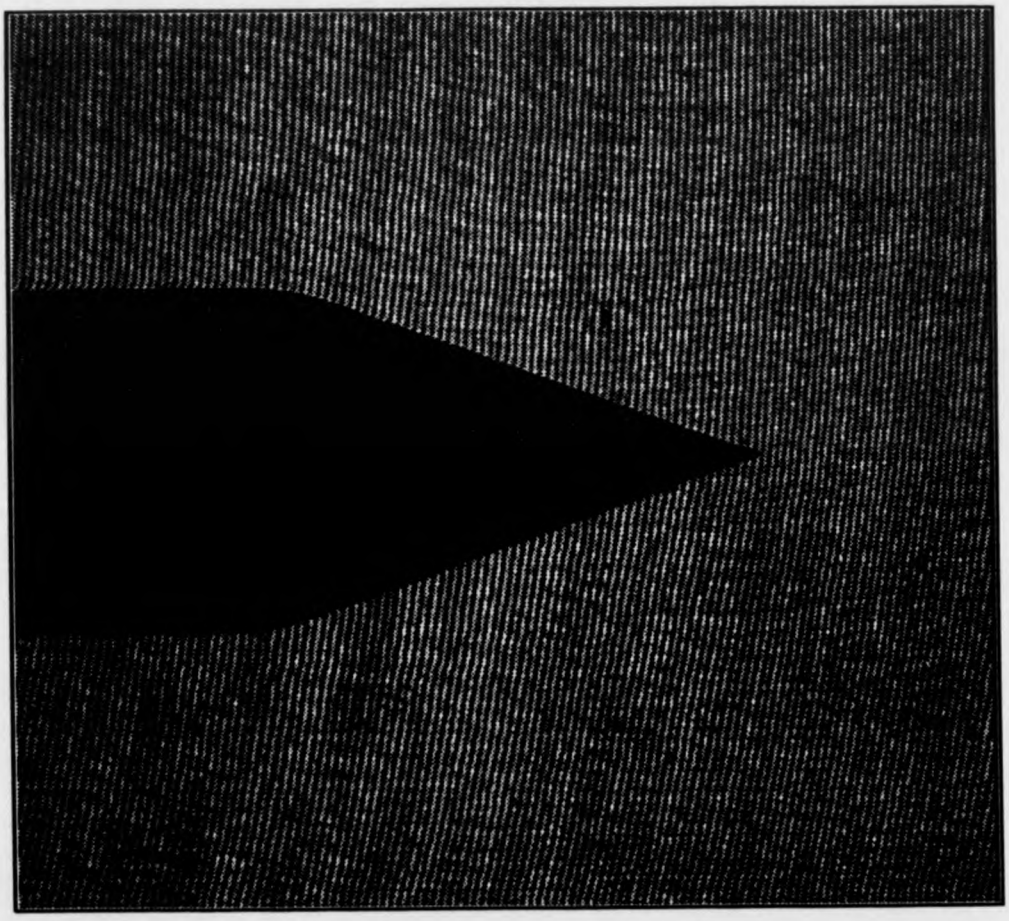


Figure 121: Finite fringe reconstruction at a viewing angle of 58°

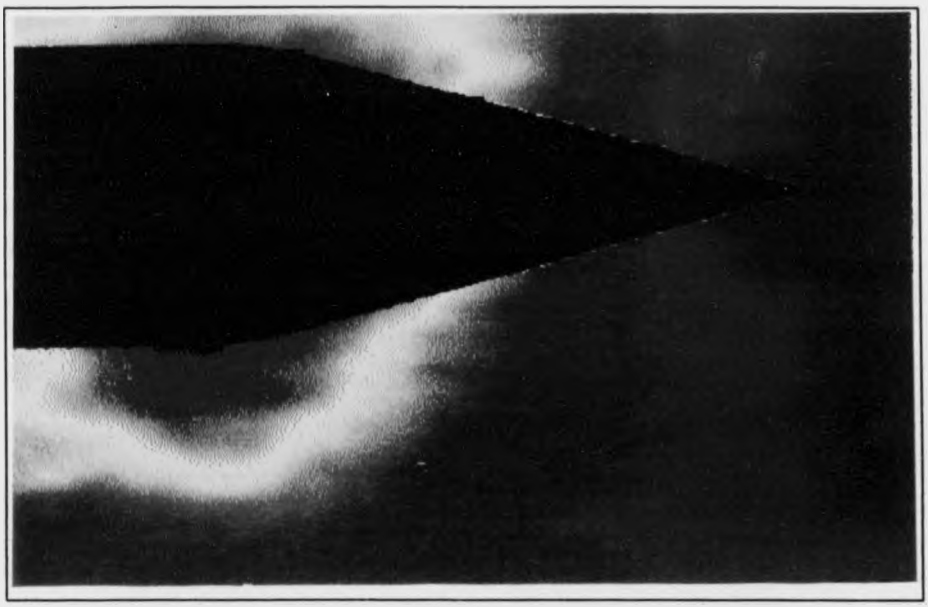


Figure 122: Unwrapped phase map with the error surface removed at a 58° viewing angle

Figure 123: Computer generated integrated density field at a viewing angle of 58°

7.5 Conclusions

Three-dimensional holographic flow analysis is only viable if an optical system can be developed that enables the recording of multi-directional interferograms. Previous attempts have either involved diffuse illumination, or replication of schlieren based specular geometries around the working section. Diffuse systems are optically inefficient and yield low contrast noisy interferograms, due to fringe localisation and laser speckle. Multiple specular configurations are also impractical, because they occupy a large area and are subsequently prone to air currents, vibration and misalignment.

This chapter has presented an alternative solution using HOE's. The system combines the wide angular field-of-view of diffuse illumination, with the fringe quality associated with specular wavefronts. Beam collimation and steering through the test section was achieved using diffractive elements, with the light expanded and imaged by more efficient refractive components. Aberrations were controlled by matching the reconstruction wavelength and geometry, to that used during recording. In addition, a symmetrical design was used to reduce aberrations further. Amplitude HOE's were used to minimise Rayleigh scatter, although is is shown that bleached elements are required for systems with greater than three optical channels.

Finite fringe interferograms were simultaneously recorded at 0° and ±58°, and the results were processed by Fourier analysis and phase unwrapping. Error fringes, produced by window deflection, were removed by subtracting a calibration run with the tunnel empty. The phase maps at orthogonal and oblique viewing angles were compared to an integrated Euler CFD prediction. Results were qualitatively similar, except that the experimental shock cone was elliptical due to flow confinement by the transonic liners.

8 Tomographic Reconstruction of Interferometric Flow Data

8.1 Introduction

Chapter 7 described the design and construction of a multi-channel holographic interferometer, which combined the benefits of both diffractive HOE's and conventional refractive optics. The system was highly compact and enabled the simultaneous recording of multiple interferometric projections over a viewing angle of 116°. Although tomographic reconstruction of the three-dimensional density field was not possible using only three views, the wind tunnel test demonstrated the feasibility of recording holographic measurements over a large viewing angle. Before more advanced designs are considered, that generate a greater number of views, it is necessary to investigate the accuracy of specific reconstruction algorithms for various beam geometries. This chapter reviews the theory of computed tomography and presents a study on the accuracy of the methods using data produced by a CFD code.

The first mathematical solution of three-dimensional image reconstruction from two-dimensional projection data was presented by Radon [222] in 1917. Computed tomography originated in the late 1950's when several independent research groups simultaneously considered reconstructing radiological projection data using the Radon transform. Cormack's pioneering work in the early development of computed tomography [223] was recognised when he was awarded a Nobel prize. Kalos *et al* subsequently proposed the development of an industrial system for measuring the density of steam inside atomic reactor heat transfer tubes [224]. At the time, however, tomographic imaging was still considered to be an invention in search of an application, and interestingly one of the earliest suggested applications in 1958 was for wind tunnel testing [225].

Computed tomography was slow to be exploited by industry due to practical difficulties associated with data acquisition and was surpassed by efforts in the medical field. Geoffrey Hounsfield an engineer from EMI, and James Ambrose a neurolo-

gist, collaborated to produce the first computer aided tomography (CAT) scanner [226]. The invention revolutionised non-invasive tissue examination and they were subsequently awarded a joint Nobel prize. Initially, X-ray projection data was used, although the technique was later extended to the imaging of γ emitting radionuclides. Images of similar high quality are now being produced of soft tissue by nuclear magnetic resonance (NMR) [227]. NMR has the advantage that ionising radiation is not involved and therefore the possibility of biological damage is reduced. Similar safety advantages are offered by tomographic imaging of ultrasonic data, although refraction of the acoustic field limits its use to the study of entirely soft tissues.

The number of potential industrial applications for computed tomography exceed those in the medical field. Principal areas which have been identified as likely benefactors of tomographic imaging are: radioastronomy [228], geophysical exploration [229], remote sensing [230], non-destructive testing [231, 232] electron microscopy [233], plasma diagnostics [234] and optical interferometry [235]. The considerable benefits can only be exploited, however, if the logistics of recording data over an entire 180° are resolved. Alternatively, an algorithm must be developed which can reconstruct an object from limited angle projection data. The feasibility of limited angle tomography is discussed by Rangayyan [236] and the problem is addressed in the computer simulations presented in this chapter.

This chapter is primarily concerned with the implementation of computed tomography for the measurement of three-dimensional density fields from optical projection data. Various optical techniques may be used to study cold and reactive flows. For example, deflectometry (study of refracted light) has been used by Faris and Byer to calculate the density field of a supersonic nitrogen jet [237]. Alternatively, optical absorption measurements may be reconstructed tomographically to study mixing processes. Bennett *et al* used this technique to calculate the distribution of iodine in a seeded jet [238] and Santoro analysed methane concentrations in a methane-air jet [239].

Absorption and refraction results in amplitude modulation of the electromagnetic

field and therefore these techniques are inaccurate when the optical system causes extraneous scatter or absorption. In contrast, holographic interferometry records phase information, which is related to the shape of a wavefronts after it has been distorted by a flow field. Phase measurements may be extracted with high precision from holographic recordings and the data is conducive to tomographic analysis.

Holographic interferometry produces fringe patterns that map the perturbation of optical wavefronts after they have propagated through a field of varying refractive index. Scattering processes are not involved in the formation of interferograms and therefore depth information of the phase retarding features is lost in the line integration through the flow. Three-dimensional information can only be retrieved if the flow is interrogated at a number of angles, so that tomographic reconstruction methods may be applied.

The reconstruction of flow information from interferometric projections is analogous to the recombination of X-ray shadowgrams. Important differences between the techniques are as follows:

- Most interferometric optical projections contain missing data due to opaque obstructions in the beam path, whereas X-ray shadowgrams may be recorded with complete views.
- Beam access is usually superior with X-ray configurations.
- Optical wavelengths are typically 10^4 times longer than X-rays and therefore wave properties, such as refraction, must be included in a rigorous analysis.
- Holographic interferometry records phase perturbations of a wavefront, whereas X-ray analysis measures amplitude modulation.
- The amplitude of a wavefront is attenuated exponentially as it propagates through a uniform absorption field, whereas the phase is retarded linearly.

A compressible air flow generates a three-dimensional refractive index distribution $n(x, y, z)$, which is a perturbation of the quiescent refractive index n_0. The refractive

index is related to the density of the flow $\rho(x,y,z)$ by the Gladstone-Dale relation (equation 27), which is restated here for convenience:

$$n(x,y,z) = K_g \rho(x,y,z) + 1$$

In the following discussion of reconstruction methods, the perturbation of the refractive index field $n(x,y,z) - n_0$ is denoted for simplicity as $f(x,y,z)$. The line integration of $f(x,y,z)$ through a flow is directly proportional to the optical phase change measured in an interferometer and is readily accessible by fringe analysis.

If a single horizontal plane in the flow is considered (figure 124), the parallel projection data $l(p,\theta)$ measured by a holographic interferometer, is given by the forward Radon transform:

$$\begin{aligned} l(p,\theta) &= \int_{\text{line}} f(x,y)ds \\ &= \int\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - p)dxdy \end{aligned} \tag{126}$$

Where,

- p - one-dimensional projection co-ordinate
- θ - projection angle
- ds - element of line integration
- δ - Dirac delta function

The two-dimensional Dirac delta function of equation 126 defines a linear integration path through the refractive index field. The equation of the line through the plane is given by:

$$p = x\cos\theta + y\sin\theta \tag{127}$$

Recovery of $f(x,y)$ from the projection data requires an inversion of the Radon transform. This problem has been studied extensively (e.g. Herman [240] and Dean

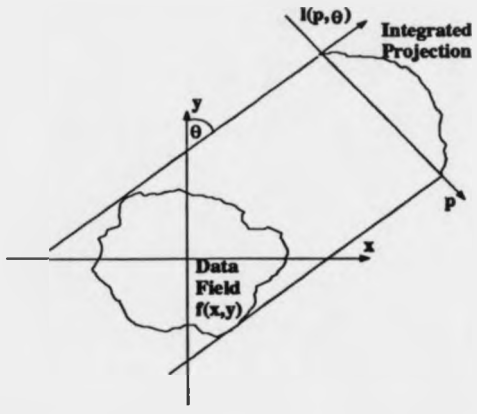


Figure 124: Pictorial representation of the Radon transform

[241]) and reviews of the various inversion strategies are given by Gordon [242]. Analytical inversion of the Radon transform is given by:

$$f(x,y) = \frac{1}{2\pi^2} \int_{-\pi/2}^{\pi/2} d\theta \int_{-\infty}^{\infty} \frac{l(p,\theta)}{p + x\sin\theta - y\cos\theta} dp \tag{128}$$

Unfortunately, direct inversion is only possible for continuous data, which are complete and self-consistent. In practice, experimental measurements are unsuitable for direct inversion and other methods must be sought.

When the flow field is axisymmetric, the projection data is described by the Abel transform:

$$l(p) = 2 \int_{p}^{R} \frac{f(r) r dr}{(r^2 - p^2)^{1/2}} \tag{129}$$

r - radial co-ordinate in the density field
R - maximum extent of the density field in the radial direction

This transform may also be inverted analytically:

$$f(r) = -\frac{1}{\pi} \int_{r}^{R} \frac{dl}{dp} (p^2 - r^2)^{-1/2} dp \tag{130}$$

Equation 130 can be used to calculate the density field of an axisymmetric flow from a single projection (e.g. Kalal [243]). The Abel inversion is used infrequently, however, because the integrand diverges at $p = r$ and noise is increased when the projection data is differentiated. Dasch has compared various strategies for inverting axisymmetric data, including Abel inversion, 'onion peeling' and filtered back projection [244]. He concludes that these algorithms have similar accuracy and are susceptible to noise, although this may be reduced by correctly spacing the data. Unfortunately most practical data fields are asymmetric and axisymmetric approximations are unsatisfactory. Hence, a more general method for solving complex three-dimensional phase objects is required.

Practical inversion strategies may be subdivided according to whether they employ series expansion or Fourier transform methods. In the development of CAT for the medical field, research has shown that the choice of reconstruction algorithm is relatively unimportant whenever there is sufficient data to over-determine the problem. Commercial scanners are highly automated, and images are routinely reconstructed with an accuracy of 0.1% when there is complete angular coverage and the data is sampled evenly. Algorithms that have been implemented include: basic and filtered back projection, direct Fourier inversion, and algebraic techniques. These methods produce images of comparable quality and differ only in speed, cost and complexity.

Tomographic reconstruction of interferometric phase information is complicated by the practical difficulties associated with collecting the required projection data. The particular problems imposed by a wind tunnel geometry may be summarised as follows:

- **Insufficient Independent Data:** In theory, n independent measurements are required to reconstruct a field containing n data points. Hence, if a flow contains 512×512 resolvable elements, and each projection consists of 512 data points, then at least 512 separate views are needed to satisfy this condition. If fewer projections are available then the problem is under-determined and the solution is less accurate. In practice, however, good reconstructions can be obtained

even when the data is very sparse and sampled over a limited angle [245]. Also, reconstruction accuracy may be improved if a *priori* knowledge is available, although an algorithm must be devised which can use this data efficiently.

- **Limited Angular Coverage:** It is better to have the projections spaced equally over the full 180°, rather than confined to a narrow fan of angles. If the angular coverage is incomplete then regions of the flow will be sparsely sampled and will therefore reconstruct less accurately than areas which are over-determined.

- **Data Occlusion (Shadowing):** In a wind tunnel there are regions in the flow where it is impossible to obtain projection data due to shadowing. The most apparent occlusion of projected rays is caused by the wind tunnel model and sting. Consequently, the tomographic reconstruction algorithm must recognise these regions either by deduction or from a *priori* knowledge. The final reconstructed field will be deficient in density information in front of and behind the model in the direction of ray propagation. This situation may be resolved if reflected rays are utilised.

- **Partial Sampling of the Data:** If the flow field does not possess a clearly defined boundary (condition of closure) then some of the projections will include data which has not been sampled in other views.

- **Refraction:** Identification of the correct voxels through which the rays pass is complicated if the light does not obey rectilinear propagation. Refraction is most pronounced when the rays encounter transverse density gradients (e.g. when shocks and boundary layers are present). Wavefront trajectories may be calculated by ray tracing, although analysis is difficult when the data is not confined to a single two-dimensional plane.

- **Errors and Noise in the Projection Data:** Holographic fringe patterns contain coherent noise, aliasing errors from the phase retrieval process, erroneous fringes from vibration and air currents, and optical aberrations. Consequently,

the projection data is slightly inconsistent and a unique solution cannot be calculated.

Numerous research groups have addressed these issues and several flow geometries have been solved. The first reported tomographic reconstruction of holographically recorded flow data was by Matulka and Collins in 1971, when they calculated the density distribution of a jet [246]. Numerous studies were subsequently made by Vest and his co-workers in the study of temperature fields [247, 248]. A particularly impressive demonstration of the technique was when Yu *et al* investigated transonic flow around helicopter rotor tips in an anechoic hover chamber [249, 96]. Other notable examples include the analysis of plasmas [250] and the study of high speed flows in a hypersonic shock tunnel [251]. In general, however, these experiments were a *proof-of-concept* and did not provide routine diagnostic information. Recently, marked improvements have occurred in image processing hardware, solid state cameras, high speed computing and tomographic reconstruction algorithms. With this technology fully matured, it should now be feasible to develop a fully automated system for operation in a production wind tunnel facility.

8.2 Tomographic Reconstruction Algorithms

8.2.1 Introduction

The Radon transform may be inverted by series decomposition, Fourier transform methods or by statistical analysis of the projected views. The choice of algorithm is highly dependent on the type of projection data available and is influenced by the following factors:

- Number of independent views
- Angular coverage
- Presence of occlusions within the views
- Level of *a priori* knowledge
- Number of data points in each view
- Sampling uniformity
- Systematic errors and noise

The principles behind the various computed tomography algorithms are outlined in this section and their suitability to reconstructing holographic projection data is discussed. Parallel beam, opposed to fan-beam, projection data is considered because holographic interferograms are usually recorded with collimated illumination.

8.2.2 The Fourier Slice Theorem and Direct Fourier Inversion

The Fourier slice theorem, also known as the central slice theorem, is fundamental to the understanding and implementation of many tomographic reconstruction algorithms. Derivation of the theorem is given in many texts (e.g. Kak [252]) and it is most easily described, without recourse to detailed mathematics, by reference to

figure 125. This diagram shows the relationship between a planar data field $f(x, y)$, possessing a two-dimensional Fourier transform $F(u, v)$, and the associated parallel projections $l(p, \theta)$ and their one-dimensional Fourier transforms $L(q, \theta)$.

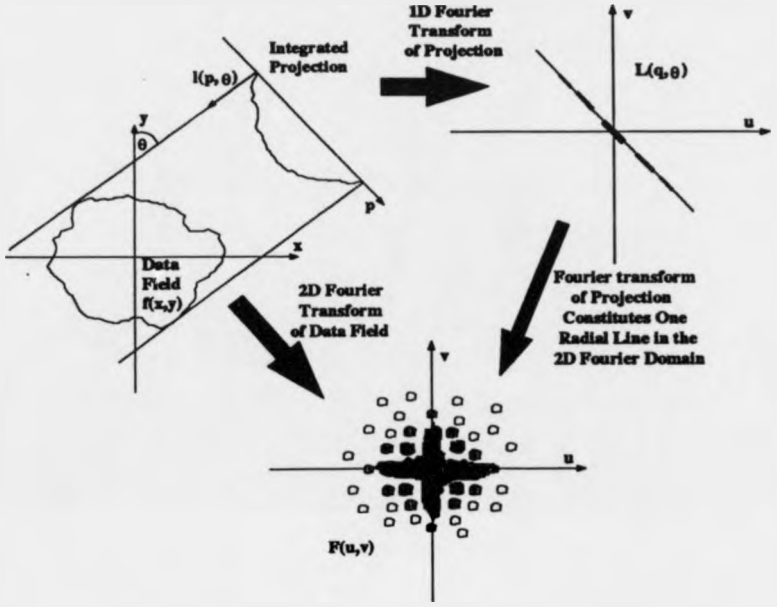


Figure 125: Illustration of the Fourier slice theorem

The central slice theorem states that the Fourier transform of a projection $L(q, \theta)$, constitutes a single radial line in the two-dimensional Fourier transform of the data field $F(u, v)$. Hence, it is possible to reconstruct the whole frequency domain of the original field by building up radial contributions from a number of views. The resulting two-dimensional frequency space may be inverse Fourier transformed to calculate the two-dimensional object space. This reconstruction technique is known as direct Fourier inversion and was first proposed by Bracewell for analysing radio astronomy data [228].

The most complicated aspect of digital implementation of direct Fourier inversion, besides calculation of the forward and reverse transforms, is the interpolation of the radially sampled spatial frequency information onto a cartesian grid. A 'nearest neighbour' approach is often used to project the values; however this method introduces

errors and reconstruction accuracy is significantly poorer than with filtered back projection. Other more sophisticated interpolation strategies are described by Stark *et al* [253]. The Fourier reconstruction method is not used in commercial scanners due to its complexity and inferior accuracy. Nevertheless, the technique offers significant advantages for NMR and radio astronomy, where the physical measurements are already a Fourier transform of the required object field. This benefit was recognised by Gmitro *et al*, who developed an analogue tomographic reconstruction system using coherent Fourier optics [254].

The Fourier slice theorem provides insight into the information contained in projection data. If the projections are widely spaced, then the only commonality between the views in the Fourier domain is the zero frequency component where the radial contributions intersect (figure 126). Hence, in this instance the projections contain largely independent measurements and will all contribute to the final reconstruction accuracy. From this description it is also seen that the high spatial frequencies, located at the periphery of the Fourier domain, are sparsely sampled and therefore reconstruct less accurately. Consequently, a large number of widely spaced views are required to increase the resolution of an image, whereas the low frequency components can be reconstructed with only a few projections.

Figure 126: Commonality of the low spatial frequencies in each of the projections

If the views are gathered over a limited angle then the frequency information is

clustered into two sectors (figure 127) and reconstruction errors occur. A computer simulation was undertaken to estimate these errors and determine the theoretical minimum viewing angle which may be used to reconstruct the density field in a wind tunnel. An Euler CFD code, modelling flow over a cone-cylinder body, was used to generate a field of density values. Figure 128 shows a pseudo-colour representation of a plane of data two cone diameters from the centreline. An off-axis plane was chosen so that a shadow of the model was not present in the image. This phantom (mathematical object) was used throughout this thesis as a representative test case for transonic flows containing a shock. The density field is normalised so that the average pixel value is unity.

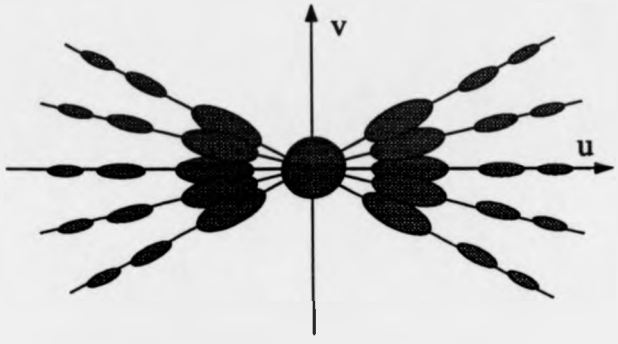


Figure 127: Loss of spatial frequency information due to the use of a limited viewing angle

The two-dimensional Fourier transform of the CFD phantom is shown in figure 129. The information lost when gathering data over a viewing angle of $\pm 30°$ and $\pm 45°$, was determined by removing the masked regions shown in figure 130 from the Fourier transform. These two modified frequency spectra were subsequently inverse transformed to produce the simulated reconstructions. The reconstructed image in figure 131 demonstrates that the data field is largely unaffected when the angular field-of-view is $\pm 45°$ (two quadrants of the Fourier domain removed). A difference map of the original and filtered images (figure 132) shows that the modifications are typically $< 1\%$. Hence, if 'perfect' projections are recorded then, in principle, a reconstruction accuracy approaching 1% could be expected with a viewing angle of $\pm 45°$.

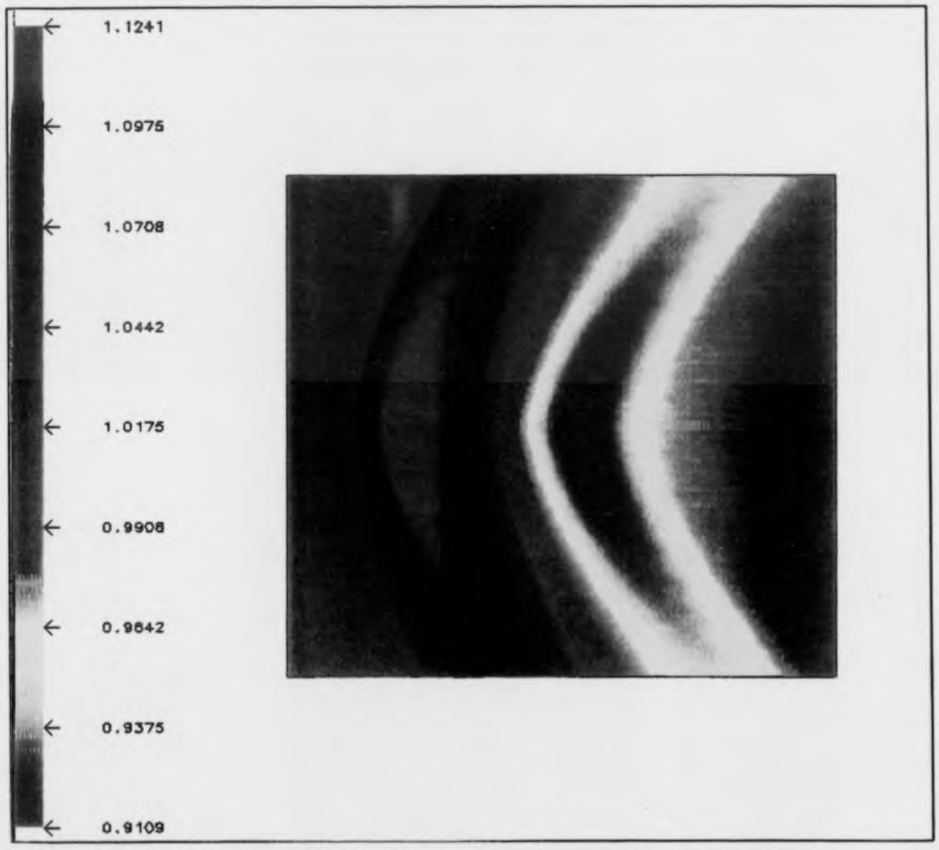


Figure 128: Plane through the CFD data field

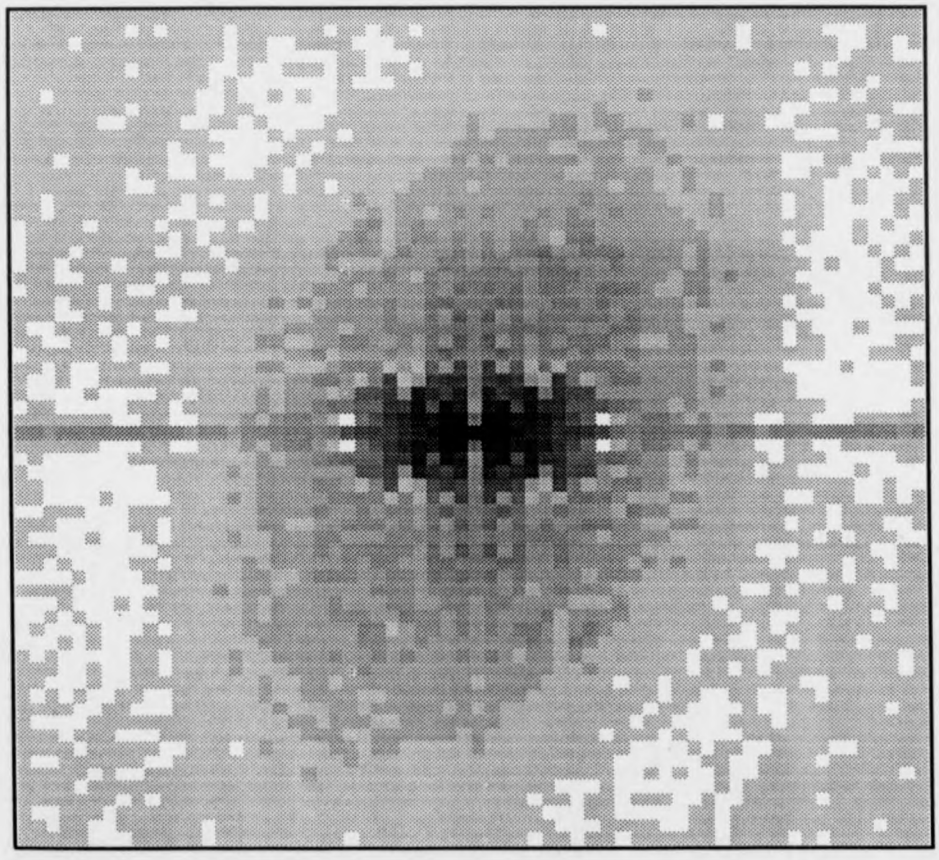


Figure 129: Fourier transform of the CFD phantom

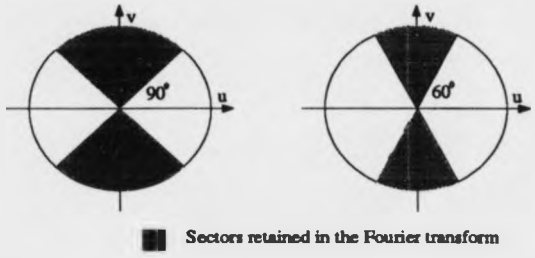


Figure 130: Removal of sectors in the Fourier domain to simulate the effect of a limited viewing angle

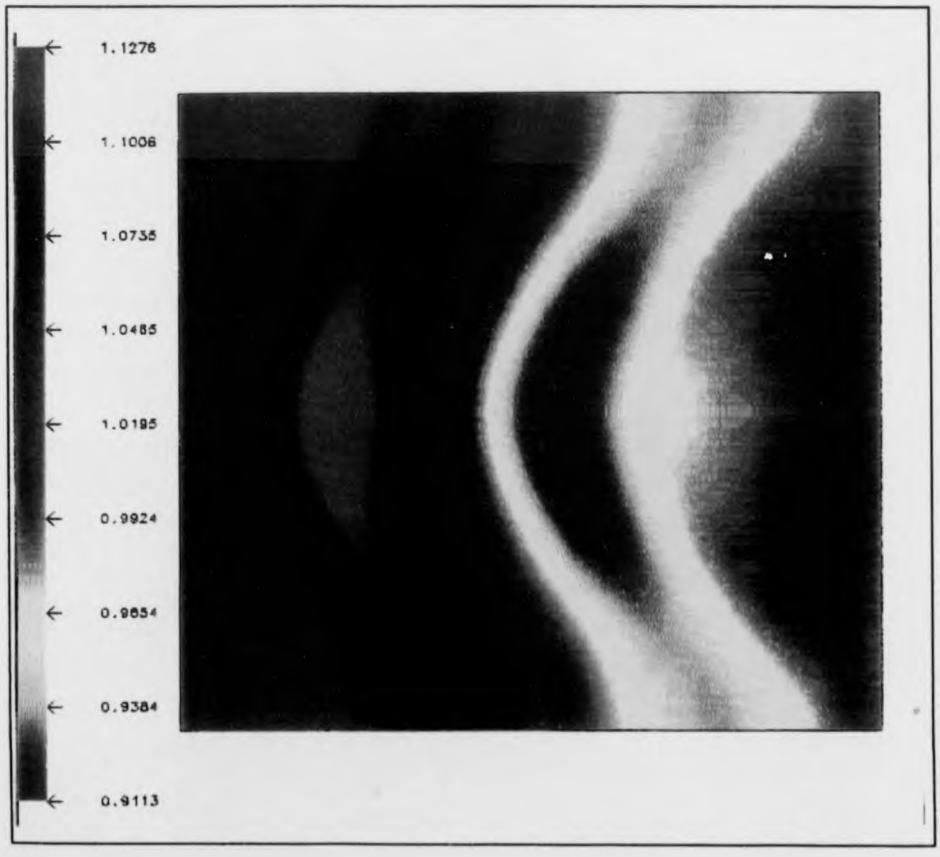


Figure 131: Reconstructed image when the viewing angle is constrained to $\pm 45°$

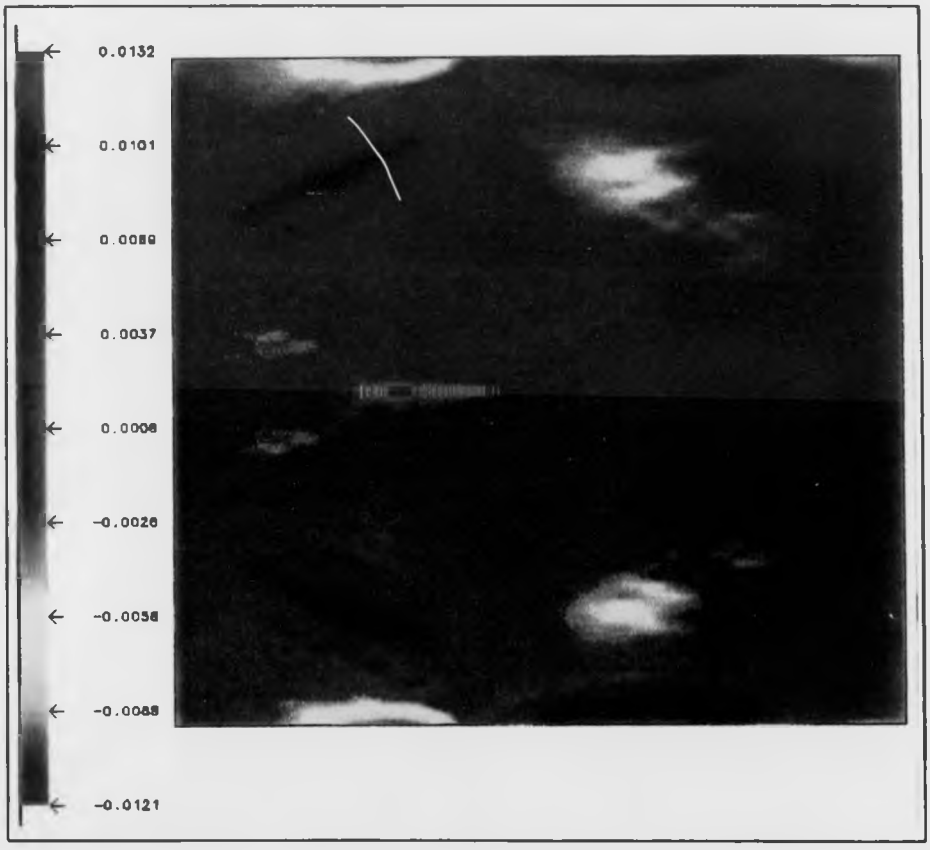


Figure 132: Reconstruction errors with the viewing angle limited to ±45°

Figure 133 shows that if the viewing angle is reduced to ±30° then, although the rms error is only 1.6%, the data is seriously degraded (4-5%) in parts of the image. The error map (figure 134), shows that the positive and negative discrepancies form alternate 'islands', which cancel when the integration paths are made along the allowed directions. The red region in the picture, that represents a zone of accelerated flow, has bifurcated due to inadequate sampling of the field. Similarly, a yellow artefact (area of reduced density values) has developed at the left of the picture, which signifies a physically unrealistic increase in flow velocity ahead of a shock wave.

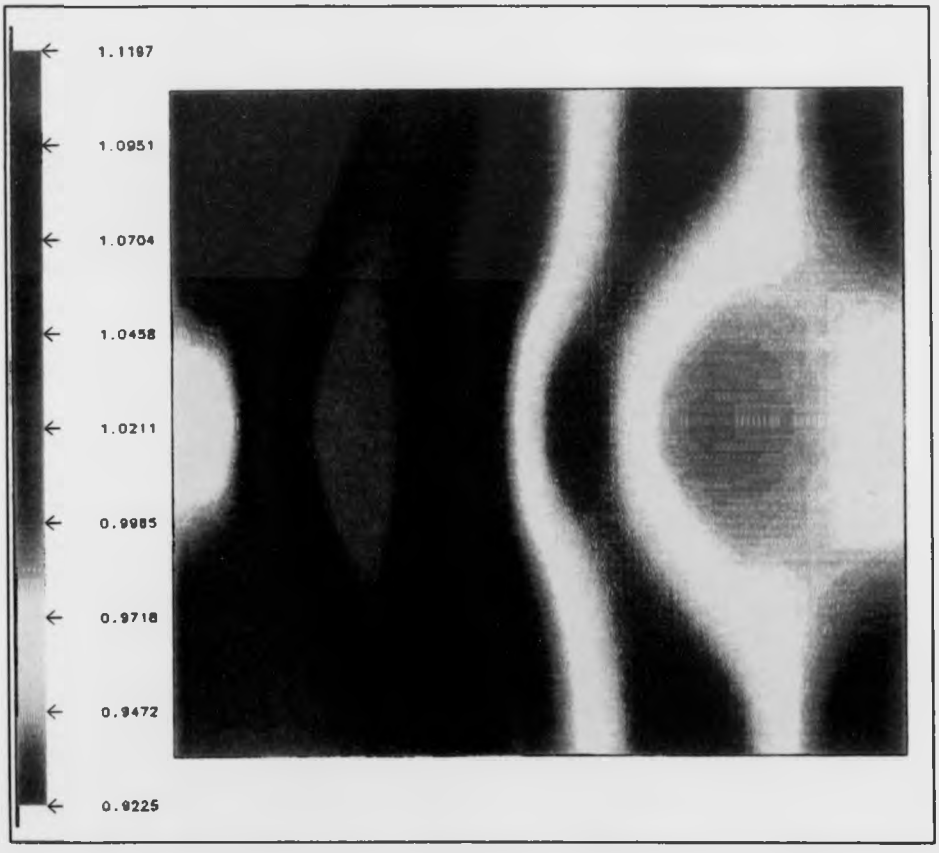


Figure 133: Reconstructed image when the viewing angle is constrained to ±30°

These examples show that care must be exercised when interpreting tomographically reconstructed flow field images. However, Hanson and Wecksung have shown that reconstruction accuracy may be improved, if knowledge about the fluid dynam-

ics is used to constrain the solution [255]. The use of this a *priori* information is equivalent to estimating the missing values in the Fourier domain.

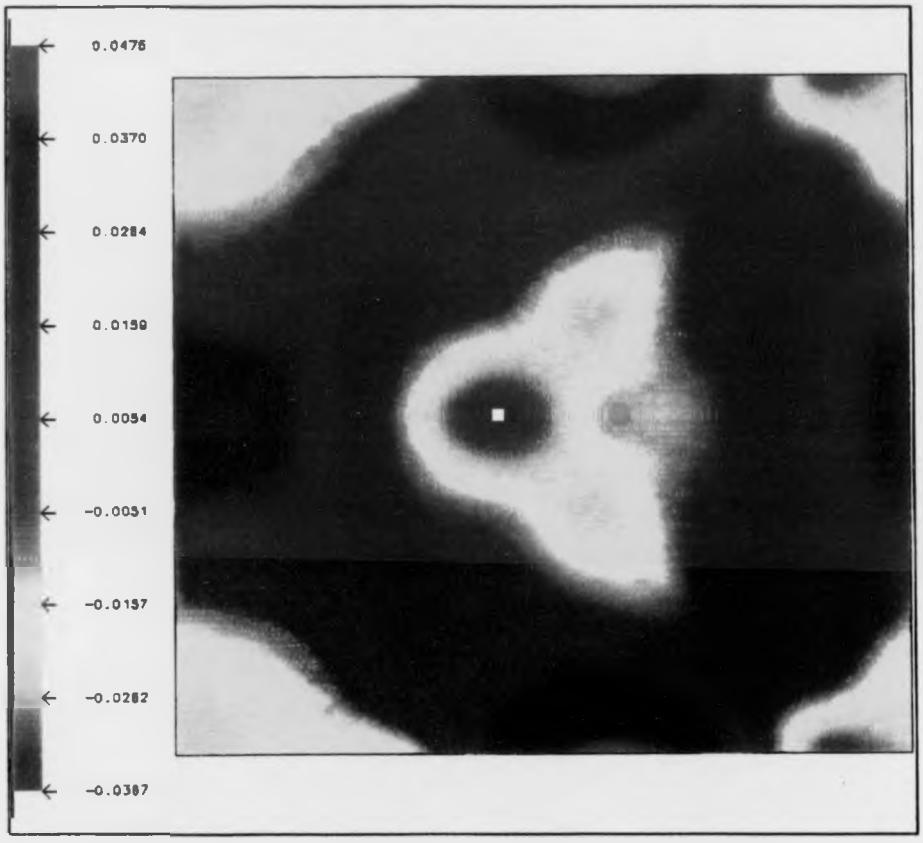


Figure 134: Reconstruction errors with the viewing angle limited to ±30°

In conclusion, a viewing angle of ±30° provides a theoretical reconstruction accuracy of better than 5% over the entire field, with an rms error of approximately 1.6%. Although the accuracy could be improved by conditioning the solution with a *priori* knowledge, it is likely that a viewing angle approaching ±45° would be required due experimental errors and discrete sampling of the data. Nevertheless, this result is encouraging, because a viewing angle of ±58° was provided by the multi-channel holographic interferometer used at City University (chapter 7).

8.2.3 Back Projection and Filtered Back Projection

The most intuitive method of reconstructing a field of data is by back projecting the integrated views so that they overlap in object space. If the field consists of binary amplitude objects, that simply block or transmit light, then the back projected views cast shadows which intersect to reconstruct the object. Mathematically this process is described by equation 131, where M_{proj} is the number of projected views:

$$f(x,y) = \frac{\pi}{M_{proj}} \sum_{all\ \theta} l_\theta(x\cos\theta + y\sin\theta) \tag{131}$$

Figure 135 illustrates the forward and back projection processes, and shows how errors arise when the field contains opaque obstructions that prevent the acquisition of data in recessed regions.

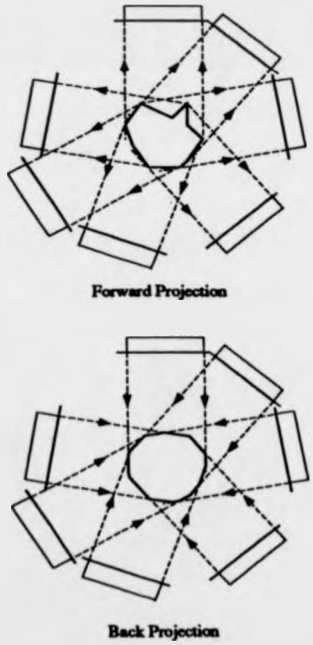


Figure 135: Illustration of forward and back projection

In practice, a finite number of views is available and therefore back projection produces streaking and spoked artefacts. Similar artefacts are generated when trans-

parent phase objects are reconstructed. These artefacts arise because the projection data lie on radii in the Fourier domain (Fourier slice theorem) and therefore the high spatial frequencies are undersampled.

Streaking artefacts may be suppressed by filtering the views prior to back projection, so that the high spatial frequencies are amplified with respect to the lower frequencies. This is accomplished by multiplying the Fourier transform of each of the parallel projections by a transfer function. Filtered back projection, also known as convolution back projection, was initially suggested in 1967 by Bracewell and Riddle [256], and subsequently re-introduced in 1971 by Ramachandran and Lakshminarayanan [257]. The first widespread use of the technique originated in 1974, after it was popularised by Shepp and Logan [258]. Filtered back projection is the most accurate reconstruction technique ($> 0.1\%$ accuracy) when the data is evenly sampled over the entire 180°. A review of the subject, with particular attention to iterative schemes, is given by Medoff *et al* [259].

Determination of the optimum transfer function for filtered back projection requires detailed mathematics. However a heuristic argument suggests that, because the sampling rate decreases linearly with increasing radius in the Fourier domain, the filter function should counteract this attenuation by increasing linearly with frequency. Spatial frequencies exceeding the Nyquist limit do not contain useful information and therefore this defines a cut-off frequency $1/2\delta$, where δ is the sampling interval. The resulting transfer function is illustrated in figure 136. The filtered projections $l'(p,\theta)$ are described by a convolution of the unmodified views with an impulse response function $h(p)$:

$$\begin{aligned} l'(p,\theta) &= l(p,\theta) \otimes h(p) \\ &= \int_{-\infty}^{\infty} l(p,\theta) h(p-\alpha) d\alpha \end{aligned} \tag{132}$$

The impulse response function is the inverse Fourier transform of the transfer function, which is given analytically as:

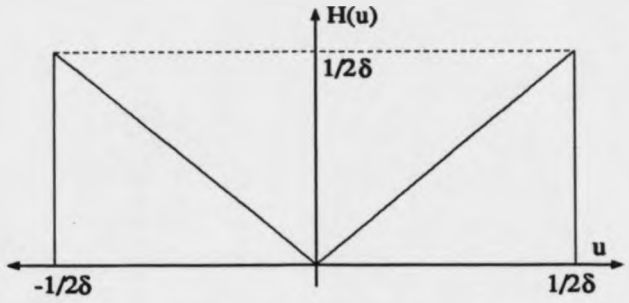


Figure 136: Transfer function used in filtered back projection

$$h(p) = \frac{1}{2\delta^2}\left[\frac{\sin(\pi p/\delta)}{(\pi p/\delta)}\right] - \frac{1}{4\delta^2}\left[\frac{\sin(\pi p/2\delta)}{(\pi p/2\delta)}\right]^2 \tag{133}$$

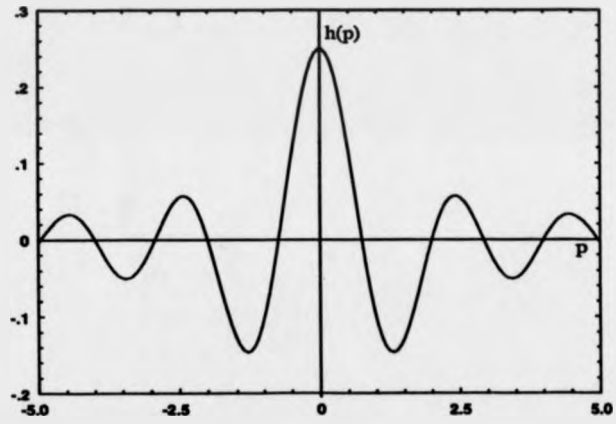


Figure 137: Impulse response function used in filtered back projection

Equation 133 is oscillatory and contains negative values (figure 137). When $h(p)$ is convolved with the projections, negative values are generated at the edges of the data. Hence, when these filtered views are back projected, the artefacts interfere destructively and a more compact object field is reconstructed. The convolution operation may be implemented directly on a digital computer, with the continuous impulse response (equation 133) replaced by a discrete function defined by equation 134 and shown graphically in figure 138 for $\delta = 1$.

$$h(n\delta) = \begin{cases} 1/4\delta^2 & \text{if } n = 0 \\ 0 & \text{if } n \text{ is odd} \\ -1/(n\pi\delta)^2 & \text{if } n \text{ is even} \end{cases} \tag{134}$$

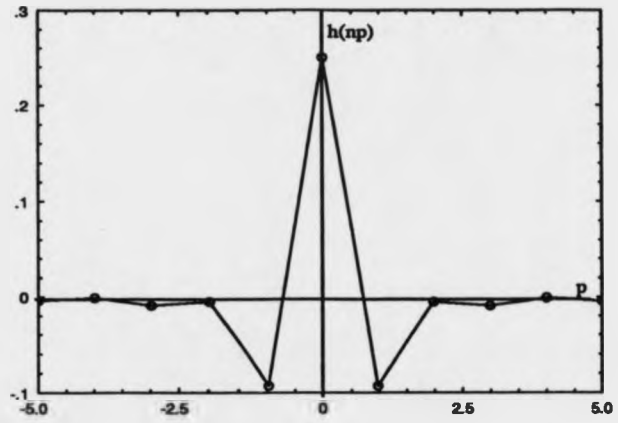


Figure 138: Discretised impulse response function

In practice, the filtered projections are simpler to compute in the Fourier domain by use of a discrete fast Fourier transform (FFT) algorithm. To ensure that the projections appear as aperiodic functions it is necessary to pad the edges of the views with zeroes. It may be shown [260] that interperiod interference is avoided if the view is extended from N to $(2N-1)$ data points. When an FFT routine is used, the number of input values must equal a power of 2 and therefore it is sometimes necessary to insert more zeroes than defined by this criterion. Padding each projection with zeroes reduces the effect of convolving the data with the Fourier transform of the window function. Noise may be further suppressed by applying a low-pass filter to eliminate spurious high frequencies. A review of the aliasing errors introduced by filtered back projection is given by Kak and Crawford [261]. Once the views have been filtered, they are inverse Fourier transformed and back projected by applying equation 131.

In conclusion, the filtered back projection algorithm has been widely accepted as the best method for reconstructing CT projections, because of its simplicity and quality of results ($> 0.1\%$ accuracy). Direct Fourier inversion, by use of the central slice theorem, is less accurate and produces mottled images; however it is popular

for applications where the measurand is already transformed (e.g. radio astronomy and electron microscopy). The greatest drawback of filtered back projection is that it performs badly when the angular field-of-view is limited, the views are spaced unequally, or the projections are partially obscured by an opaque object. All of these scenarios arise in the recording of interferometric flow data and therefore it is necessary to investigate algorithms which are more stable when the problem is ill-posed. Series decomposition and complimentary field techniques show promise in these areas, and therefore these algorithms are studied in the following sections.

8.2.4 Algebraic Reconstruction Methods

An alternative to using Fourier transform based reconstruction methods, such as direct Fourier inversion and filtered back projection, is to describe the field as a linear summation of basis functions. This technique provides a class of algorithms, which offer many advantages over Fourier transform methods for reconstructing typical wind tunnel interferometric data.

One particular implementation, known as singular value decomposition, uses basis functions to formulate a matrix equation describing the relationship between the object field and projection data. The object field is determined from this equation by computing the singular value decomposition of the projection matrix. Singular value decomposition was first applied to the analysis of interferometric data by Sweeney and Vest [262] in 1973, with the field expanded as a summation of sinc basis functions. Recent developments are reviewed by Verhoeven [263] and Decker [264].

In wind tunnel testing, the angular field-of-view is restricted and the number of projections that may be recorded is limited. If singular value decomposition is adopted, the matrix equation is ill-conditioned and an accurate solution cannot be obtained. Also, the algorithm is computationally inefficient, because the sparseness of the projection matrix is not exploited and a *priori* knowledge is not utilised. Consequently, singular value decomposition is rarely used for analysing holographic data.

Decomposition of the object field into simple basis functions is also employed in the family of algebraic reconstruction algorithms. This approach is related to that used by Kaczmarz in 1937 for solving systems of linear equations [265] and was adopted for tomographic imaging by Gordon *et al* in 1970 [266]. Algebraic reconstruction methods have subsequently gained in popularity for analysing interferometric data, because they provide a simple mechanism for incorporating a *priori* knowledge and produce good results when the problem is ill-conditioned. The principal algebraic reconstruction algorithms are:

- Algebraic reconstruction technique (ART)
- Multiplicative algebraic reconstruction technique (MART)
- Simultaneous iterative reconstruction technique (SIRT)
- Simultaneous algebraic reconstruction technique (SART)

These techniques, with the exception of SART, use the pixel as the basis function. In SART, noise is reduced by employing a basis function which extends from each pixel to its four nearest neighbours. Another popular basis function is the cubic B-spline which is defined in one-dimension by equation 135, and shown graphically in figure 139.

$$b(x) = \begin{cases} \frac{(2\Delta_x - |x|)^3 - 4(\Delta_x - |x|)^3}{4\Delta_x^3} & \text{if } |x| < \Delta_x \\ \frac{(2\Delta_x - |x|)^3}{4\Delta_x^3} & \text{if } \Delta_x \leq |x| \leq 2\Delta_x \\ 0 & \text{if } |x| > 2\Delta_x \end{cases} \tag{135}$$

The advantage of using a basis function that has similar properties to the flow (e.g. smoothness) is that fewer coefficients are required to describe the field to a given accuracy.

With algebraic reconstruction methods, a trial solution is initially proposed for the object field. A null or unity field may be used if no a *priori* knowledge of the solution

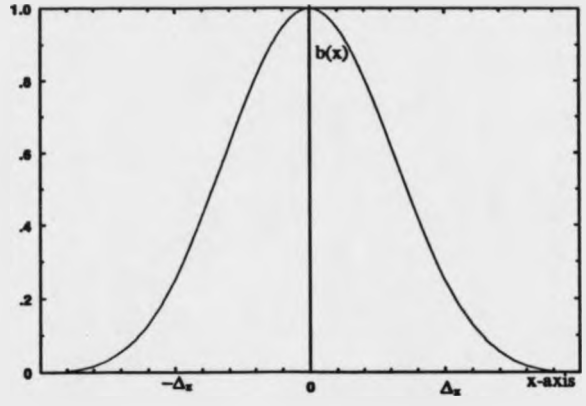


Figure 139: Cubic B-spline basis function

is available, or a more accurate 'guess' may be provided by a computer model. This is particularly convenient for the validation of CFD predictions, because the computed field can be used as the initial estimate. The trial solution is subsequently forward projected to produce a set of integrated views. The basis function is of finite extent and therefore the continuous line integration (equation 126) must be replaced by a discrete summation over all pixels intersected by each ray (denoted by $\sum_{ray}$):

$$l_i \approx \sum_{ray} w_{ij} f_j \tag{136}$$

Where,

- l_i - projection value for the ith ray
- w_{ij} - weighting factor for the ith ray passing through the jth pixel
- f_j - value of the field at the jth pixel

A discrete approximation to integration through the field is inaccurate if all pixel values intersected by the rays are summed equally. The estimate is improved by weighting the contributions with a factor w_{ij}, which is related to the geometrical intersection of the ray with each pixel. Numerous methods for assigning w_{ij} have been devised [267] and a compromise is usually made between accuracy and computational efficiency. In general, when a crude approximation is employed, the reconstructed image is speckled due to exaggeration of data discretisation.

The simplest method for allocating weighting factors is illustrated in figure 140. With this scheme, a binary weight is assigned according to the ray's proximity to the centre of the pixel. Binary weights were originally adopted to simulate the effect of using a detector array with elements of finite width and spacing. Experimentally, a null signal (zero weight) was obtained if the ray did not pass close to the centre of a pixel, because it missed its associated detector. In the example shown (figure 140), three of the pixels have a unity weighting and the rest of the array $w_{i,j}$ is set to zero. This technique is computationally efficient and requires minimal data storage. However, the process exaggerates discretisation of the image and produces poor results when the number of rays or pixels is small. Reconstruction accuracy may be improved if the number of samples in each projection is increased, so that discretisation errors may be spatially averaged.

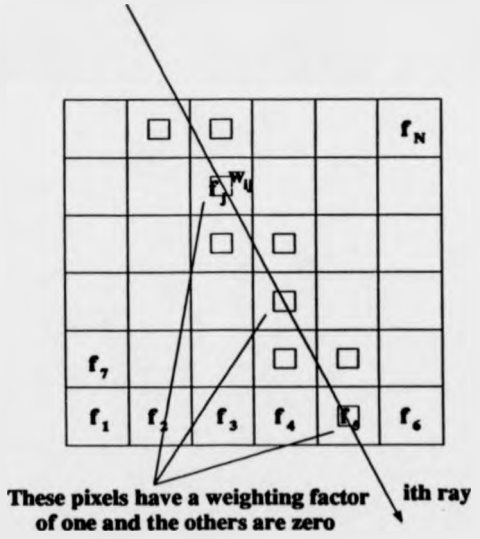


Figure 140: Allocation of binary weights $w_{i,j}$ according to the proximity of the ray to the centre of each pixel

An improved estimate for the weights, which does not exaggerate spatial discretisation of the data, is given by calculating the overlap area of a ray of finite width with each pixel (figure 141). The method may be simplified, without a loss in accuracy, by calculating the path length through the pixel instead of the overlap area (figure 142).

This integration technique was adopted for most of the computer simulations in this thesis, because of its high accuracy and relatively good computational efficiency.

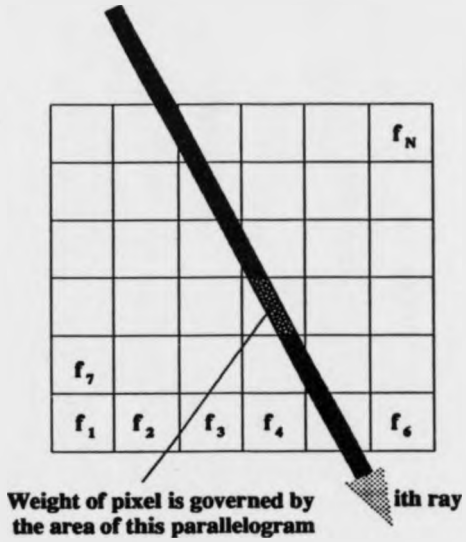


Figure 141: Allocation of weights w_{ij} according to the area of intersection of the ray with each pixel

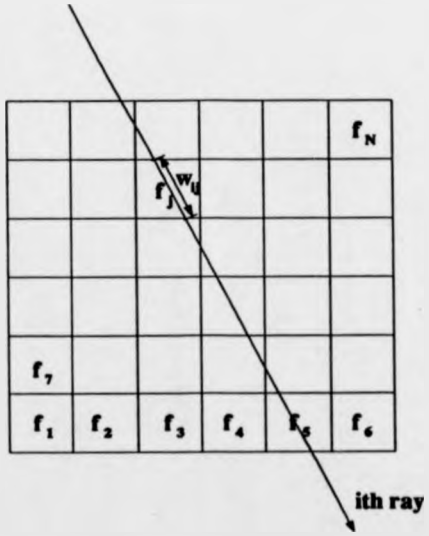


Figure 142: Allocation of weights w_{ij} according to the path length of the ray through each pixel

Once the forward projections have been calculated, the accuracy of the trial solution is determined by differencing the projections from the measured views. The discrepancy between the computed and measured views is subsequently 'streaked' back across the field for each ray. The pixel values in the trial solution are modified by the back projected residuals in proportion to the weighting factors w_{ij}. In some schemes the entire residual is divided between the intercepted cells, whereas in others it is multiplied by a relaxation factor so that only a fraction of the residual is used. The modified field may then be used to compute a new set of forward projections, which can be re-assessed against the measured views. An iterative process may be initiated to further improve the consistency between the proposed field and the observed projections.

The original ART algorithm proposed by Gordon *et al* [266] is described mathematically by:

$$f_j^{k+1} = f_j^k + \left[\frac{l_{exp} - \sum_{ray} f_j^k w_{ij}}{\sum_{ray} w_{ij}}\right] \tag{137}$$

Where,

f_j^k - trial solution at pixel j
f_j^{k+1} - improved field value at pixel j after one iteration
l_{exp} - measured experimental projection for the ith ray

With Gordon's original algorithm, binary weights were used and the calculation was terminated after only a single iteration. In words, the algorithm computes the difference between the ray sums of the estimated and measured fields, and distributes this discrepancy equally between the pixels involved in the integration.

The revised estimate f_j^{k+1} may be further improved by repeated iterations until a convergence criteria is satisfied. A number of ART algorithms have been developed on this basis, with differences only in the way the pixel values are updated and how the optimum convergence is judged. A refinement of equation 137 is given in equation 138. In this algorithm, the ray sum discrepancy is distributed across the field in accordance with the weighting factors and a user-definable relaxation factor λ has

been introduced. This method is useful when the weights have been determined more accurately by calculating the length of the ray through each pixel (figure 142).

$$f_j^{k+1} = f_j^k + \lambda \left[\frac{l_{exp} - \sum_{ray} f_j^k w_{ij}}{\sum_{ray} w_{ij}^2} \right] w_{ij} \tag{138}$$

Equation 138 is referred to as ART2 and it has been shown to produce a minimum variance solution [268]. ART2 is one of the most popular algebraic reconstruction algorithms and therefore it was chosen for investigation in this study. The relaxation factor incorporated in equation 138 reduces the amount the field is altered at each iteration. The advantage of using such a factor is that oscillations between successive iterations are damped and therefore convergence is both accelerated and stabilised. The optimum relaxation factor is governed by the amount of noise in the measurements and the conditioning of the system. If the data is noisy then it is better to use a small value ($\sim$0.1), although this means that more iterations are required to obtain convergence. When the data is noisy, it is common for the reconstruction to diverge from the true solution after several iterations. A decaying relaxation factor may then be used to delay the onset of divergence.

An alternative algorithm to ART is multiplicative ART (MART), which amends the pixel values by a multiplicative rather than additive process. The simplest MART algorithm, introduced by Gordon *et al* [266] and referred to in this thesis as basic MART, is described by equation 139.

$$f_j^{k+1} = f_j^k \left[\frac{l_{exp}}{\sum_{ray} f_j^k w_{ij}} \right] \tag{139}$$

The bracketed term in equation 139 tends to unity as convergence is reached, provided that none of the data points in the forward projections equal zero. If a singularity occurs, then the field value must be left unchanged for that particular iteration. Although computationally efficient, Verhoeven has shown that this simple algorithm exhibits poor convergence for experimental data [269].

A more sophisticated MART algorithm (equation 140), developed by Lent [268] and referred to as Lent MART, converges to the maximum entropy reconstruction if a solution exists. Physically this is desirable, because it results in the 'smoothest' image that is consistent with the projection data, thereby reducing reconstruction artefacts. However, Herman has compared this algorithm with ART2, in the context of tumour detection [270], and has demonstrated that it does not converge to the maximum entropy solution for a practical number of iterations. Also, this algorithm is shown to produce less accurate results than ART2 for his medical imagery. Furthermore, his study shows that the minimum variance and maximum entropy solutions are not particularly effective for detecting tumours that are perturbations of an otherwise smooth field. It is likely that this result would also apply in the visualisation of small flow features.

$$f_j^{k+1} = f_j^k \left[\frac{l_{exp}}{\sum_{ray} f_j^k w_{ij}} \right]^{\lambda w_{ij}} \tag{140}$$

Verhoeven has demonstrated that the relative performance of ART and MART algorithms is a function of the spatial frequencies contained in the reconstructed image [263]. His results show that ART is best suited to the reconstruction of smooth objects, whereas MART algorithms perform better when there are large gradients in the data. This indicates that MART should be used if there are shocks in the flow, whereas ART is better suited at high subsonic speeds where shock formations are weak.

Herman's and Verhoeven's results suggest that the performance of algebraic reconstruction algorithms must be assessed empirically. The Lent MART algorithm has been shown to produce a particularly stable convergence [268] and therefore it was chosen for investigation with flow field data. GH MART, a multiplicative version of ART2 (equation 141), was also studied because, although its convergence has no special mathematical significance, the algorithm has been shown to accurately reconstruct experimental data [263].

$$f_j^{k+1} = f_j^k \left\{ 1 - \lambda \frac{w_{ij}}{w_{max}} \left[1 - \frac{l_{exp}}{\sum_{ray} f_j^k w_{ij}} \right] \right\} \tag{141}$$

8.3 Investigation into the Accuracy of Algebraic Reconstruction Algorithms

8.3.1 The Reconstruction of Random Data using Basic MART with Binary Weights

Algebraic reconstruction techniques are the most popular methods for analysing three-dimensional interferometric data, because of their stability when the data is ill-conditioned. The accuracy of these algorithms was investigated using the phantom of 64×64 CFD generated values (figure 128). The conventions used for numbering the pixels and labelling the projection angle, are illustrated in figure 143. This geometry was adopted to maintain consistency with an initial subcontracted study undertaken by Bolton Institute of Higher Education (BIHE). In this preliminary investigation, the feasibility of reconstructing interferometric flow data from a limited viewing angle was explored. The rays traversing the grid in figure 143 are arranged to pass through the centres of the pixels on the main diagonal. When the projection angle changes, this condition is maintained by rotating the rays about the centres of these diagonal pixels. There are 64 pixels on the diagonal and therefore each of the projections contains 64 values, regardless of its physical width.

In the BIHE study, a field of random integers in the range [0,100] was forward projected, by the method of binary weights (figure 140), to generate synthetic projection data. The basic MART algorithm was subsequently used to 'blindly' reconstruct the original data field. These results are presented at the start of this investigation of ART techniques to demonstrate the inaccuracies of this approach.

Throughout this investigation, the accuracy of the algebraic reconstruction algorithms was assessed by calculating the rms error e_{rms} (equation 142). In addition, the errors of the largest and smallest values in the field were calculated, to determine the effect of reconstruction artefacts on the extremes of the dynamic range. A visual comparison of the reconstructed fields was provided by displaying the results as pseudo-colour images.

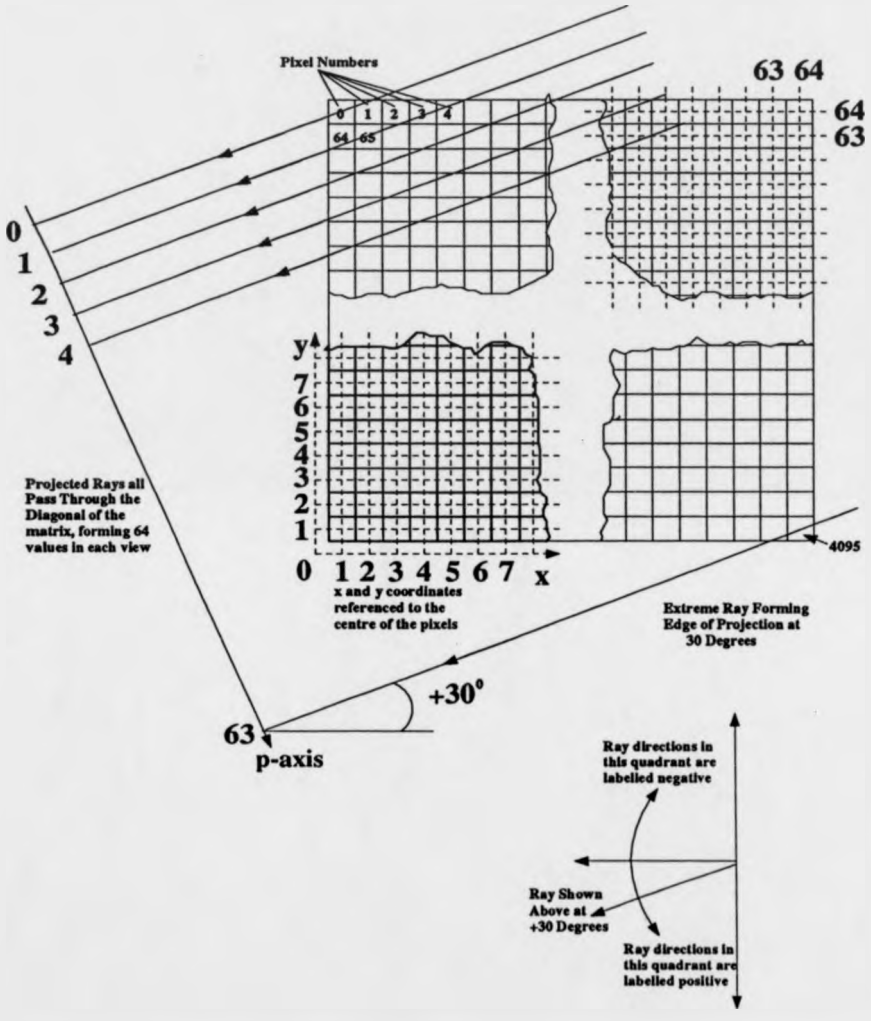


Figure 143: Conventions used for labelling the pixels and defining the projection angle

$$e_{rms} = \frac{1}{N}\sqrt{\sum_{j=1}^{N} e_j^2} \tag{142}$$

Where,

N - total number of values in the field (4096)
e_j - error in the jth reconstructed pixel value

Two different trial solutions were used as inputs to the iterative algorithms. The first of these consisted of a uniform unity field, to represent a case where no *a priori* knowledge was available. The rms error between this field and the CFD phantom is 5.3%. A null field was not adopted, because this would have terminated the iterative process for the multiplicative algorithms. The second image, representing a case where a CFD prediction has been used to provide *a priori* knowledge about the flow, was generated by modifying the original CFD phantom. The modifications consisted of both a random perturbation and a systematic component covering half of the image. Initially, signal dependent gaussian noise, with a standard deviation 0.5% of the pixel value, was added to the image. Subsequently, the systematic errors were implemented by amending pixels in the range $32 \leq x \leq 64$ by equation 143. This function distorts half of the image and simulates the failure of a CFD model to predict flow recovery behind a shock wave (figure 144). The rms error between this approximation and the original phantom is 2.5%.

$$\rho'(x,y) = \rho(x,y) + [\rho(x,y) - \rho_{min}(x,y)](x-32)/4 \tag{143}$$

Where,

$\rho'(x,y)$ - amended density field
$\rho(x,y)$ - CFD density field phantom
ρ_{min} - minimum density value in the field

In the BIHE study [245], random integers in the range [0,100] were reconstructed using the basic MART algorithm with binary weights. Single precision floating point arithmetic was used at all stages in the calculation. The weights were assigned a unity

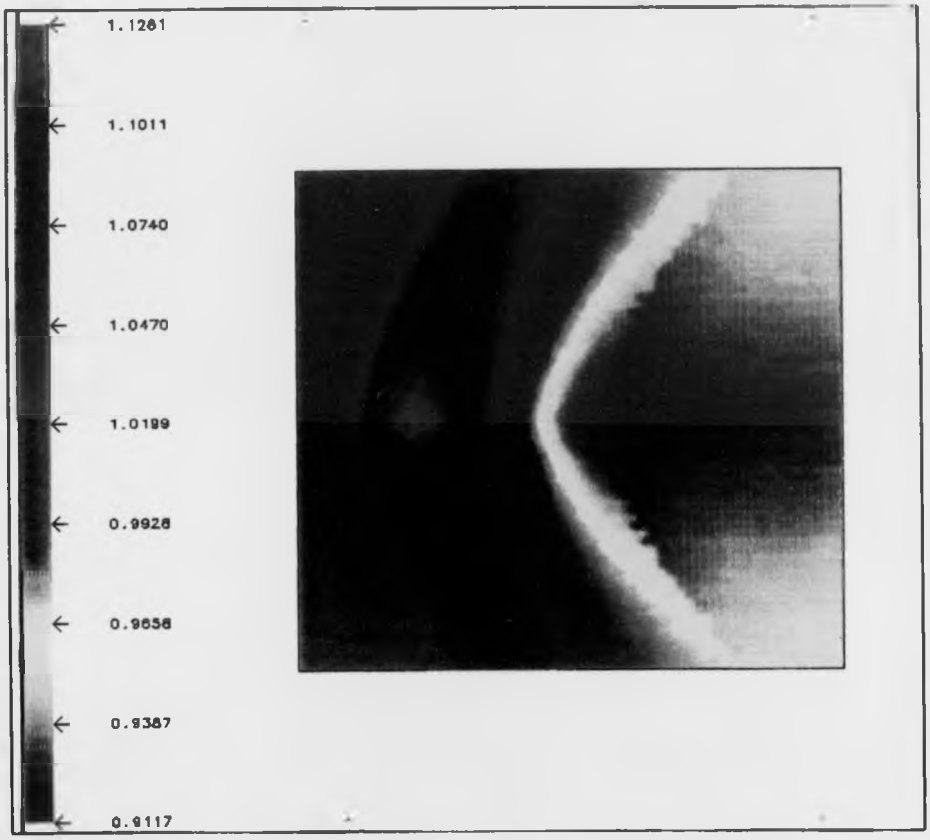


Figure 144: Modified CFD data used as a starting approximation to iterative reconstruction algorithms

value if the projected rays passed within 10% of the centre of the pixels and the viewing angle was restricted to ±60°. The algorithm was terminated after ten iterations and the reconstruction accuracy was investigated as the number of projections and rays in each view were varied.

No. of projections	No. of rays per projection	rms error (%)
64	64	±14.432
32	128	±12.541
128	64	±2.986
64	128	±2.147
64	256	±0.261

Table 11: Reconstruction of random integers using the basic MART algorithm with binary weights and a viewing angle of ±60°

The results in table 11 show that when there are only 64 rays in each projection, the errors are large due to discretisation of the line integration. The reconstruction accuracy is improved by over-determining the system with either a greater number of views, or by increasing the number of beams in each projection. For holographic applications the latter option is preferred, because it does not involve increasing the number of holographic recordings. Nevertheless, the results suggest that a prohibitively large number of views (~ 64) is required to obtain a sufficient reconstruction accuracy.

8.3.2 Comparison of ART Algorithms and the Influence of Viewing Angle and Projection Number on Reconstruction Accuracy

The software developed by BIHE was modified to accept CFD density values. Initially, the accuracy of discrete integration with binary weights, was assessed by forward projecting the CFD phantom. In figure 145 the 64 forward projections are shown as a composite image for compactness (this picture should not be confused with the reconstructed field). The first vertical line at the left of the image was generated at a viewing angle of -60° and the last one on the right for an angle of 60°. All of the views are shown to have an equal width (square picture), because 64 data points were

sampled in each projection, whereas the true physical widths are a function of the viewing angle.

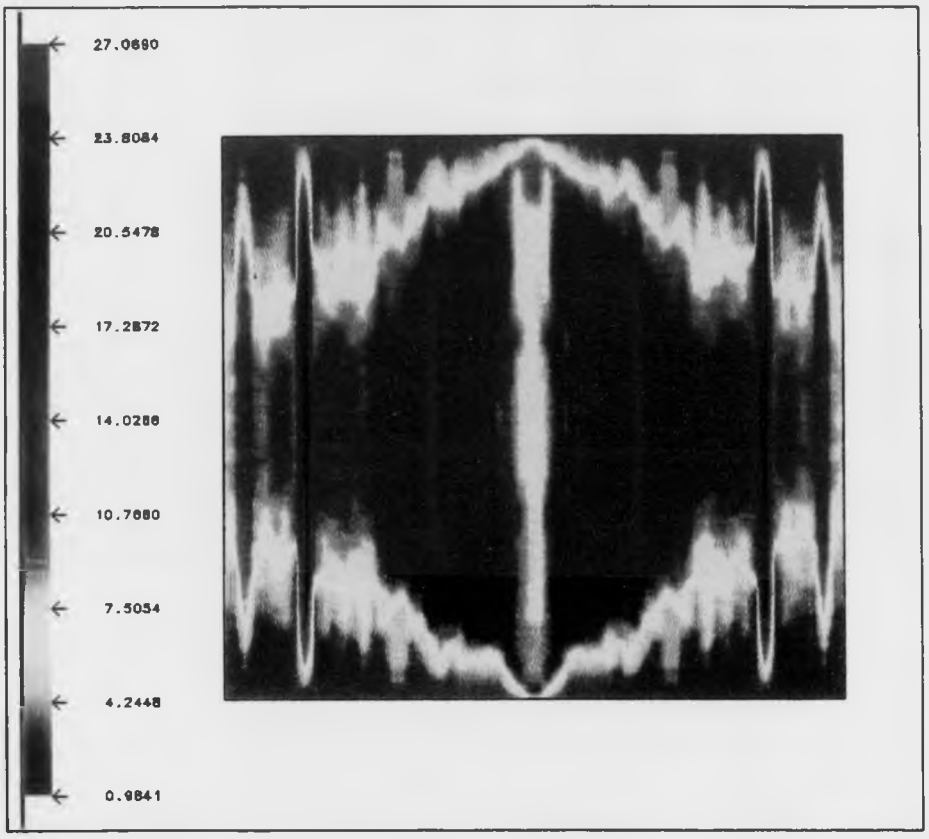


Figure 145: Composite image of the projected CFD phantom using binary weightings

The composite image exhibits streaking, which suggests that these are geometrical artefacts of the integration process, rather than true features of the projected flow field. This was confirmed when similar effects were observed after a field of unity values was forward projected. Streaking occurs when the spacing of the rays coincides with the centre of the cells, as shown by the abnormally large integrated values at viewing angles of $\pm 45°$.

The basic MART algorithm was subsequently used to reconstruct the CFD phantom from both the unity field and the improved trial solution. After ten iterations,

the rms error of the data reconstructed from the unity trial solution was 1.96%. When the more accurate trial solution was used, the rms error was reduced to 0.92%, with the maximum and minimum field values differing by 3.5% and 2.3% respectively (figure 146). The reconstructed field is mottled and contains particularly large artefacts along the diagonals. Nevertheless, this image is a significant improvement over the initial guess and the rms error is reduced by a factor of 2.76. The reconstruction could have been improved by increasing the number of beams in each projection to suppress aliasing artefacts. However in this investigation all of the algorithms were compared with 64 beams in the projections to reduce the computer run-time.

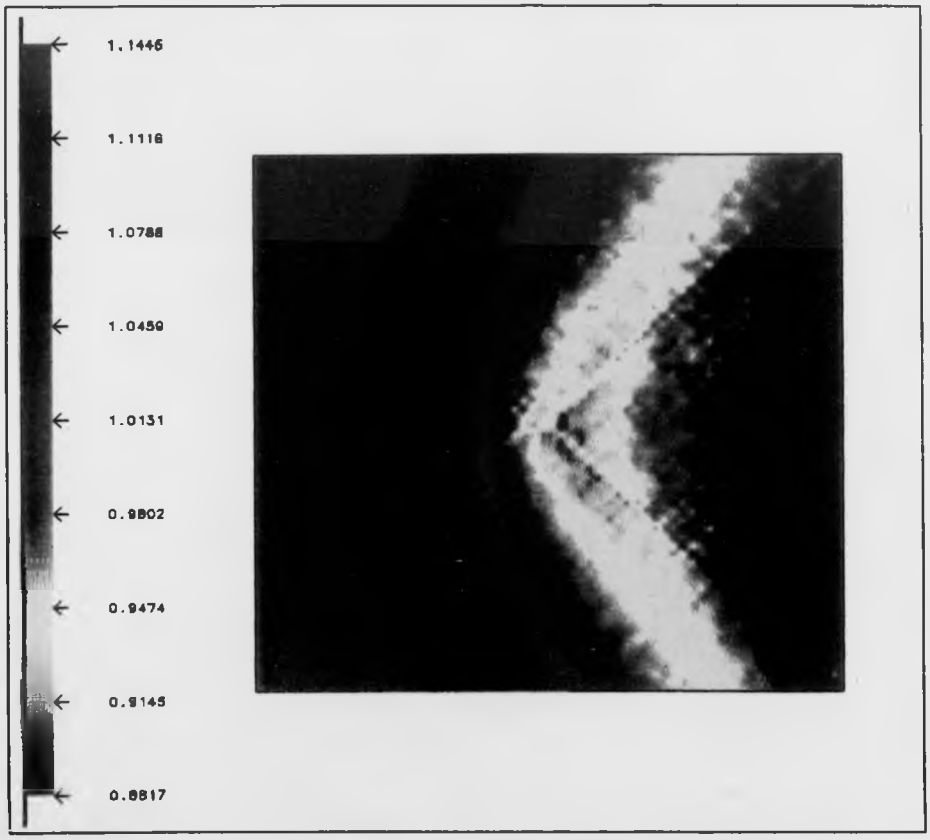


Figure 146: Reconstructed field using basic MART with binary weights and a viewing angle of $\pm 60°$

The results from BIHE and the study with CFD data, show that the use of binary

weights with basic MART is inaccurate, unless the number of sampled points exceeds the number of unknown field values. In most practical scenarios it would be difficult to record this amount of interferometric data holographically. The simulation was therefore repeated for the Lent MART algorithm, with the weights evaluated according to the length of the ray through each pixel. The results from this study will determine whether a sufficiently accurate solution can be obtained with a substantially reduced number of views. The forward projection of the trial solution (amended CFD phantom), over a viewing angle of $\pm 60°$, is illustrated in figure 147. Comparison of this result with figure 145 shows that the streaked artefacts have been eliminated by using an improved forward projection method.

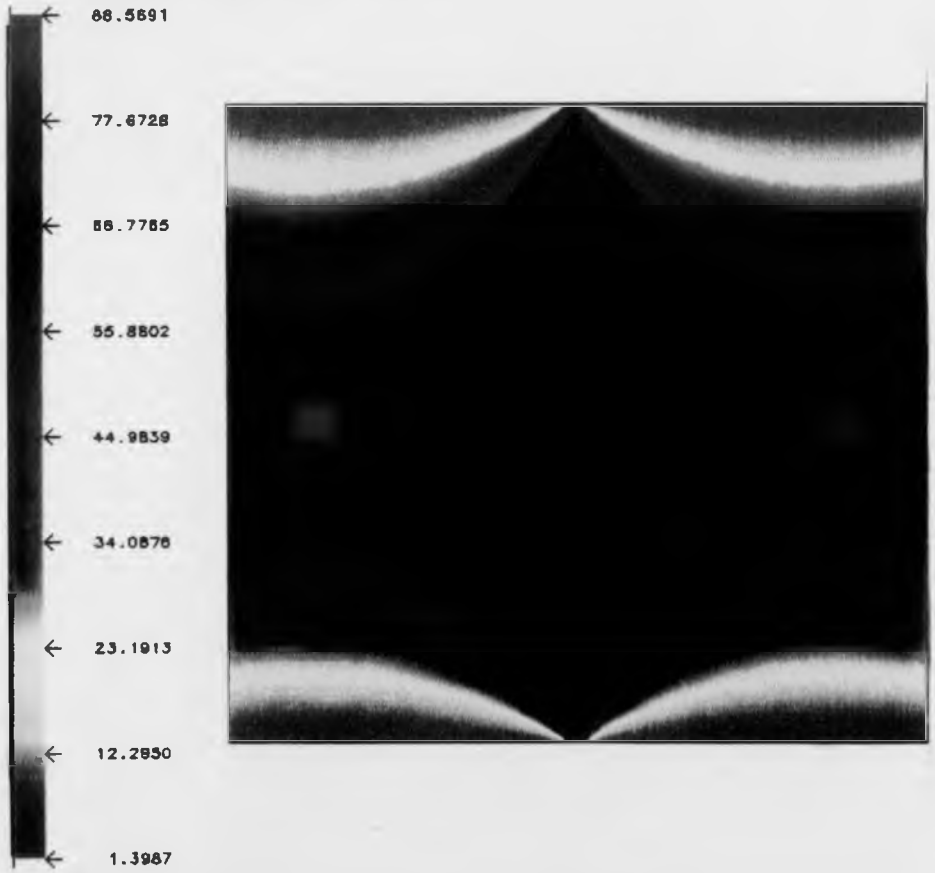


Figure 147: Forward projection of the trial solution, with the integration weights determined by the length of the ray through each pixel

The reconstructed field after ten iterations of the Lent MART algorithm, with figure 144 used as the trial solution, 64 projections, and a relaxation factor of 0.5, is shown in figure 148. The rms reconstruction error is 0.66% and the discrepancies of the maximum and minimum values in the field are 0.37% and 0.18% respectively. Although the reconstruction exhibits periodic noise, the image is a faithful reproduction of the CFD phantom.

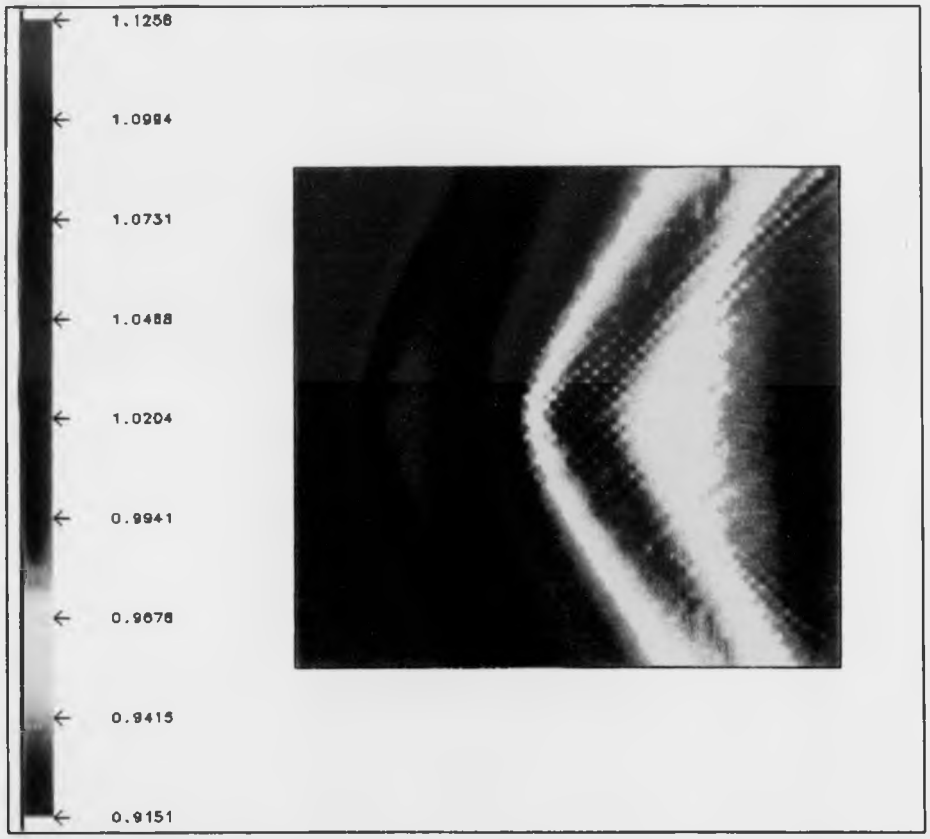


Figure 148: Reconstructed field with the Lent MART algorithm

A study was then made to determine how reconstruction accuracy varies with viewing angle. Initially, the CFD data field was reconstructed using projections calculated for the complete $\pm 90°$, to determine whether this would eliminate the reconstruction artefacts lying along the diagonals of figure 148. The result (figure 149) exhibits markedly reduced noise (rms error of 0.187%) and suggests that mottling of

reconstructed images is caused by aliasing of the data field.

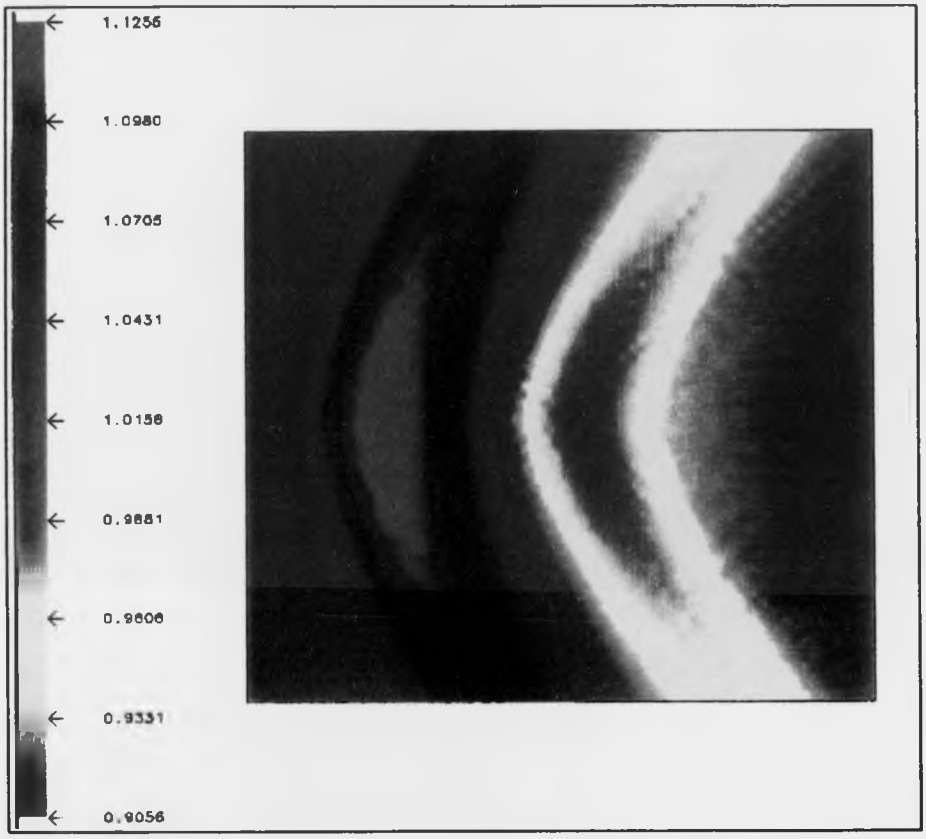


Figure 149: Reconstruction of the CFD phantom using Lent MART with complete angular coverage

The field was then reconstructed for projections in the range $[\pm 40°, \pm 90°]$, to establish the critical viewing angle. The variation of reconstruction accuracy with angle is shown in figure 150. This graph illustrates that an angle of $\pm 47°$ is sufficient to obtain a reconstruction accuracy of approximately 1%. A similar result was obtained in the theoretical study presented in section 8.2.2, which shows that ART algorithms perform close to the limitations imposed by information theory. A curve fit relating the percentage rms reconstruction accuracy to the viewing angle in degrees θ (solid line of figure 150), is described by equation 144. The data points and curve are in good agreement, which shows that the rms reconstruction accuracy is approximately

inversely proportional to the viewing angle.

$$e_{rms} = 0.844/\theta + 0.0077 \tag{144}$$

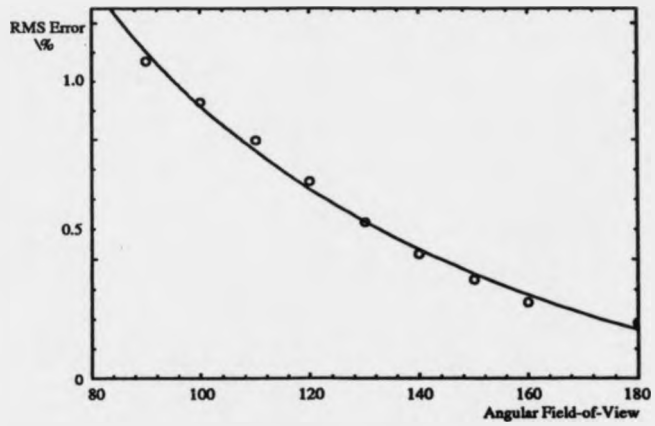


Figure 150: Reconstruction accuracy versus viewing angle for 64 projections

A similar investigation was then made to determine the variation of reconstruction accuracy with the number of equally spaced projections. A viewing angle of $\pm 90°$ was used and the number of views was incremented in the range [4,64]. The results (figure 151) show that there is little improvement in accuracy after approximately sixteen views have been recorded. Increasing the number of views beyond this value, increases image resolution without significantly improving the average error. The curve fit to this data is defined by equation 145, which shows that the rms reconstruction accuracy is also approximately inversely proportional to the number of views.

$$e_{rms} = 0.032/M_{proj} + 0.0025 \tag{145}$$

Although figure 151 shows that it is unnecessary to record more than sixteen views when the viewing angle is $\pm 90°$, the result does not indicate whether a smaller angle may suffice. Consequently, the variation of reconstruction accuracy versus viewing

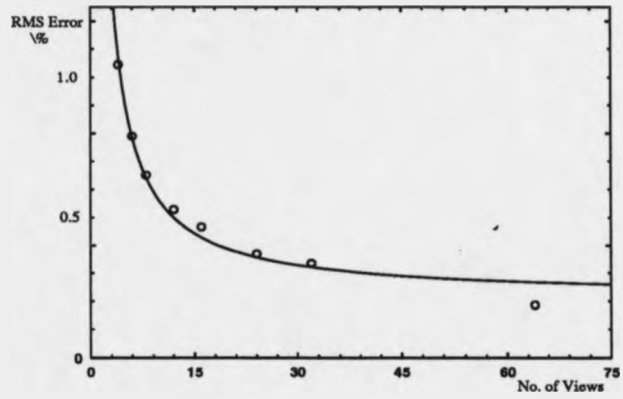


Figure 151: Reconstruction accuracy versus the number of projections

angle, was determined for sixteen projections. The resulting graph (figure 152) exhibits a similar inverse proportionality to figure 150 and shows that a reconstruction accuracy of 1% is possible with a viewing angle of approximately ±58°. An angular field-of-view of ±58° was achieved in the wind tunnel tests at City University (chapter 7) and therefore this result shows that in the absence of systematic errors and noise, the density field could have been calculated to an accuracy of 1% if sixteen views had been recorded.

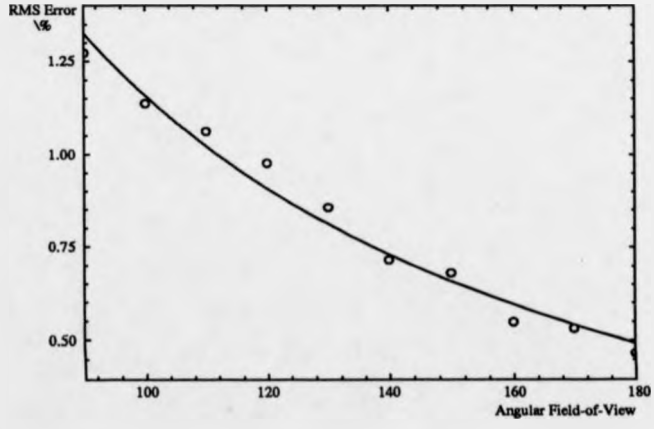


Figure 152: Reconstruction accuracy versus viewing angle for sixteen projections

The effect of the relaxation parameter on reconstruction accuracy, with a viewing angle of ±90° and 64 views, is illustrated by the convergence graph of figure 153. When a small factor is used (0.04), the data field is modified only slightly between successive iterations and therefore the convergence is slow but stable. A relaxation factor of 0.4 produces rapid convergence, with little improvement in the solution after five iterations. When a factor of 0.5 is used, the convergence is unexpectedly slower, although the final reconstruction accuracy is slightly higher. The reduced rate of convergence may be attributed to the solution 'overshooting' with successive iterations. In the limit where a unity relaxation factor is used, the solution does not converge. It is likely that critically damped convergence occurs for a relaxation factor in the range $0.5 > \lambda > 0.4$. The minor discrepancy in the final reconstruction accuracies, for relaxation factors of 0.4 and 0.5, is difficult to explain and may be an effect peculiar to the data field.

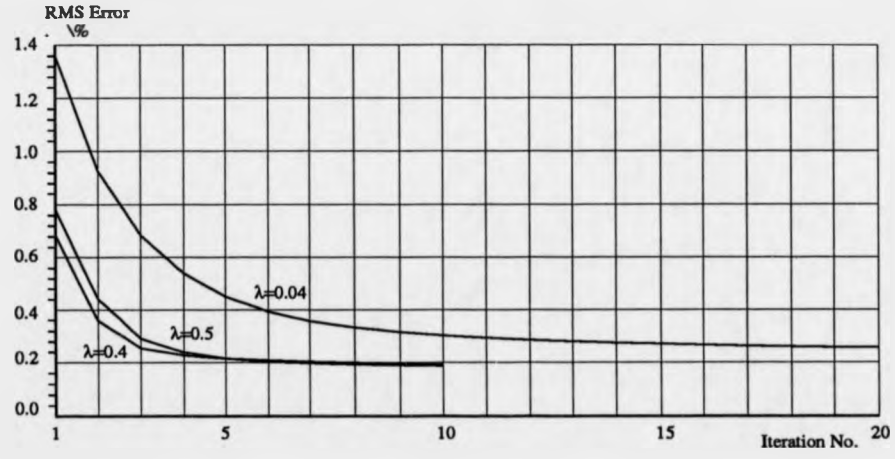


Figure 153: Effect of the relaxation parameter λ on convergence

The accuracy of GH MART was subsequently compared to the Lent MART algorithm, with a viewing angle of ±90° and 64 projections. The GH MART algorithm gave a visually identical reconstruction to Lent MART, with an rms error after ten iterations of 0.196%. The relative performance of these two algorithms is best demonstrated by examining the respective convergence graphs of rms error versus iteration

number (figure 154). After a single iteration the GH MART algorithm is more accurate, with an error of 0.64% compared to 0.78% for Lent MART. The accuracies of the two algorithms converge with further iterations and Lent MART ultimately attains a slightly more accurate solution (0.009% better). However the difference is negligible and the best algorithm is probably a function of the input data.

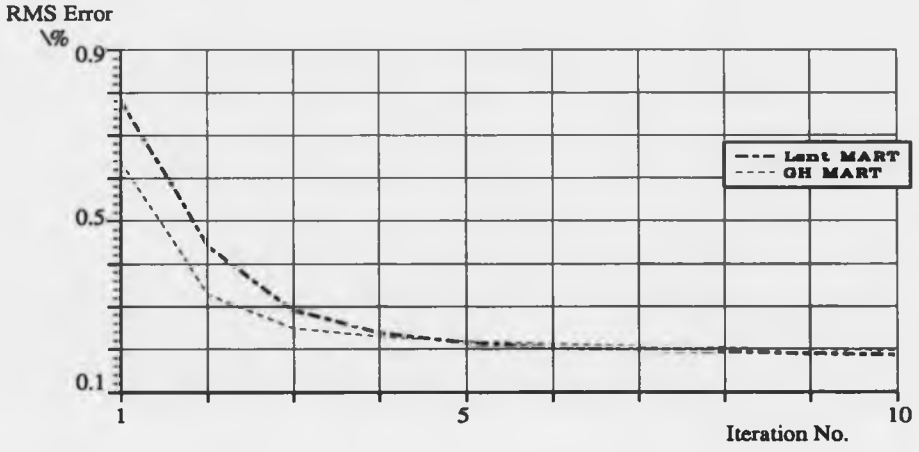


Figure 154: Convergence of Lent MART and GH MART reconstruction algorithms

Similar results were obtained when Lent MART was compared to ART2, with 64 projections gathered over ±90°. Figure 155 shows the convergence histories of Lent MART and ART2, when a unity field was used as the initial trial solution. After a single iteration, the rms error is 1.92% for ART2 (no relaxation factor) and 1.25% for Lent MART (relaxation factor of 0.5). These errors are larger than those from the previous study, because a unity field was used as the initial guess. The rms reconstruction accuracy after this first iteration is higher for Lent MART, even though a relaxation factor was used for this algorithm.

Stable convergence was reached after approximately ten iterations and the rms errors after twenty iterations were 0.24% for Lent MART and 0.22% for ART2. These final error values are larger than those from the previous study, which shows that reconstruction accuracy at convergence is influenced by the quality of the trial solution. In addition, the results show that Lent MART converges faster, but ART2 reaches a

more accurate final solution. This conclusion was also reached in a study by Herman [270].

Figure 156 shows the difference between the true field and the reconstructed image, after twenty iterations of the ART2 algorithm. This error map shows that the reconstruction noise is predominantly along the diagonals of the image. Removal of these artefacts by Fourier filtering is described in the next section.

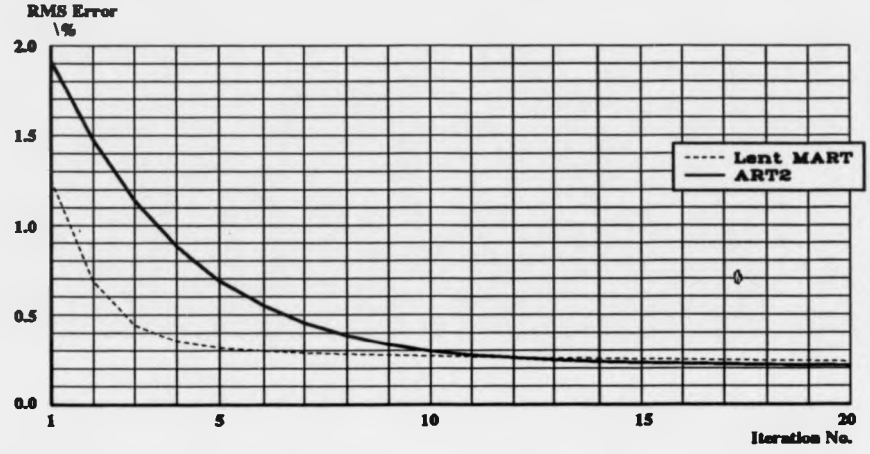


Figure 155: Convergence of Lent MART and ART2 reconstruction algorithms

8.3.3 The Effect of Experimental Errors on Reconstruction Accuracy and the Reduction of Aliasing Artefacts by Spatial Filtering

The reconstructed images of the CFD data field presented in the section 8.3.2, contain periodic high frequency noise. The density field contains lower spatial frequencies and therefore it should be possible to filter this noise. This problem was investigated by calculating the Fourier transform of the reconstructed image after fifteen iterations of the ART2 algorithm. A comparison of this image with the spectrum generated from the original CFD phantom, showed that the noise was localised in the frequency plane. The noise was subsequently removed by a windowing operation and the inverse Fourier transform of the spectrum calculated to produce the filtered image. A further

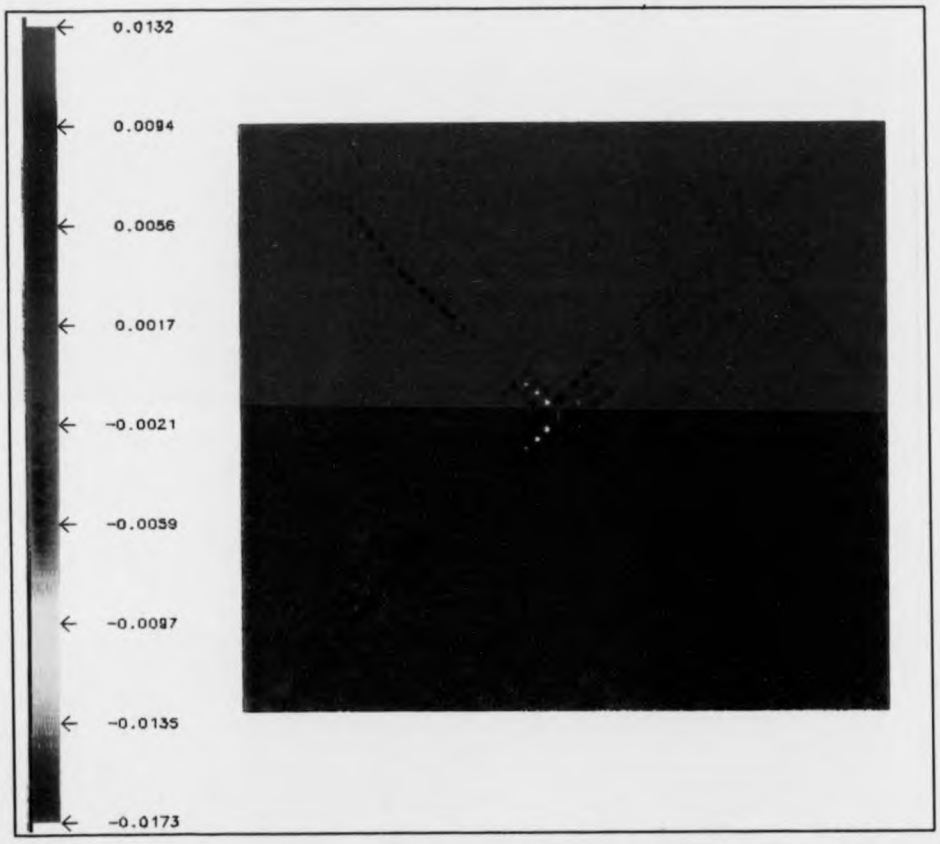


Figure 156: Error Map after Twenty Iterations of the ART2 Algorithm

fifteen iterations of the ART2 algorithm were then made to assess the convergence characteristics after filtering.

The convergence history before and after Fourier plane filtering is shown in figure 157. This graph shows that the filtering operation reduces the rms phase error from 0.62% to 0.38%. In conclusion, spatial filtering is effective when the noise spectra is confined in the Fourier plane. A disadvantage of the technique is that the Fourier transform operation is computationally intensive.

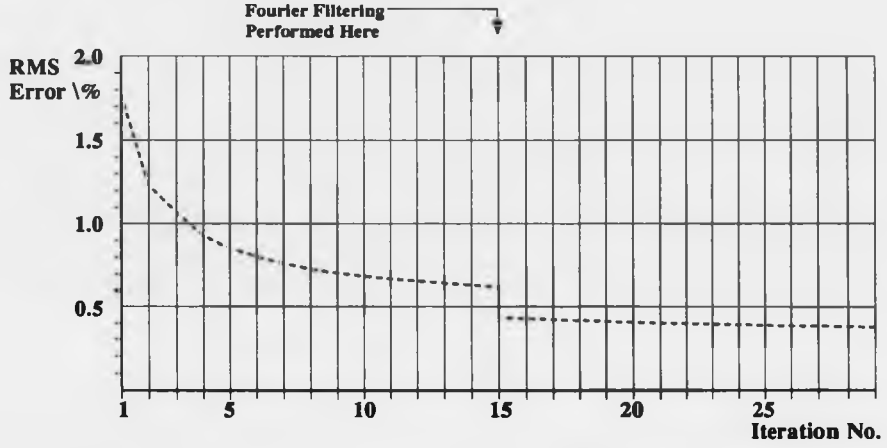


Figure 157: Convergence of the ART2 reconstruction algorithm before and after Fourier plane filtering

In addition to aliasing, reconstruction errors also occur when the projections are inconsistent due to experimental error and noise. Sources of experimental inaccuracy are summarised below:

- Spatial errors arising from refraction of the light through large density gradients.
- Presence of error fringes in the interferograms due to vibration and air currents.
- Fringe analysis errors.
- Speckle and other coherent noise.
- Inaccuracies in measuring the projection angles.

- Occlusion of data by shadows cast by the wind tunnel model.

The effect of experimental noise on reconstruction accuracy was simulated by adding random noise to the input projections. Although this study does not model the sources of experimental error precisely, it should provide insight into the sensitivity of MART algorithms to inconsistencies between projections. In routine wind tunnel testing, a CFD prediction of the flow may be available and therefore the superior trial solution was used (figure 144). The Lent MART algorithm was used as it has been shown to produce a stable convergence when noise is present [268]. A viewing angle of $\pm 60°$, and 64 projections, were used so that the results could be compared to the earlier studies.

The 64 forward projections of the CFD phantom were corrupted by adding random noise with a maximum amplitude of $\pm 5.0\%$. This noise level was chosen because it is representative of the discrepancies between typical holographic measurements and a CFD prediction (chapter 6). Initially, the relaxation factor was set to 0.5 as this was the optimum value in computer simulations with noise-free data. However, convergence was not obtained until λ was reduced to approximately 0.01.

The reconstructed field is shown in figure 158 and the convergence of the solution is compared graphically with the result obtained for noise-free data in figure 159. The graph shows that optimum convergence is obtained after three iterations, whereafter the solution diverges. The maximum rms reconstruction accuracy is 1.7% and after ten iterations the error has increased to 2.2%. Also shown in figure 159 is the convergence graph obtained when the projection data has been pre-processed by a three pixel averaging mask. The convergence of the smoothed projections is more stable, with a minimum rms error of 1.4% after approximately seven iterations.

The computer runs were repeated with the random noise reduced to an amplitude of $\pm 2.5\%$. The reconstructed field with a reduced noise level is visually similar to figure 158, but the convergence graph (figure 160) shows that the iterations are more stable and exhibit lower minimum errors (1% for the smoothed projections).

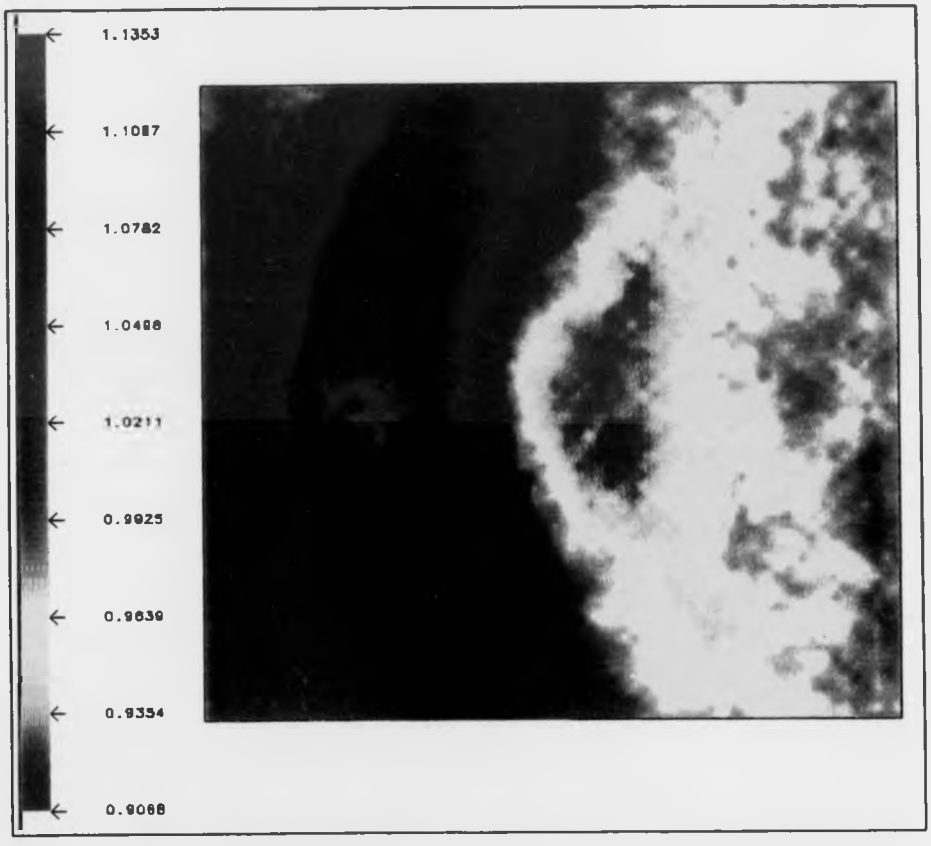


Figure 158: Lent MART reconstruction with the projections corrupted by ±5% random noise

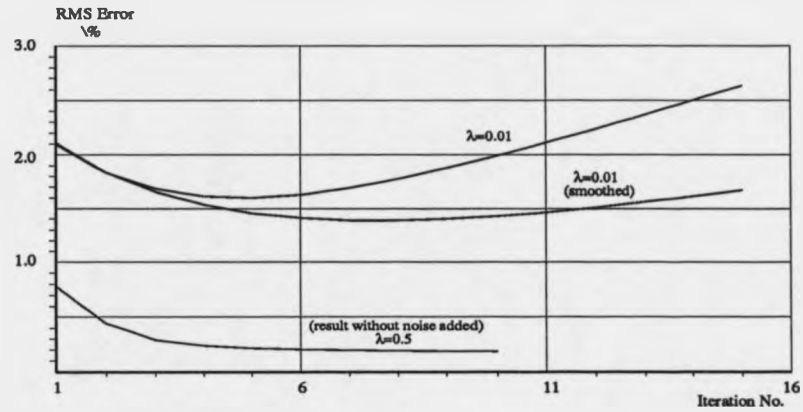


Figure 159: Convergence of Lent MART for uncorrupted, noisy (±5%) and smoothed projections

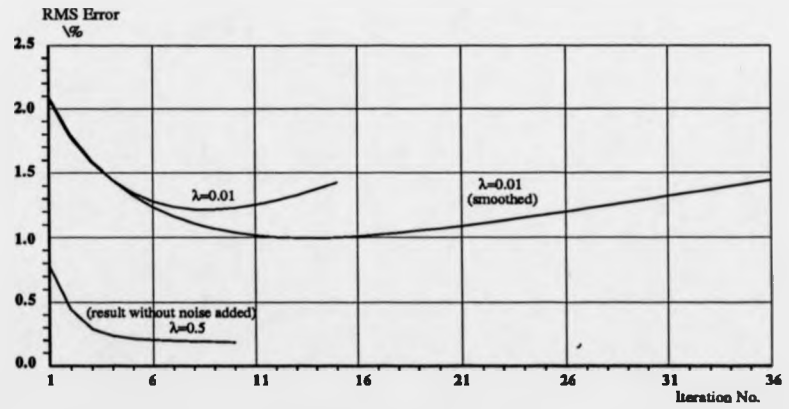


Figure 160: Convergence of Lent MART for uncorrupted, noisy (±2.5%) and smoothed projections

8.4 The use of Tomographic Imaging in Production Wind Tunnel Testing

8.4.1 Optical Systems for Recording Multi-Directional Interferograms

To record the multiple interferometric views required for tomographic reconstruction of a flow, the input illumination must be multiplexed in either the spatial or temporal domains. The wind tunnel test performed at City University (chapter 7), used three spatially multiplexed beams generated by HOE's, to interrogate the flow at angles of 0° and ±58°. A schematic diagram of a similar double-pass configuration, for use in a transonic wind tunnel with a perforated liner, is illustrated in figure 161. This geometry simplifies the optics by using both transmission and reflection HOE's to generate counter-propagating beams. An additional advantage of this arrangement is that it reduces the effect of vibration by locating the transmitting and receiving optics on a common mass. Complete angular coverage could be achieved by generating two fans of beams (figure 162).

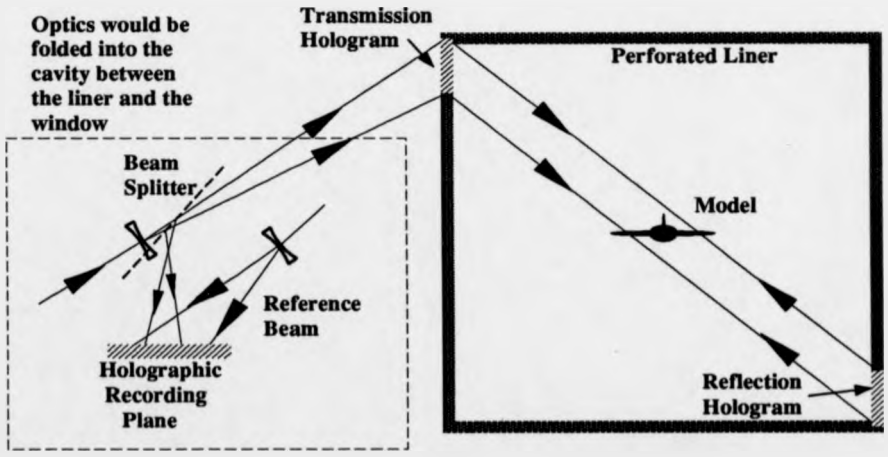


Figure 161: The use of transmission and reflection HOE's to generate a collimated beam crossing a wind tunnel

The principal advantage of this design over temporally multiplexed systems, is that all of the views are recorded simultaneously, so that unsteady flows may be studied

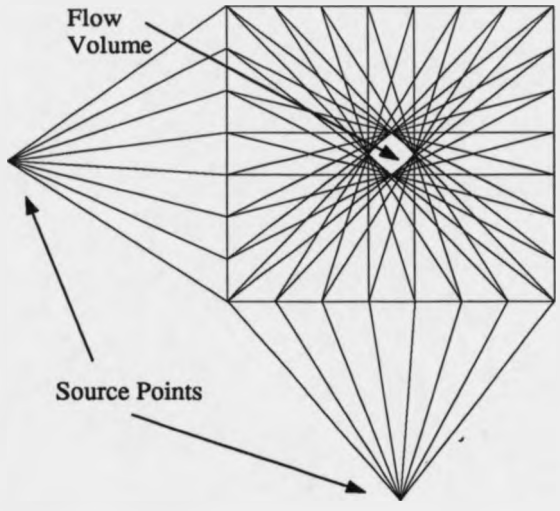


Figure 162: Intersection of the two fans of laser beams

and wind tunnel run-time is minimised. However, there are also numerous drawbacks which must be considered:

- The optical system is highly complicated and costly to manufacture (duplicated optics).
- The intersection area of the beams is small.
- Parts of the flow may not be interrogated by all of the beams.
- The optical efficiency and resolution is different for each of the channels.
- Beam alignment is difficult.
- A powerful laser is required to illuminate all of the channels simultaneously.
- The flow quality at transonic speeds would be affected by insertion of the HOE's in the perforated liner.

These practical problems are not easily resolved for a typical production wind tunnel. Nevertheless, a spatially multiplexed design may be feasible for studying an open jet, where optical access is simplified.

In general, a more elegant solution is to use a single optical channel, with the exposures recorded sequentially as the wind tunnel model is rotated about a longitudinal axis. This configuration was used (without model rotation) during the test at Cranfield Institute of Technology (chapter 6) and the resulting interferograms were of high contrast and simple to reconstruct. The recording system may be automated if a film drive is used and instabilities in the flow may be examined by recording a series of holograms at each viewing angle. The main disadvantage of this geometry is that the sting must be re-orientated to obtain the correct views. Also, the constraints of a square working section may generate inconsistent views. These problems would be most severe when the flow is highly three-dimensional and are topics of future research.

A more fundamental restriction to tomographic data analysis arises due to the opacity of the wind tunnel model. When the model is rotated, a body of revolution is formed and areas in the flow are sampled over a limited angle even when there is complete external optical access. For example, if a typical aircraft configuration is considered, optical access close to the wing surfaces is severely limited due to the silhouette cast by the fuselage. When there are stores present, this shadowed region encroaches further into the freestream. Unfortunately, these occluded regions often contain important flow features, such as points where shock waves impinge on the model's surface.

Regions close to the surface of a wind tunnel model may be visualised by utilising reflected light. This capability is unique to holographic interferometry, because conventional interferometers require specular illumination. An additional simplification is provided if the light traversing the flow is retro-reflected by a planar mirror, so that both transmitted and reflected beams are recorded by a single hologram. This arrangement is illustrated schematically in figure 163, where a hologram of static air has been recorded and replaced in the system to enable real-time observation of the 'live' fringe pattern. Unlike double-exposure techniques, real-time interferometry does not require photographic processing of the exposures holograms as the fringe

pattern is viewed directly by the CCD camera. Rigid body motion of the optics results in finite fringes, which may be removed during Fourier transform fringe analysis. High frequency vibration, that causes blurring of the fringes, may be tolerated if the interferograms are captured using pulsed illumination.

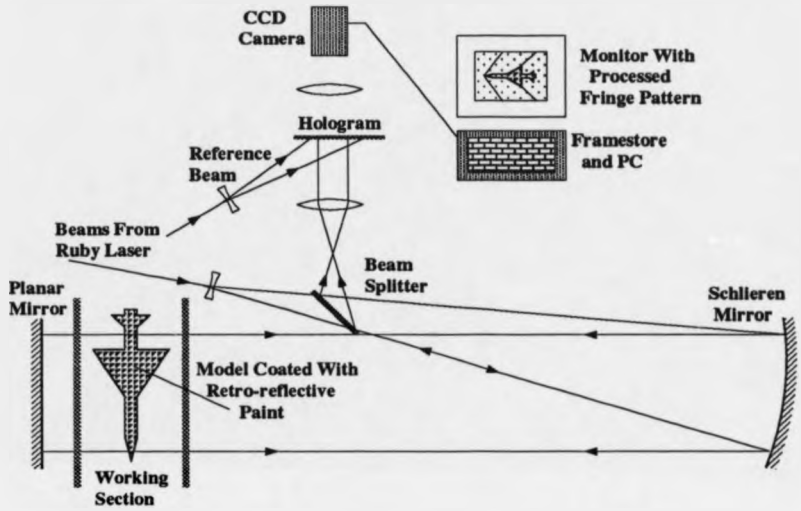


Figure 163: Schematic representation of a practical real-time holographic interferometer

The precise system proposed in figure 163 is impractical, because there would be a large mismatch in the intensities of the beams reflected from the model and planar mirror. A metallic model produces both specular and diffuse reflections. A large proportion of the specular contribution would be lost from the optical system, because of reflection from curved surfaces of the model. The diffuse component would be scattered into a solid angle of 4π steradians for a Lambertian surface and would also be severely depleted before arrival at the recording plane.

The reflected beam may be enhanced by coating the model with retro-reflective paint or tape. This material consists of a suspension of glass micro-spheres which retro-reflect the incident beam by total internal reflection. The reflected beam should therefore be highly directional and significantly brighter than a diffuse reflection. Also, speckle noise and fringe localisation should be less severe than for diffuse geometries,

because most of the light is retained within the system and the rays travel through the phase object in predominantly the same direction.

The performance of retro-reflective paint was investigated by building a laboratory based demonstrator (figure 164). A lenticular wind tunnel model, equipped with fins, was sprayed with 3M 'Scotchlite' 7210 silver retro-reflective paint. A phase object was created in front of the model by using a heated soldering iron to produce a plume of hot air. The fringe pattern was generated by recording a pulsed ruby double-exposure hologram, with the soldering iron heated for the second exposure.

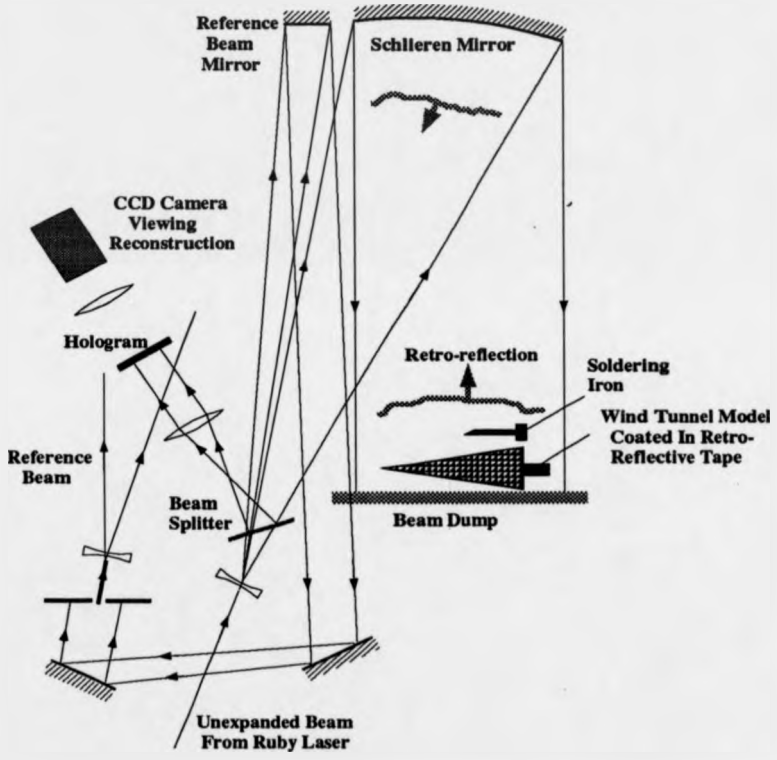


Figure 164: Double-pass holographic interferometer utilising retro-reflected beams

The hologram was reconstructed with a He-Ne laser and the interferogram was captured using a CCD camera and framestore. In the centre of the image (figure 165), the soldering iron is shown silhouetted against the wind tunnel model (33mm width), with the fins visible at the right and left of the picture. The fringes produced by the

lower density hot air are clearly defined and a set of finite fringes are also visible due to rigid body motion of the apparatus between exposures.

Although this result demonstrates the feasibility of imaging fringe patterns in retro-reflected light, the interferogram exhibits speckle noise because an aperture stop was required to obtain a sufficient depth-of-field. This is a consequence of the fringe localisation that occurs when light is scattered from the optical system. A large reduction in speckle noise, and corresponding increase in the reflected signal, was subsequently achieved by attaching faceted retro-reflective tape to the model. Unfortunately, this material prevented fringe analysis, because the facets segmented the image. Numerous other retro-reflective paints and tapes are commercially available, and it recommended that future research is spent examining these materials.

In conclusion, both spatially and temporally multiplexed holographic interferometers have a role to play in routine flow analysis. The spatially multiplexed design is suitable when the flow volume is compact and easily accessible from large angles; whereas temporal multiplexing is more practical for typical production wind tunnels. In addition to enabling the recording of interferometric data, both optical systems provide multi-directional imaging of the flow volume. Hence, an integrated system, combining two-wavelength holographic interferometry and stereoscopic PIV, could be used to measure two state variables and velocity. The interferometer could be operated in real-time, if rigid body motion is removed during Fourier fringe analysis, and shadowed regions could be interrogated by coating the model with retro-reflective paint or tape.

8.4.2 The Reconstruction of Incomplete Projections by Complimentary Field Methods (CFM)

When the density field of a compressible flow is measured by holographic interferometry, the viewing angle is usually restricted by the test section and the projections are incomplete due to the opacity of the wind tunnel model. If filtered back projection,

Figure 165: Fringe pattern around a soldering iron illuminated by retro-reflected light

Fourier inversion, or algebraic reconstruction techniques are applied in these circumstances, the reconstructed field contains streaked artefacts and is too inaccurate for comparison to CFD predictions.

If the missing data occupies only a small percentage of the views, then series expansion methods may be used successfully [271]. In practice, however, the occluded regions are usually too extensive for successful implementation of this technique. Lewitt *et al* [272, 273] have shown that reconstruction accuracy is dramatically improved by interpolating across the missing data, so that the empty field is augmented by fictitious values.

Other methods which have been tried include iterative schemes that alternate between object, projection and Fourier space. One particular technique which appears to offer the most promise is the complementary field method (CFM). This approach was originally proposed by Heffernan and Rob [232], and subsequently developed by Cha and Sun [274]. The CFM algorithm may be summarised by a series of simple operations:

1. Determine an estimate for the unknown field $f_e(x, y)$.
2. Forward project the estimated field, ignoring the missing data, to give $l_e(p, \theta)$.
3. Calculate the difference between the measured and estimated views ($l_{exp}(p, \theta)$-$l_e(p, \theta)$).
4. Reconstruct this difference (complementary) field $f_c(x, y)$.
5. Revise the estimated field by adding the complementary field ($f_e(x, y)+f_c(x, y)$).
6. Repeat iterations with the revised field until a convergence criterion is satisfied.

In many respects this technique is similar to ART, except that the error projections are used to reconstruct the complementary field, instead of being used to directly modify the estimated field. This process prevents oscillations of the solution, thereby

aiding convergence. The missing data in the views is not required if a series expansion method is used to reconstruct the complementary field (step 4). If filtered back projection is employed, then interpolation must be used to augment the missing data.

Cha and Sun have demonstrated that the CFM algorithm produces a stable convergence when most other techniques diverge or produce erroneous data [275]. The technique provides a simple mechanism for incorporating a *priori* knowledge and combines the virtues of both ART and filtered back projection strategies. They conclude that although convergence may be slow, and the computing overheads are large, the technique enables the reconstruction of severely ill-posed data when a reasonable estimate of the unknown field is available. Although this method is not covered in this thesis, it is a topic of current research and will be used to extend the tomographic reconstruction of interferometric data to occluded flows.

8.5 Conclusions on the Reconstruction of Interferometric Flow Data

This chapter has demonstrated the feasibility of reconstructing multi-directional interferometric data using computer aided tomography. Various reconstruction algorithms have been reviewed and the Fourier slice theorem has been invoked to show that, in principle, a reconstruction accuracy of approximately 1% may be obtained when the viewing angle is limited to $\pm 45°$.

A computer study using CFD input data was then undertaken to establish the variation in reconstruction accuracy with viewing angle and number of projections. Algebraic reconstruction techniques were used, as these perform better than alternative algorithms when the data is sampled over a limited angle and provide a simple mechanism for incorporating *a priori* knowledge. The forward projection process was approximated by a discrete integration through the data, with the length of the rays through each pixel used to determine the weights of each of the contributions. A simpler approach, which uses binary weights, was shown to perform badly because it exaggerates the discrete nature of the data. This study confirmed the previous result that a viewing angle of $\pm 45°$ is adequate if the projections are free from noise. Also, it showed that sixteen equally spaced views are sufficient to reconstruct a field of $64 \times 64 \times 64$ values.

When the views contain experimental noise, the solution may diverge after several iterations because of the inconsistencies between projections. Divergence of the reconstruction may be delayed, and a more accurate solution obtained, if a smaller relaxation factor is used in the algorithms. The solution is also improved when an accurate initial estimate is used to initiate the iterative process. This trial solution may be provided by a CFD prediction and was modelled in this study by amending the original data field.

This study shows that the algebraic reconstruction algorithms GH MART, Lent MART and ART2 produce results of similar accuracy. It is likely that the best algorithm is a function of the imagery used, with Lent MART performing comparatively

better when the field contains discontinuities and ART2 preferred when the data contains lower spatial frequencies. The Lent MART algorithm has been shown to converge to the maximum entropy solution and ART2 produces data with minimum variance from the correct field. Although these conditions are mathematically elegant, they are not definitive methods for optimising the reconstruction of flow images.

When the flow contains opaque obstructions, such as the wind tunnel model or sting, the reconstruction problem is ill-posed even when the data is adequately sampled around the full 180°. Consequently, the reconstructed images are inaccurate and streaked due to aliasing. Artefacts may be suppressed if the missing data is approximated by interpolation and the field is reconstructed by complementary field methods.

9 Conclusions and Recommended Areas of Future Research

At present, there is no single whole-field flow diagnostic technique which can simultaneously measure two state variables and velocity at transonic Mach numbers. Without this measurement capability, an understanding of flow physics is incomplete and CFD predictions cannot be accurately evaluated. This research has made progress towards establishing this complete experimental capability, by developing holographic interferometry to a level where it may be used routinely for measuring three-dimensional density information. If this system is extended to dual wavelength operation, and integrated with PIV, then all of the flow variables may be determined.

Interferometry is particularly effective at transonic speeds, because fringe densities are optimal, and the flow is unperturbed by the measurement process. Alternative diagnostics usually require particle seeding, which is inconvenient and difficult to administer in sufficient quantities at compressible Mach numbers. Also, particle inertia reduces measurement accuracy at these speeds. Holographic interferometry has the additional advantage that schlieren, shadowgraph and interferometric images may be extracted from a single recording for qualitative flow evaluation.

The principal disadvantage of holographic interferometry is that the encoded flow information is line integrated. Hence, when three-dimensional wind tunnel models are tested, multi-directional image capture and tomographic inversion are required to extract the density field. Also, automated fringe analysis has been hampered by practical difficulties of data acquisition and noise reduction. These problems have recently been ameliorated by the availability of inexpensive high quality CCD cameras, framestores and computer workstations.

The experimental results presented in this thesis demonstrate how high quality holograms can be recorded in harsh wind tunnel environments. Initially, a NACA 0012 aerofoil was tested in the Cranfield continuous running transonic wind tunnel, to illustrate data extraction for two-dimensional flows. Specular and diffuse illumination

geometries were implemented and the results show that fringe contrast is higher when specular wavefronts are used. Additional advantages are that speckle noise is reduced, optical efficiency is high, and the fringes are delocalised.

Experimental fringe patterns vary in contrast and background intensity. Consequently, data extraction is inaccurate unless amplitude noise is filtered by phase sensitive processing. The phase of the reconstructed wavefronts may be manipulated by recording the two objects waves on different holographic plates, or by encoding the exposures with different reference beam angles. During the wind tunnel testing, sandwich holography was found to be impractical and therefore the dual reference beam method was adopted. Both wide and narrow angle configurations were examined, with the interferograms processed by phase stepped and Fourier transform fringe analysis.

It was found that a wide angle reconstruction system is susceptible to vibration and misalignment, and that the theoretical accuracy of phase stepped fringe analysis is difficult to achieve in an industrial environment. In contrast, holograms recorded with a narrow angle can be reconstructed with a single replay beam and therefore stability is increased. Also, these interferograms may be reconstructed with a thermal source instead of laser, because the object wavefronts are automatically path matched. The phase data may be extracted from a single finite fringe pattern using Fourier transform fringe analysis, whereas three or four images are required for phase stepped strategies.

An interferogram of transonic flow over a NACA 0012 two-dimensional aerofoil was processed by Fourier analysis and the principal phase values were unwrapped along a path that minimises inconsistencies on a hierarchy of scales. The phase map was converted into density information by calibrating the measurements against wind tunnel parameters. The resulting data field enables direct evaluation of CFD codes, whereas previous research in this subject has been limited to a subjective correlation of fringe centres with predicted iso-density contours.

The holographic measurements were compared quantitatively with a prediction from a Navier-Stokes CFD code. The respective data fields were differenced and the

results shown pictorially as a pseudo-colour error map. In general, the holographic measurements were in close agreement with the computational prediction. Significant discrepancies occurred in the wake, because a laminar boundary layer model was used in the CFD code, whereas the experimental boundary layer had undergone turbulent transition.

The accuracy of Fourier transform fringe analysis was assessed using a combination of experimental imagery and computer generated phase data. The results from this study showed that a measurement accuracy of typically 0.1% is possible, if the carrier frequency of the fringes and pixel spacing of the imaging sensor, are carefully tailored to the interferogram. If either of these conditions are violated then aliasing occurs and moiré fringes are generated. Recording non-linearity was shown to reduce the resolution of the phase map, although the accuracy at lower spatial frequencies was relatively unaffected.

A transonic wind tunnel test was conducted at London's City University, to extend the technique to three-dimensional flows. A simple cone-cylinder model was tested, so that the results could be compared to theoretical and computational predictions. Previous attempts by other researchers at capturing multi-directional interferometric data, have involved duplicating the optical system for different viewing angles. This arrangement is impractical in confined test environments, and is prone to vibration and air currents. A more elegant solution was found by combining the versatility of diffractive optics with the efficiency of refractive components.

A prototype optical system was built, which used HOE's to interrogate the flow over a viewing angle of 116°. The geometry was compact and the reconstructed interferograms had a high signal-to-noise ratio. A spatial resolution of approximately 250μm was obtained and, if necessary, this could have been improved by carefully controlling emulsion shrinkage.

The optical efficiency of the interferometer was predetermined theoretically and it was discovered that adequate exposure densities could be obtained using amplitude HOE's. Bleached phase elements were not used, as these produce significantly

higher Rayleigh scatter. The reconstructed interferograms were of high contrast and demonstrated the suitability of HOE's recorded in silver halide.

The City University wind tunnel operated intermittently and it was found that the wind tunnel windows distorted when the pressure in the working section dropped at transonic speeds. A mechanical solution was sought, but attempts were unsuccessful within the available time. Consequently, a series of empty tunnel calibration runs were made to enable subtraction of the window deformation from the data during image processing. Finite fringe interferograms were recorded of a pointed cone-cylinder model at Mach 1.2 and the error surface was removed from the processed phase maps.

From interferograms recorded at various Mach numbers, it became apparent that the conical shock envelope over the model was elliptical due to the boundary constraints of the working section. The major axis of the ellipse gave shock angles which agreed with cone theory, whereas the minor axis gave angles consistent with a higher freestream Mach number. Although the asymmetry of the shock cone prevented accurate flow analysis, the holographic projections at Mach 1.2 were compared with an Euler CFD prediction. Orthogonal and oblique integrated views were calculated from the CFD solution and the results were qualitatively in good agreement with the experimental measurements.

Computed tomography was reviewed to determine which reconstruction algorithm is most suitable for processing three-dimensional interferometric data. The study showed that algebraic reconstruction methods are the most accurate, when the problem is ill-conditioned due to limited optical access. In addition, these methods provide a simple mechanism for incorporating *a priori* flow information from CFD predictions. A theoretical study, based on the central slice theorem, showed that the critical viewing angle is approximately $\pm 45°$. A similar result was obtained in a computer simulation using CFD density data of transonic flow around a cone. In practice, the angular field-of-view needs to be increased to approximately $\pm 60°$ when noise is present, to prevent divergence of the tomographic solution. Also, opaque obstructions in the flow, such as those generated by the wind tunnel model, complicate analysis

and induce further ill-conditioning. Future research should explore complimentary field techniques, based on both ART and filtered back projection strategies, to enable interpolation of data into these regions. Also, measurement inaccuracies produced by refraction should also be evaluated. If there are no large density gradients in the flow, then ray-tracing could be used to reduce these errors during ART reconstruction.

Although this thesis has demonstrated the feasibility of three-dimensional holographic flow analysis, there are still numerous problems which must be addressed before the technique can fulfill its true potential. Initially, the use of retro-reflective wind tunnel models should be investigated, to enable the acquisition of flow data in occluded regions. This subject was considered briefly in chapter 8, but a more extensive study should be made to determine the best type of retro-reflective coating. Also, an optical system should be developed that records the interferograms sequentially as the model is rotated. A principal part of this research would be the design of a sting that re-orientates the model between exposures.

At present, transonic flows cannot be fully characterised by a single diagnostic tool and therefore it is recommended that a hybrid system is developed, which integrates stereo PIV with dual wavelength holographic interferometry. This thesis has shown how Fourier analysis may be used to remove error fringes produced by rigid body motion and therefore real-time interferometry could be employed once frame-rate fringe analysis has been realised. Currently this is only feasible with dedicated hardware, but soon computer workstations should provide adequate processing power.

Once research into multi-directional holographic interferometry has been completed, there will need to be a development period to integrate the technique with production testing. In conclusion, the commercial importance of experimental flow analysis at transonic speeds, justifies further allocation of resources to refine this flow diagnostic.

Appendix 1

An infinite fringe interferogram may be described by equation 146, with the approximation that the spatial dependence is sinusoidal:

$$i(x,y) = a(x,y) + b(x,y)\cos[\phi(x,y)] \tag{146}$$

Where,

$a(x,y)$ - background fringe intensity
$b(x,y)$ - fringe modulation
$\phi(x,y)$ - phase of the interferogram

Quasi-heterodyne fringe analysis involves the calculation of the phase function from a series of k images, which are sequentially incremented in phase by δ_p. The intensity distribution of the kth image may be described mathematically by equation 147, where the spatial dependence is omitted for reasons of clarity.

$$i_k = a + b\cos(\phi - k\delta_p) \tag{147}$$

Equation 147 may be expanded in terms of the compound angles:

$$i_k = a + b\cos\phi\cos(k\delta_p) + b\sin\phi\sin(k\delta_p) \tag{148}$$

A general solution for the phase of the interferogram may be obtained by considering an infinite summation of these fringe patterns, where each successive term in the series is incremented in phase by δ_p:

$$\sum_{k=1}^{\infty} i_k(k\delta_p) = \sum_{k=1}^{\infty} a + b\cos\phi \sum_{k=1}^{\infty} \cos(k\delta_p) + b\sin\phi \sum_{k=1}^{\infty} \sin(k\delta_p) \tag{149}$$

Equation 149 is a Fourier series expansion of the infinite sum of phase shifted images. This is made apparent if equation 149 is compared to a general definition of the Fourier series of a periodic function $f(x)$:

$$f(x) = a_0/2 + \sum_{k=1}^{\infty}[a_k\cos(kx) + b_k\sin(kx)] \tag{150}$$

The terms a_k and b_k are the Fourier coefficients, which are defined over an arbitrary period λ as:

$$a_k = \frac{1}{\pi}\int_{\lambda} f(x)\cos(kx)dx \tag{151}$$

$$b_k = \frac{1}{\pi}\int_{\lambda} f(x)\sin(kx)dx \tag{152}$$

In the case of the infinite summation of phase stepped fringe patterns, the coefficients are all equal and given by inspection as:

$$a_k = b\cos\phi \tag{153}$$

$$b_k = b\sin\phi \tag{154}$$

Hence, by the definitions used in equations 152 and 151, expressions relating the kth constant coefficient to the summed images, are given by:

$$b\cos\phi = \frac{1}{\pi}\int_{\lambda}\sum_{N=1}^{\infty} i_N(N\delta_p)\cos(k\delta_p)d\delta_p \tag{155}$$

$$b\sin\phi = \frac{1}{\pi}\int_{\lambda}\sum_{N=1}^{\infty} i_N(N\delta_p)\sin(k\delta_p)d\delta_p \tag{156}$$

Where a dummy variable N has been used for clarity.

These expressions may be simplified because the fringe patterns vary periodically when the phase is shifted by a multiple of 2π:

$$i_N(N\delta_p) = i_N(N\delta_p + 2n\pi) \tag{157}$$

Hence, the infinite summations may be replaced by finite summations over a single period, multiplied by an integer constant:

$$\sum_{N=1}^{\infty} i_N(N\delta_p) = c\sum_{\lambda} i_N(N\delta_p) \tag{158}$$

The constant c describes the number of periods in the infinite domain and is meaningless (infinite) except when divided by itself.

An additional simplification is afforded from the fact that δ_p is not a continuous variable so that the integrations in equations 155 and 156 may be replaced by discrete summations. Hence, if a ratio is taken of the sum of N coefficients ($N\delta_p < 2\pi$) within a single period, the expressions become:

$$\sum_{k=0}^{N} \sin\phi = \sum_{k=0}^{N} i_k(k\delta_p)\sin(k\delta_p) \tag{159}$$

$$\sum_{k=0}^{N} \cos\phi = \sum_{k=0}^{N} i_k(k\delta_p)\cos(k\delta_p) \tag{160}$$

However, as the coefficients are constant, the left-hand sides are simply $N\sin\phi$ and $N\cos\phi$ respectively. Consequently, if a ratio is taken of equations 159 and 160, the wrapped phase of the interferogram is given as:

$$\phi = \tan^{-1}\left\{\sum_{k=0}^{N} i_k(k\delta_p)\sin(k\delta_p) / \sum_{k=0}^{N} i_k(k\delta_p)\cos(k\delta_p)\right\} \tag{161}$$

Appendix 2

During recording of the holograms, a sheet of black glass was index-matched with white spirit to the back of the holographic plate, to suppress multiple internal reflections within the substrate. If this step is omitted then the recording is modulated by a macroscopic fringe pattern which contours the thickness variations of the glass and increases the amplitude noise of the flow field interferogram. Residual white spirit was removed from the holograms prior to developing, by immersing them in a water bath containing Ilfotol wetting agent. This process has the additional advantage of hypersensitising the photosensitive layer, so that emulsion speed is increased and the final diffraction efficiency is higher.

The holograms were developed in Neofin Blue at $\sim 26°$ until an exposure density of 0.6D was achieved for the superimposed object and reference beams. This proprietary developer enhances the speed of the emulsion and therefore enables the use of lower laser powers to minimise the risk of damage to the optical components. In addition, Neofin Blue creates very fine filaments of developed silver, which reduces the amount of Rayleigh scatter and therefore increases the contrast of the interferograms.

Appendix 3

Flow quality in the City University transonic wind tunnel was investigated by Sykes [216], shortly after installation of the facility, by measuring side-wall pressure variations. Studies made with the slots in the liners covered, showed that the mean Mach number along the length of the working section was actually 10% higher than that given by plenum pressures at Mach 0.8, with a Mach number gradient of $0.16m^{-1}$. With the slots exposed, however, the local Mach number along the side wall was approximately uniform along the length of the working section, with a maximum Mach number gradient of 0.005 over 400mm. During individual runs, the Mach number could be held constant to within 1% by manual control of the supply valve.

Shortly after tunnel installation in 1968, runs were made with a cone-cylinder model to identify the nature of transient disturbances and mechanical problems of the support structure. A strain gauge balance was used to measure the axial force on the model and the results identified a resonance at 49Hz with a negligible amplitude of $\pm 0.05N$. Turbulence characteristics of the tunnel were studied by calculating the critical Reynolds number, defined as the value at which the difference between stagnation pressure and rear body pressure gives a pressure coefficient of 1.22, for a 5.8mm diameter sphere. The turbulence factor was measured at a single point on the centre line of the working section, 450mm downstream from the end of the sidewall, with the slots in the liner both open and closed. The sealed wall working section had a turbulence factor of 1.025 when corrected for blockage and the slotted wall section had a factor of 1.13 based on a free-air critical Reynolds number of 0.385×10^6. These turbulence factors correspond to turbulence intensities of 0.16% and 0.03% respectively, and were considered low enough to enable a quantitative comparison of holographic measurements with a CFD prediction.

Appendix 4

The aberrations affecting the holographic lens combination used in the wind tunnel system, may be analysed by considering the geometries used to record and reconstruct the HOE's (figure 166). During the recording process two spherical waves, emanating from source positions $r_o(x,y,z)$ and $r_r(x,y,z)$, are superimposed at the hologram plane with radii of curvature R_o and R_r. When the beams are mutually coherent, with a wavelength λ_1, the interference term encoded on the hologram is proportional to:

$$I(x_h, y_h) \propto \exp\left\{j\frac{2\pi}{\lambda_1}\left[r_o(x,y,z) - r_r(x,y,z)\right]\right\} \tag{162}$$

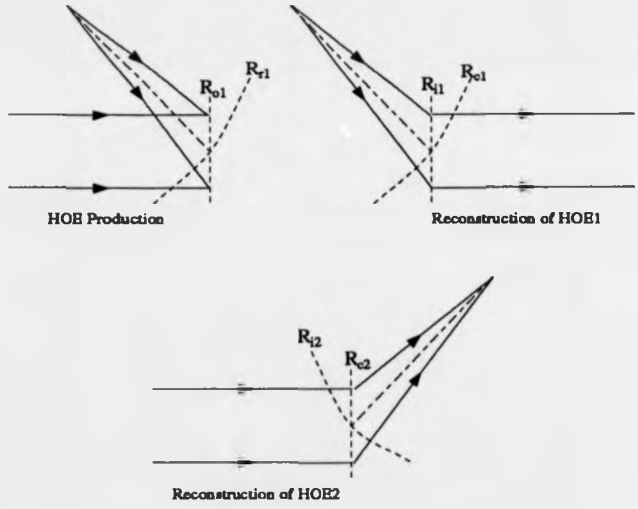


Figure 166: Definition of wavefront curvatures

Reconstruction with a point source located at $r_c(x,y,z)$ generates a series of diffraction orders, with the reconstructed object wave given by:

$$A(x,y,z) = A_0 \exp\left\{j\frac{2\pi}{\lambda_1}\left[r_o(x,y,z) - r_r(x,y,z)\right]\right\} \exp\left\{j\frac{2\pi}{\lambda_2}\left[r_c(x,y,z)\right]\right\} \tag{163}$$

Where A_0 is the peak amplitude of the diffracted wave.

To generate a perfect image of the point source, equation 163 must be of the form:

$$A_i(x,y,z) = A_0 \exp\left[j\frac{2\pi}{\lambda_2} r_i(x,y,z)\right] \tag{164}$$

The imaging properties may be calculated by expanding the exponents of equations 163 and 164 as a Binomial series, with the condition that the longitudinal co-ordinates are an order of magnitude larger than the transverse co-ordinates. The position of the gaussian focus $r_i(x,y,z)$ is determined by equating the first order Binomial terms of equations 163 and 164. Champagne [276] has shown that this defines the focal position as:

$$\frac{x_i}{R_i} = \frac{x_c}{R_c} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{x_o}{R_o} - \frac{x_r}{R_r}\right) \tag{165}$$

$$\frac{y_i}{R_i} = \frac{y_c}{R_c} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{y_o}{R_o} - \frac{y_r}{R_r}\right) \tag{166}$$

Where the radial distances are related by:

$$\frac{1}{R_i} = \frac{1}{R_c} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{1}{R_o} - \frac{1}{R_r}\right) \tag{167}$$

and the positive and negative signs denote real and virtual images respectively.

Similarly, the third order terms of the expansions may be equated to determine the Seidel aberrations. These are described by the following coefficients:

1. Spherical Aberration

$$S = \frac{1}{R_c^3} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{1}{R_o^3} - \frac{1}{R_r^3}\right) - \frac{1}{R_i^3} \tag{168}$$

2. Coma

$$C_x = \frac{x_c}{R_c^3} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{x_o}{R_o^3} - \frac{x_r}{R_r^3}\right) - \frac{x_i}{R_i^3} \tag{169}$$

$$C_y = \frac{y_c}{R_c^3} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{y_o}{R_o^3} - \frac{y_r}{R_r^3}\right) - \frac{y_i}{R_i^3} \tag{170}$$

3. Astigmatism

$$A_x = \frac{x_c^2}{R_c^3} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{x_o^2}{R_o^3} - \frac{x_r^2}{R_r^3}\right) - \frac{x_i^2}{R_i^3} \tag{171}$$

$$A_y = \frac{y_c^2}{R_c^3} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{y_o^2}{R_o^3} - \frac{y_r^2}{R_r^3}\right) - \frac{y_i^2}{R_i^3} \tag{172}$$

$$A_{xy} = \frac{x_c y_c}{R_c^3} \pm \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{x_o y_o}{R_o^3} - \frac{x_r y_r}{R_r^3}\right) - \frac{x_i y_i}{R_i^3} \tag{173}$$

Distortion and field curvature also occur, but these effect only the location of image points rather than their sharpness.

The initial reference beam, with co-ordinates (x_r, y_r, z_r) and radius R_r, is replaced by an identical reconstruction wave during replay, with matching co-ordinates (x_c, y_c, z_c) and radius R_c. In addition, the object and image waves are collimated and therefore have an infinite radius of curvature. Substitution of these conditions into the Seidel coefficients shows that spherical aberrations, coma and astigmatism are zero when the replay conditions match those during recording. When an extended object is present instead of a single source point, for instance due to the scatter and refraction of a model in an air flow, aberrations are generated by both of the component HOE's. However, due to the symmetry of the optical system, the aberrations are minimised.

References

[1] P.C. Steinback. A review of hot wire anemometry in transonic flows. *ICIASF record*, pages 67–78, 1985.

[2] P.W. Bearman, M.M. Freestone, P.F. Hall, and J.K. Harvey. Particle image velocimetry applied to three-dimensional wake flows. In *Flow Mapping using Laser Techniques*, I. Mech. E. HQ, London, 1991.

[3] R.L. Ladenburg, C.C. Van Voorhis, and J. Winckler. Interferometric studies of faster than sound phenomena. Part II. analysis of supersonic air jets. *Physical Review*, 76(5):662–675, 1949.

[4] D.G. Fletcher and R.L. McKenzie. Simultaneous measurements of temperature and density in air flows using UV laser spectroscopy. *ICALEO*, pages 11–19, 1990.

[5] A. Arnold, H. Becker, W. Ketterle, M. Schafer, and J. Wolfrum. Combustion diagnostics by two-dimensional laser induced fluorescence using tunable excimer lasers. *ICALEO*, pages 70–81, 1990.

[6] F. Ahlborne. On the mechanism of hydrodynamic drag. *Abhandl Gebiete Naturwiss, Hamburg*, 17, 1902.

[7] L. Prandtl and O.G. Tietjens. *Applied hydro and aeromechanics.* Mc Graw Hill, New York, 1934.

[8] A. Melling. Seeding gas flows for laser anemometry. *AGARD*, CP-395, 1986.

[9] W. Merzkirch. *Flow visualisation.* Academic Press, New York, 1974.

[10] R. Elder. Introduction to laser anemometry: Short course notes. Technical report, Cranfield Institute of Technology, 1985.

[11] J.C. Agui and J. Jiminez. On the performance of particle tracking. *J. Fluid Mech*, pages 447–468, 1985.

[12] G. Mie. *Ann. Phys*, 15:377, 1908.

[13] M. Born and E. Wolf. *Principles of Optics. 6th Edition.* Pergamon Press, 656, 1980.

[14] N.G. Jerlov and E.S. Nielson. *Optical aspects of oceanography.* Academic Press, New York, 1974.

[15] R. Barat, J.P. Longwell, A.F. Sarofim, S.P. Smith, and E. Bar-Ziv. Laser Rayleigh scattering for flame thermometry in a toroidal jet stirred combustor. *Appl. Opt*, 30(21):3003–3010, 1991.

[16] R.W. Dibble and R.E. Hollenbach. Non-intrusive temperature measurements in a turbulent diffusion flame. *AIAA-80-1362*, 1980.

[17] A. Corney. *Atomic and laser spectroscopy.* Clarendon Press, Oxford, 1977.

[18] P. Erbland and R. Bogue. Taking the measure of aerodynamic testing. *Aerospace America*, November 1992.

[19] T.D. Dudderar and P.G. Simpkins. Laser speckle photography in a fluid medium. *Nature*, 270:45, 1977.

[20] C. Gray and C.A. Greated. The application of particle image velocimetry to the study of water waves. *Opt. and Lasers in Eng.*, 9:265–276, 1988.

[21] M.P. Arroyo, T. Yonte, M. Quintanilla, and J.M. Saviron. Particle image velocimetry in Rayleigh-Bernard convection: photographs with a high volume of exposures. *Opt. and Lasers*, 9:295–316, 1988.

[22] P.J. Bryanston-Cross, C.E. Towers, T.R. Judge, D.P. Towers, S.P. Harasgama, and S.T. Hopwood. The application of particle image velocimetry (piv) in a short duration transonic annular turbine cascade. *ASME*, 221:1–7, 1991.

[23] C.E. Towers, D.P. Towers, P.J. Bryanston-Cross, and J. Gass T.R. Judge. Particle imaging techniques and applications. *SPIE*, 1553:404–417, 1991.

[24] O.G. Raabe. The generation of aerosols of fine particles. In *Fine Particles*, pages 57–110. New York Academic Press, 1976.

[25] C.N. Davies. *Aerosol science.* Academic Press. London, 1966.

[26] J.B. Moss. Simultaneous measurements of concentration and velocity in an open premixed turbulent flame. *Comb. Sci. and Tech*, 22:115, 1980.

[27] D. Kunii and O. Levenspiel. *Fluidization engineering.* Wiley, New York, 1969.

[28] P.J. Bryanston-Cross *et al.* Particle imaging techniques and applications. *SPIE*, 1553:404–417, 1991.

[29] B.C.R. Ewan. Particle velocity distribution measurement by holography. *Appl. Opt*, 18(18), 1979.

[30] L.M. Lourenco. Velocity bias technique for particle image velocimetry measurements of high-speed flows. *Appl. Opt*, 32(12):2159–2162, 1993.

[31] W. Lauterborn and A. Vogel. Modern optical techniques in fluid mechanics. 16:223–244, 1984.

[32] I. Kimura and T. Toshi. *Image processing of flow around a circular cylinder using correlation technique*, pages 221–226. IV. 1986.

[33] C.E. Towers, D.P. Towers, P.J. Bryanston-Cross, J. Gass, and T.R. Judge. Particle imaging techniques and applications. *SPIE*, 1553:404–417, 1991.

[34] J.M. Coupland, C.J.D. Pickering, and N.A. Halliwell. Particle image velocimetry: theory of directional ambiguity removal using holographic image separation. *Appl. Opt*, 26(9):1576–1578, 1987.

[35] D.H. Barnhart, G.C. Papen, and R.J. Adrian. Holographic particle image velocimetry (HPIV). *SPIE*, 1600:357–365, 1991.

[36] K. Hinsch, H. Hinrichs, G. Kuhfahl, and P. Meinlschmidt. Holographic recording of 3D flow configurations for particle image velocimetry. *ICALEO*, pages 121–130, 1990.

[37] M.P. Arroyo and C.A. Greated. Stereoscopic particle image velocimetry. *Meas. Sci. Tech*, 2:1181–1186, 1991.

[38] P.J. Bryanston-Cross, M. Funes-Gallanzi, C. Quan, and T.R. Judge. Holographic particle image velocimetry. *Opt. and Laser Tech*, 24(5):251–256, 1992.

[39] J.W. Cooley and S.W. Tuckey. An algorithm for the machine calculation of complex Fourier series. *Math. Comput*, 19:297–301, 1965.

[40] J.M. Coupland and N.A. Halliwell. Three-dimensional PIV. In *Optical Methods and Data Processing in Heat and Fluid Flow*, London, 1992. I. Mech. E.

[41] J. Stefanini *et al.* The use of a CCD camera for a coloured PIV technique. In *Optical Methods and Data Processing in Heat and Fluid Flow*, London, 1992. I. Mech. E.

[42] Y. Yeh and H.Z. Cummins. *Appl. Phys. Lett*, 4:176–178, 1964.

[43] J.W. Foreman *et al.* *Appl. Phys. Lett*, 17:77–78, 1965.

[44] L.E. Drain. *The laser Doppler technique.* John Wiley and Sons, 1980.

[45] F. Durst and J.H. Whitelaw. *J. Phys. E*, 4:804–808, 1971.

[46] J.H. Bell, L.C. Rodman, and R.D. Mehta. Aspects of the design and performance of a 3-component LDV system. *ICIASF record*, pages 288–296, 1985.

[47] F.K. Owen, G.M. Orngard, and T.K. Mc. Devitt. A new LDV system for the NASA Ames 6'×6' wind tunnel. *ICIASF Record*, pages 257–262, 1985.

[48] J.B. Abbiss. *Photon correlation spectroscopy and velocimetry.* Plenum Press, New York, 1977.

[49] D.H. Thompson. A tracer particle fluid velocity meter incorporating a laser. *J. Phys. E*, 1:929–932, 1968. Series 2.

[50] L. Lading. Analysis of a laser correlation anemometer. In *3rd Symposium on Turbulence in Liquids*, pages 205–219, University of Missouri-Rolla, 1973.

[51] W.T. Mayo, A.E. Smart, and T.E. Hunt. Laser transit anemometer with microcomputer and special digital electronics: measurements in supersonic flows. In *8th ICIASF annual meeting*, Naval Postgraduate School, Monterey California, 1979.

[52] R.G.W. Brown. Velocimetry signals and data reduction in simulation and practice. In *Photon Correlation Techniques*. Springer-Verlag, New York.

[53] R. Schodl. A laser dual beam method for flow measurement in turbomachines. *ASME*, 74-GT-159, 1974.

[54] R. Schodl, H. Selbach, and H.G. Lossau. Comparison of signal processing by correlation and by pulse-pair timing in laser dual focus velocimetry. In *Symposium on Long Range and Optical Velocity Measurements*. ISL German-French Research Institute, St Louis, France, 1980.

[55] H. Hayami, Y. Senoo, and H. Ueki. Flow in the inducer of a centrifugal compressor measured with a laser velocimeter. *ASME*, 84-GT-74, 1984.

[56] R. Schodl. Laser-two-focus velocimetry for use in aeroengines. *AGARD-LS-90*, pages 4.1–4.34, 1977.

[57] A.E. Smart, D.C. Wisler, and W.T. Mayo. Optical advances in laser transit anemometry. In *Measurement Methods in Rotating Components of Turbomachinery: Proc. of Joint Fluids Engineering Gas Turbine Conference and Products Show*, pages 149–156, New Orleans, Loisiana, 1980. ASME.

[58] R.L. Elder, C.P. Foster, and M.E. Gill. Application of Doppler and transit laser anemometry in small turbomachines. *AGARD*, CP-399:10.1–10.13, 1986.

[59] H.A. Schreiber. Supersonic exit flow measurements downstream of a compressor cascade by the laser two focus method. In *Measuring Techniques in Transonic and Supersonic Cascade Flows*. Central Electricity Research Labs, 1979.

[60] R.G.W. Brown and P.H. Richards. Base flow measurements in a transonic cascade using a laser transit anemometer. In *5th Int. Conf. Kiel-Damp 1982*, pages 192–196. Springer-Verlag, Berlin, 1983.

[61] R.G.W. Brown and P.N. Inman. Direct comparisons of laser Doppler, transit, hot wire anemometer measurements in an axi-symmetric jet. In *10th Int. Cong. on Instrum. in Aerospace Simulation Facilities, ICIASF '83 record*, pages 158–164, Saint Louis, France, 1983. German-French Research Institute (ISL).

[62] R. Schodl. Laser-two-focus velocimetry. *AGARD*, CP-399:7.1–7.30, 1986.

[63] A.Y. Chang, M.D. DiRosa, D.F. Davidson, and R.K. Hanson. Rapid tuning cw laser technique for measurements of gas velocity, temperature, pressure, density, and mass flux using NO. *Appl. Opt*, 30(21):3011–3022, 1991.

[64] P. Desgroux, E. Domingues, and M.J. Cottereau. Measurements of OH concentration in flames at high pressure by two-optical path laser induced fluorescence. *Appl. Opt*, 31(15):2831–2838, 1992.

[65] P. Andresen, A. Bath, W. Gröger, H.W. Lülf, G. Meijer, and J.J. ter Meulen. Laser induced fluorescence with tunable excimer lasers as a possible method for instantaneous temperature field measurements at high pressures: checks with atmospheric flame. *Appl. Opt*, 27(2):365–378, 1988.

[66] M.G. Allen and S.J. Davis. Planar laser-induced fluorescence diagnostics. *ICALEO*, 1990:39–48, 1990.

[67] R. Falco *et al.* Vorticity measurements using laser induced photochemical anemometry (LIPA). *ICALEO*, pages 57–69, 1990.

[68] D. Greenhalgh. CARS thermometry for low and high pressure combustion systems. *AGARD*, CP-399:2.1–2.24, 1986.

[69] D. Klick, K.A. Marco, and L. Rimai. Broadband single pulse CARS spectra in a fired internal combustion engine. *Appl. Opt*, 20(7):1178–1181, 1981.

[70] M. Alden and S. Wallin. CARS experiments in a full scale (10×10m) industrial coal furnace. *Appl. Opt*, 24:3434–3437, 1985.

[71] R. Price. Technical evaluation report, session I - Coherent Anti-Stokes Raman Spectroscopy (CARS). *AGARD*, CP-399:R1:1–R1:6, 1986.

[72] R. Hooke. Of a new property in air. *Micrographia Observation, London*, LVIII:217–219, 1665.

[73] A. Toepler. *Beobachtungen nach einer neuen optischen methode.* Maximilian Cohen und Sohn, Bonn, 1864.

[74] H. Schardin. Das toeplersche schlierenverfahren: Grundlagen für seine anwendung und quantitative auswertung. *Forschungsheft 367, Beilage zu forschung auf dem gebiete des ingenieurwesens, Ausgabe B*, Band 5, 1934.

[75] M. Born and E. Wolf. *Principles of Optics. 6th Edition.* Pergamon Press,122, 1980.

[76] D.W. Holder and R.J. North. A schlieren apparatus giving an image in colour. *Nature*, 169:466, 1952.

[77] G. Settles. Colour schlieren optics: A review of techniques and applications. In *Int. Symp. of Flow Vis*, Bochum FRG, 1980.

[78] D.R. Buchele and D.W. Griffin. Compact color schlieren optical system. *Appl. Opt.*, 32(22):4218–4222, 1993.

[79] R.J. Parker and J.B. Brownell. Application of a copper vapour laser to investigation of shock and blade interaction in a shock tube. In *9th Symposium on Measuring Techniques for Transonic and Supersonic Flows in Cascades and Turbomachines*, Oxford, March 1988.

[80] C.M. Vest. Interferometry of strongly refracting axisymmetric phase objects. *Appl. Opt.*, 14(7):1601–1606, 1975.

[81] T. Lanen. *Digital holographic interferometry in compressible flow research.* PhD thesis, Delft University, 1992.

[82] P.J. Bryanston-Cross. The application of holography as a transonic flow diagnostic to rotating components in turbomachinery. *AGARD*, CP-399:32.1–32.22, 1986.

[83] P.J. Bryanston-Cross. The use of holographic interferometry to visualise edge flow structures at the wheel exit of a turbo-charger. *J. Photographic Science*, 34:77–80, 1986.

[84] P.J. Bryanston-Cross and J.D Denton. Comparison of measured and predicted transonic flow around an aerofoil. *AIAA*, 22(8):1025–1026, 1984.

[85] D.A. Johnson and W.D. Bachalo. Transonic flow past a symmetrical airfoil - inviscid and turbulent flow properties. *AIAA*, 18(1):16–24, 1978.

[86] F.W. Spaid and W.D. Bachalo. Experiments on the flow about a supercritical airfoil including holographic interferometry. *J. Aircraft*, 18(4):287–294, 1981.

[87] W.D. Bachalo and D.A. Johnson. Transonic, turbulent boundary layer separation generated on an axisymmetric flow model. *AIAA*, 24(3):437–443, 1986.

[88] J.E. Craig and C. Allen. Aero-optical turbulent boundary layer / shear layer experiment on the kc-135 aircraft revisited. *Opt. Eng.*, 24(3):446–454, 1985.

[89] G. Havener, M.S. Holden, and D. Azevedo. Holographic visualizations of viscous interactions in hypersonic flows. *SPIE*, 788:26–53, 1987.

[90] A. Abe and K. Takayama. Numerical simulation and density measurement of a shock wave discharged from the open end of a shock tube. *JSME*, 33(2):216–223, 1990.

[91] H. Sugiyama, K. Takayama, T. Shirota, and H. Doi. Shock-wave-induced flow past a circular cylinder in a dusty-gas shock tube. *JSME*, 33(2):224–228, 1990.

[92] R.C. Kim and L.F. Rubin. Diffuse holographic interferometry system for Rocketdyne hypersonic shock tunnel. *SPIE*, 1162:351–359, 1989.

[93] S.P. Sharma and C. Park. Survey of simulation and diagnostic techniques for hypersonic nonequilibrium flows. *J. Thermophysics*, 4(2):129–142, 1990.

[94] V.K.H. Cheung, T.R. Judge, and P.J. Bryanston-Cross. Automatic video inspection of a diesel spray. *SPIE*, 1163:194–219, 1989.

[95] R.B. Lal, J.D. Trolinger, W.R. Wilcox, and R.L. Koes. Holographic flow field analysis in spacelab-3 crystal growth experiments. *SPIE*, 788:62–72, 1987.

[96] R. Snyder and L. Hesselink. Optical tomography for flow visualisation of the density field around a revolving helicopter rotor blade. *Appl. Opt.*, 23(20):3650–3656, 1984.

[97] I.S. Dring, R. Devonshire, J. Meads, H. Boysan, and D.A. Greenhalgh. Observation of non-equilibrium effects in free-convective flows using cars. Technical Report 12660, UKAEA Harwell, 1987.

[98] J.P. Singh and F.Y. Yueh. Multiplex CARS for simultaneous measurement of temperature and co_2 and h_2 concentrations in a combustion environment. *Appl. Opt.*, 30(15):1967–1975, 1991.

[99] D. Gabor. A new microscopic principle. *Nature*, 161:771, 1948.

[100] D. Gabor. Microscopy by reconstructed wavefronts. *Proc. Roy. Soc. (London)*, A197:454, 1949.

[101] D. Gabor. Microscopy by reconstructed wavefronts. *Proc. Roy. Soc. (London)*, B64:449, 1951.

[102] E.N. Leith and J. Upatnieks. Wavefront reconstruction with diffused illumination and three-dimensional objects. *JOSA A*, 54:1295, 1964.

[103] E.N. Leith and J. Upatnieks. Reconstructed wavefronts and communications theory. *JOSA A*, 52:1123, 1962.

[104] E.N. Leith and J. Upatnieks. Wavefront reconstruction with continuous tone objects. *JOSA A*, 53:1377, 1963.

[105] R. Powell and K. Stetson. Interferometric analysis by wavefront reconstruction. *JOSA A*, 55:1593–1598, 1965.

[106] R. Powell and K. Stetson. Interferometric vibration analysis of three-dimensional objects by wavefront reconstruction. *JOSA A*, 55:612, 1965.

[107] R. Powell and K. Stetson. Interferometric hologram evaluation and real-time vibration analysis of diffuse objects. *JOSA A*, 55:1694–1695, 1965.

[108] R.E Brooks, L.O. Heflinger, R.F. Wuerker, and R.O. Brione. Holographic photography of high speed phenomena with conventional and Q-switched ruby lasers. *Appl. Phys. Lett*, 7:92–94, 1965.

[109] P.J. Bryanston-Cross and J. W. Gardner. Holographic visualisation of a combustion flame. *Opt. and Lasers in Eng*, 9:85–100, 1988.

[110] J. Surget. Holographic interferometer for aerodynamic flow analysis. In *Flow Visualization II*, pages 743–747, Bochum, West Germany, September 1980.

[111] T. Uyemura, Y. Yamamoto, K. Tenjimbayashi, and N. Yokoyama. Real-time holographic interferometry with pulsed lasers. *SPIE*, 192:190–195, 1979.

[112] D.R. MacQuigg. Hologram fringe stabilization method. *Appl. Opt.*, 16(2):291–292, 1977.

[113] R.N. Smartt and J. Strong. *JOSA*, 62:737, 1972.

[114] W.D. Bachalo and M.J. Houser. A real-time interferometer technique for compressible flow research. In *Plasma dynamics and laser conference*, pages 1–13, Snowmass, USA, June 1984.

[115] A.J. Decker. Evaluation of diffuse illumination holographic cinematography in a flutter cascade. Technical Report TP-2593, NASA, 1986.

[116] D.P. Towers, T.R. Judge, and P.J. Bryanston-Cross. Analysis of holographic fringe data using the dual reference approach. *Optical Engineering*, 30(4):452–460, 1991.

[117] A.J. Decker. Measurement of fluid properties using double-exposure and time-averaged holographic interferometry. *AIAA*, 84-1461:1–8, 1984.

[118] A.J. Decker. Beam modulation methods in quantitative and flow visualisation holographic interferometry. *AGARD*, CP-399:34.1–34.15, 1986.

[119] W. Koechner. *Solid-state laser engineering.* Topics in Applied Physics, Volume 1. Springer-Verlag, Berlin, 1977.

[120] L. Fang *et al.* The repetitively pulsed argon laser and its application in a hypersonic shock tunnel. In *2nd Int. Symposium on Flow Visualisation*, volume 2, pages 769–775, 1980.

[121] J.E. Craig, G. Lee, and W.D. Bachalo. Nd:YAG holographic interferometer for aerodynamic research. *SPIE*, 353:96–103, 1983.

[122] A.J. Decker. Holographic cinematography of time varying reflecting and time varying phase objects using a Nd:YAG laser. *Opt. Lett*, 7:122–123, 1982.

[123] R.T. Pitlak and R. Page. Pulsed lasers for holographic interferometry. *Opt. Eng*, 24(4):639–644, 1985.

[124] R. Page and R.T. Pitlak. Review of pulsed holography in NDT. *SPIE*, 604:63–70, 1986.

[125] M. Mazakova, M. Pantcheva, G. Spassov, and P. Sharlandjiev. Dichromated gelatin with high sensitivity in the visible region. *SPIE*, 1183:124–127, 1989.

[126] L.T. Blair and L. Solymar. Double exposure planar transmission holograms recorded in nonlinear dichromated gelatin. *Appl. Opt.*, 30(7):775–779, 1991.

[127] R.W. Evans. Transmission holographic optical elements in dichromated gelatin. *SPIE*, 600:38–39, 1985.

[128] H.M. Smith. *Holographic recording materials.* Topics in Applied Physics, Volume 20. Springer-Verlag, Berlin, 1977.

[129] R.T. Ingwall and M. Troll. The mechanism of hologram formation in dmp-128 photopolymer. *SPIE*, 883:94–101, 1988.

[130] A. Hafiz, R. Magnusson, J.S. Bagby, D.R. Wilson, and T.D. Black. Visualisation of aerodynamic flow fields using photorefractive crystals. *Appl. Opt*, 28(8):1521–1524, 1989.

[131] J.A. Leendertz and J.N. Butters. *J. of Meas. and Control*, 4:344–350, 1971.

[132] J.R. Tyrer. Critical review of recent developments in electronic speckle pattern interferometry. *SPIE*, 604:95–111, 1986.

[133] R. Jones and C. Wykes. *Holographic and speckle interferometry.* Cambridge University Press, 1983.

[134] W. Schumann. Fringe localisation in holographic interferometry in the case of a transparent object with a nonuniformly varying index of refraction. *Opt. Lett*, 7(3):119–121, 1982.

[135] C.M. Vest. *Holographic interferometry.* Wiley, New York, 1979.

[136] J.C. Dainty. *Laser speckle and related phenomena*, volume 9 of *Topics in Applied Physics.* Springer Verlag, 1984.

[137] Y. Kawagoe, N. Takai, and T. Asakura. Speckle reduction by a rotating aperture at the Fourier transform plane. *Optics and Lasers in Eng.*, 3:197–218, 1982.

[138] J.D. Trolinger. Conversion of large schlieren systems to holographic visualisation systems. In *15th National Aerodynamics Instrumentation Symposium*, pages 89–96, Las Vegas, 1969.

[139] P.J. Bryanston-Cross, T. Lang, M.L.G. Oldfield, and R.J. Norton. *J. of Eng. for Power*, 103:697, 1981.

[140] W.D. Bachalo. An experimental investigation of supercritical and circulation control airfoils at transonic speeds using holographic interferometry. *AIAA*, 1793:1–17, 1983.

[141] C. Shakher. Use of holographic optical elements in speckle metrology. *Appl. Opt.*, 26(4):654–657, 1987.

[142] B.J. Chang and C.D. Leonard. Dichromated gelatin for the fabrication of holographic optical elements. *Appl. Opt.*, 18(14):2407–2417, 1979.

[143] D. Sheel. Dichromated gelatin for holographic optical elements. *SPIE*, 1212:2–12, 1990.

[144] A.M. Weber, W.K. Smothers, T.J. Trout, and D.J. Mickish. Holograms recorded in Du Pont's new photopolymer materials. *SPIE*, 1212:30–39, 1990.

[145] S.V. Pappu. Holographic optical elements state-of-the-art review: Part 1. *Optical Laser Technology*, 21(5):305–313, 1989.

[146] S.V. Pappu. Holographic optical elements state-of-the-art review: Part 2. *Optical Laser Technology*, 21(6):365–375, 1989.

[147] C.V. Raman and N.S.N. Nath. The diffraction of light by high frequency sound waves. Part 1. *Indian Academy of Science*, A2:402–412, 1935.

[148] R.R.A. Syms. *Practical volume holography*, volume 24 of *Oxford Engineering*. Clarendon Press, 1990.

[149] N. Abramson. Sandwich hologram interferometry: a new dimension in holographic comparison. *Appl. Opt*, 13:2019–2025, 1974.

[150] N. Abramson. Sandwich hologram interferometry 2: some practical calculations. *Appl. Opt*, 14:981–984, 1975.

[151] N. Abramson. Sandwich hologram interferometry 3: contouring. *Appl. Opt*, 15:200–205, 1976.

[152] N. Abramson, H. Bjelkhagen, and P. Skande. Sandwich holography for storing information interferometrically with a high degree of accuracy. *Appl. Opt*, 18:2017–2021, 1979.

[153] G. Havener and R.J. Radley. Quantitative measurements using dual hologram interferometry. Projects 7064 and 7063, United States Airforce, 1972.

[154] J.D. Trolinger, J.E. O' Hare, R.A. Belz, and W.M. Farmer. Holographic flow visualisation. In *AIAA Workshop in Applied Measurement Technology*. University of Tennessee, 1969.

[155] J.D. Trolinger, J. Doty, and H. Umstatter. Dual thermoplastic holography recording system. *SPIE*, 523, 1985.

[156] G. Ballard. Double-exposure holographic interferometry with separate reference beams. *J. Appl. Phys*, 39:4846–4848, 1968.

[157] R. Dändliker, E. Marom, and F.M. Mottier. Two-reference-beam holographic interferometry. *J. Opt. Soc. Am*, 66(1):23–30, 1976.

[158] R. Dändliker, R. Thalmann, and J.F.Willemin. Fringe interpolation by two-reference-beam holographic interferometry: reducing sensitivity to hologram misalignment. *Opt. Comm*, 42(4):301–306, 1982.

[159] G. Lai and T. Yatagai. Dual-reference holographic interferometry with a double pulsed laser. *Appl. Opt*, 27(18):3855–3858, 1988.

[160] G.T. Reid. Automatic fringe pattern analysis: a review. *Optics and Lasers in Engineering*, 24:37–68, 1987.

[161] K. Ramesh and B.R. Pramod. Digital image processing of fringe patterns in photomechanics. *Opt. Eng*, 31(7):1487–1498, 1992.

[162] A. Gillies. Image processing approach to fringe patterns. *Opt. Eng*, 27(10):861–866, 1988.

[163] W.R.J. Funnell. Image processing applied to the iterative analysis of interferometric fringes. *Appl. Opt*, 20:3245, 1981.

[164] T. Yatagai and T. Kanou. *SPIE*, 429:136, 1983.

[165] G.A. Mastin and D.C. Ghiglia. Digital extraction of interference contours. *Appl. Opt*, 24(12):1727–1728, 1985.

[166] D.A. Tichenor and V.P. Masden. Computer analysis of holographic interferograms for non-destructive testing. *Opt. Eng*, 8:469, 1979.

[167] F. Becker, G. Meier, and H. Wegner. Automatic evaluation of interferograms. *SPIE*, 359:386–393, 1982.

[168] D.W. Robinson. Automatic fringe analysis with a computer image processing system. *Appl. Opt*, 22:2169, 1983.

[169] P. Varman. A moiré system for producing numerical data of the profile of a turbine blade using a computer and video store. *Opt. and Lasers in Eng*, 5:41–58, 1984.

[170] S. Krishnaswany. Algorithm for computer tracing of interference fringes. *Appl. Opt*, 30:1624–1628, 1991.

[171] A. Glassman and C. Orr. Automated interferogram reduction. *SPIE*, 181:64–73, 1979.

[172] G. Ben-dor and Y. Arzoan. Fully computerised evaluation of interferograms from fluid flow investigations. *Int. J. Heat and Fluid Flow*, 6:133, 1985.

[173] T. Yatagai. Interferometric testing technology developments and applications in Japan. *SPIE*, 816:58–78, 1987.

[174] J.C. Hunter. *Automatic analysis of holographic interferograms.* PhD thesis, London University, 1987.

[175] D.C. Ghiglia, G.A. Mastin, and L.A. Romero. Cellular-automata method for phase unwrapping. *JOSA A*, 4(1):267–280, 1987.

[176] J.M. Huntley. Noise immune phase unwrapping algorithm. *Appl. Opt*, 28(15):3268–3270, 1989.

[177] T. Judge. A quasi heterodyne holographic technique and automatic algorithms for phase unwrapping. *SPIE*, 1163:95–119, 1989.

[178] M. Takeda and M. Kitoh. Spatio-temporal frequency multiplex heterodyne interferometry. *SPIE*, 1553:66–76, 1992.

[179] L. Mertz. Real-time fringe-pattern analysis. *Appl. Opt*, 22(10):1535–1539, 1983.

[180] R. Dändliker and R. Thalmann. Heterodyne and quasi heterodyne holographic interferometry. *Opt. Eng*, 24(5):824–831, 1985.

[181] R. Thalmann and R. Dändliker. Dimensional profiling by electronic phase measurement. *SPIE*, 746:61–68, 1987.

[182] G.E. Sommargren. *JOSA A*, 65:690, 1975.

[183] K.D. Stumpft. *Opt. Eng*, 18:648, 1979.

[184] F.M. Mottier. *SPIE*, 153:133, 1978.

[185] M.P. Kothiyal and C. Delisle. Rotating analyzer heterodyne interferometer: error analysis. *Appl. Opt*, 24(15):2288, 1985.

[186] J.F. Willemin, S.M. Khanna, and R. Dändliker. Heterodyne interferometer for submicroscopic vibration measurements in the inner ear. *SPIE*, 746:104–112, 1987.

[187] K. Creath. Phase-measurement interferometry: beware these errors. *SPIE*, 1553:213–220, 1991.

[188] J.H. Bruning, D.R. Herriot, J.E. Gallagher, D.P. Rosenfeld, A.D. White, and D.J. Brangaccio. *Appl. Opt*, 13:2693, 1974.

[189] P. Hariharan and B.F. Oreb. Digital phase shifting holographic interferometry. In *Holographic systems, components and applications*, volume 76, pages 1–5, 1987.

[190] F.M. Santoyo, D. kerr, and J. Tyrer. Interferometric fringe analysis using a single phase step technique. *Appl. Opt*, 27(21):4362–4364, 1988.

[191] R.H. Stolen and R. De Paula. *Proc. IEEE*, 75:1498, 1987.

[192] R. Kist and R. Th. Kersten. *Laser und Optoelektronik*, 16(1):17, 1984.

[193] B. Doerband, J. Hertel, and M. Stockmann. Ein rechnergesteuertes twyman-green interferometer. *DgaO-Tagung Vortrag*, A52, 1978.

[194] P. Carré. Installation et utilisation du comparateur photoelectrique et interferentiel du Bureau International des poids et mesures. *Metrologia*, 2(1):13–23, 1966.

[195] M. Takeda, H. Ina, and S. Kobayashi. Fourier transform method of fringe pattern analysis for computer based topography and interferometry. *JOSA A*, 72(1):156–160, 1982.

[196] W.W. Macy Jr. Two-dimensional fringe-pattern analysis. *Appl. Opt*, 22(23):3898–3901, 1983.

[197] R.N. Bracewell. *The Fourier transform and its applications.* Mc. Graw Hill Book Co. New York, 1965.

[198] J.W. Goodman. *Introduction to Fourier optics.* Physical and Quantum Electronics. Mc. Graw Hill Book Co. New York, 1968.

[199] R.J. Green, J.G. Walker, and D.W. Robinson. Investigation of the Fourier transform method of fringe pattern analysis. In *Int. Symp. Loughborough of Fringe Pattern Analysis*, pages 45–76, 1986.

[200] R.H.T. Bates and M.J. Mc Donnell. *Image restoration and reconstruction.* Oxford University Press, 1986.

[201] S.B. Campana. Techniques for evaluating charge coupled images. *Opt. Eng*, 16(3):267–274, 1977.

[202] J.C. Feltz and M.A. Karim. Modulation transfer function of charge coupled devices. *Appl. Opt*, 29(5):717–722, 1990.

[203] T. Judge. PhD thesis, Engineering Department, Warwick University, 1992.

[204] G.M. Brown. Practical phase unwrapping of holographic interferograms. *SPIE*, 1553:204–212, 1991.

[205] N.G. Cooper. From Turing and Von Neumann to the present. Technical Report 9, Los Alamos National Laboratory, 1983.

[206] D.J. Bone. Fourier fringe analysis: the two-dimensional phase unwrapping problem. *Appl. Opt.*, 30(25):3627–3632, 1991.

[207] J.M. Huntley and H. Saldner. Temporal phase-unwrapping algorithm for automated interferogram analysis. *Appl. Opt.*, 32(17):3047–3052, 1993.

[208] T.R. Judge, C. Quan, and P.J. Bryanston-Cross. Holographic deformation measurements by Fourier transform technique with automatic phase unwrapping. *Optical Engineering*, 31(3):533–543, 1992.

[209] J. Gibb. *The cause and cure of periodic flows at transonic speeds.* PhD thesis, College of Aeronautics, Cranfield Institute of Technology, 1988.

[210] R.C. Lock. Test cases for numerical methods in 2D transonic flows. *AGARD*, 575, 1970.

[211] M.G. Hall. Transonic flows. In *Methods and Problems in Aerodynamics and Fluid Mechanics.* 1974.

[212] J.K. Kittleson and Y.H. Yu. Reconstruction of a three-dimensional, transonic rotor flowfield from holographic interferograms. *AIAA*, 25(2), 1986.

[213] S.R. Mohan. *Studies in transonic flows.* PhD thesis, Aerodynamics Dept. Cranfield Institute of Technology, 1983.

[214] I.H. Sneddon. *The use of integral transforms.* Mc Graw Hill, New York, 1972.

[215] R.H. Wright and V.S. Ritchie. Characteristics of a transonic test section with various slot shapes in the Langley 8ft high speed tunnel. Technical Report RM L51H10, NACA, 1951.

[216] D.M. Sykes. The slotted wall transonic wind tunnel at the City University. Technical Report 69/3, Engineering Dept, City University, London, 1969.

[217] H. Kogelnik. Coupled wave theory for thick hologram gratings. *Bell Syst. Tech. J*, 48(9):2909–2947, 1969.

[218] R.K. Kostuk. Effects of bleach constituents on the performance of silver-halide holograms. *SPIE*, 1212:55–62, 1990.

[219] N. Philips and D. Porter. An advance in the processing of holograms. *J. Phys. E*, 9:631, 1976.

[220] D.A. Duston. Optically produced cylindrical hoe's for signal processing applications. *SPIE*, 1212:220–226, 1990.

[221] A. Jameson, W. Schmitt, and E. Turnkel. Numerical solutions of the Euler equations by finite volume methods using Runge Kutta time-stepping schemes. *AIAA*, 81-1259, 1981.

[222] J. Radon. Über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. *Ber. Sächs. Akad. Wiss. Leipzig, Math. Phys. Kl.*, 69:262, 1917.

[223] A.M. Cormack. Representation of a function by its line integrals, with some radiological applications. *J. Appl. Phys.*, 34:2722, 1963.

[224] M.H. Kalos, S.A. Davis, P.S. Mittelman, P. Mastras, and J.H. Hutton. Conceptual development of a vapor volume fraction instrument. Technical Report NY, NDA 2131-34, AT(30-1)-2303(IX), US Atomic Energy Commission, 1961.

[225] R. Gordon. Industrial applications of computed tomography and NMR imaging: an OSA topical meeting. *Appl. Opt.*, 24(23):3969–3972, 1985.

[226] G.N. Hounsfield. A method of, and apparatus for, examination of a body by radiation such as X-ray or gamma radiation. Patent specification, 1972. No. 1283915.

[227] W.P. Rothwell. Nuclear magnetic imaging. *Appl. Opt.*, 24(23):3958–3968, 1985.

[228] R.N. Bracewell. Strip integration in radio astronomy. *Aus. J. Phys.*, 9:198–217, 1956.

[229] W.P. Rothwell and H.J. Vinegar. Petrophysical applications of NMR imaging. *Appl. Opt.*, 24(23):3969–3972, 1985.

[230] S.C. Solomon, P.B. Hays, and V.J. Abreu. Tomography inversion of satellite photometry. Part 2. *Appl. Opt.*, 24(23):4134–4140, 1985.

[231] S. Persson and E. Ostman. Use of computed tomography in NDT of polymeric materials. *Appl. Opt.*, 24(23):4095–4104, 1985.

[232] P.B. Heffernan and R.A. Robb. Difference image reconstruction from a few projections for ND materials inspection. *Appl. Opt.*, 24(23):4095–4104, 1985.

[233] D.J. DeRosier and A. Klug. *Nature*, 217:130–134, 1968.

[234] J.H. Williamson. Computerized tomography for sparse-data plasma physics experiments. *IEEE Trans. on Plasma Sci.*, PS-10(2):82–93, 1982.

[235] R. Snyder and L. Hesselink. High-speed optical tomography for flow visualization. *Appl. Opt.*, 24(23):4046–4051, 1985.

[236] R. Rangayyan, A.P. Dhawan, and R. Gordon. Algorithms for limited-view computed tomography: an annotated bibliography and a challenge. *Appl. Opt.*, 24(23):4000–4012, 1985.

[237] G.W. Faris and R.L. Byer. Three-dimensional beam deflection optical tomography of a supersonic jet. *Appl. Opt.*, 27(24):5202–5212, 1988.

[238] K.E. Bennett, G.W. Faris, and R.L. Byer. Experimental optical fan beam tomography. *Appl. Opt.*, 23(16):2678–2685, 1984.

[239] R.J. Santoro, H.G. Semerjian, P.J. Emmerman, and R. Goulard. Optical tomography for flow field diagnostics. In *AIAA-80-1541*, Snowmass, Colorado, 1980.

[240] G.T. Herman. *Image reconstruction from projections.* Academic Press NY, 1980.

[241] S.R. Dean. *The Radon transform and some of its applications.* John Wiley and Sons. NY, 1983.

[242] R. Gordon and G.T Herman. *Int. Review of Cytology*, chapter Three-dimensional reconstructions from projections: A review of algorithms, pages 111–151. Academic Press NY, 1974.

[243] M. Kalal and K.A. Nugent. Abel inversion using fast Fourier transforms. *Appl. Opt.*, 27(10):1956–1959, 1988.

[244] C.J. Dasch. One-dimensional tomography: a comparison of Abel, onion-peeling, and filtered backprojection methods. *Appl. Opt.*, 31(8):1146–1152, 1992.

[245] J. Landauro and R.G. Jackson. A feasibility study considering the application of tomographic techniques to holographic wind-tunnel data. Technical report, Bolton Institute of Higher Education, 1992.

[246] R.D. Matulka and D.J. Collins. *J. Appl. Phys.*, 42:1109, 1971.

[247] C.M. Vest and P.T. Radulovic. *Applications of Holography and Optical Data Processing*, chapter Measurement of three-dimensional temperature fields by holographic interferometry. Pergamon Press, Oxford, 1977.

[248] D.W. Sweeney and C.M. Vest. Measurement of three-dimensional temperature fields above heated surfaces by holographic interferometry. *Int. J. Heat Mass Trans.*, 17:1443–1454, 1974.

[249] Y.H. Yu and J.K. Kittleson. Reconstruction of a three-dimensional transonic rotor flowfield from holographic interferograms. *AIAA*, 25(2):300–305, 1987.

[250] D. Vukicevic, T. Neger, H. Jäger, J. Woisetschläger, and H. Philipp. Optical tomography by heterodyne holographic interferometry. *SPIE*, IS 8:160–193, 1990.

[251] A. He, D. Yan, and H. Wang. Three-dimensional measurement of asymmetric flow field in hypersonic shock tunnel with moiré interferometry. *SPIE*, 1553:676–681, 1991.

[252] A.C. Kak and M. Slaney. *Principles of computerized tomographic imaging.* IEEE Press NY, 1987.

[253] H. Stark, J.W. Woods, I. Paul, and Hingorani. Direct Fourier reconstruction in computer tomography. *IEEE Trans. Acoust. Speech Signal Processing*, ASSP-29:237–244, 1981.

[254] A.F. Gmitro, J.E. Greivenkamp, W. Swindell, H.H. Barrett, M.Y. Chiu, and S.K. Gordon. Optical computers for reconstructing objects from their X-ray projections. *Opt. Eng.*, 19(3):260–271, 1980.

[255] K.M. Hanson and G.W. Wecksung. Bayesian approach to limited angle reconstruction in computed tomography. *JOSA A*, 73(11):1501–1509, 1983.

[256] R.N. Bracewell and A.C. Riddle. Inversion of fan-beam scans in radio astronomy. *Astrophys. J.*, 150:427–434, 1967.

[257] G.N. Ramachandran and A.V. Lakshminarayanan. Three-dimensional reconstructions from radiographs and electron micrographs: application of convolution instead of Fourier transforms. *Proc. Nat. Acad. Sci.*, 68:2236–2240, 1971.

[258] L.A. Shepp and B.F. Logan. The Fourier reconstruction of a head section. *IEEE Trans. Nucl. Sci.*, NS-21:21–43, 1974.

[259] B.P. Medoff, W.R. Brody, M. Nassi, and A. Macovski. Iterative convolution backprojection algorithms for image reconstruction from limited data. *JOSA A*, pages 1493–1500, November 1983.

[260] A.C. Kak. *Image reconstruction from projections*, chapter Digital Image Processing Techniques, pages 111–170. Academic Press, 1984.

[261] A.C. Kak and C.R. Crawford. Aliasing artifacts in computerised tomography. *Appl. Opt.*, 18(21), 1979.

[262] D.W. Sweeney and C.M. Vest. Reconstruction of three-dimensional refractive index fields from multidirectional interferometric data. *Appl. Opt.*, 12(11):2649–2664, 1973.

[263] D.D. Verhoeven. *Application de la tomographie numérique et de l'interférométrie holographique à l'étude des milieux transparents.* PhD thesis, Institut Francais du Pétrole, France, 1990.

[264] A.J. Decker and S.H. Izen. Three-dimensional computed tomography from interferometric measurements within a narrow cone of views. *Appl. Opt.*, 31(36):7696–7706, 1992.

[265] S. Kaczmarz. Angenaherte auflosung von systemen linearer gleichungen. *Bul. Int. Acad. Pol. Sci. Lett. A*, pages 355–357, 1937.

[266] R. Gordon, R. Bender, and G.T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *J. Theor. Biol.*, 29:471–481, 1970.

[267] R. Gordon. A tutorial on ART. *IEEE Trans. Nucl. Sci.*, NS-21:78–93, 1974.

[268] A. Lent. A convergent algorithm for maximum entropy image restoration, with a medical X-ray application. In *Image Analysis and Evaluation*, pages 249–257. SPIE, 1977.

[269] D.D Verhoeven. MART-type CT algorithms for the reconstruction of multidirectional interferometric data. *SPIE*, 1553:376–387, 1991.

[270] G.T. Herman. Mathematical optimization versus practical performance: a case study based on the maximum entropy criterion in image reconstruction. In *Tenth Symposium on Mathematical Programming*, pages 96–112, Montreal, Quebec, Canada, August 1979.

[271] S. Cha. Interferometric tomography for three-dimensional flow fields via envelope function and orthogonal series decomposition. *Opt. Eng.*, 27:557–563, 1988.

[272] R.M. Lewitt and R.H.T. Bates. Image reconstruction for projections IV: Projection completion methods. *Optik*, 50:269–278, 1978.

[273] R.A. Brooks, G.H. Weiss, and A.J. Talbert. A new approach to interpolation in computed tomography. *J. Comput. Assisted Tomogr.*, 2:577–585, 1978.

[274] S.S. Cha and H. Sun. Interferometric tomography of continuous fields with incomplete projections. *Opt. Lett.*, 14(6):299–301, 1989.

[275] S.S. Cha and H. Sun. Tomography for reconstructing continuous fields from ill-posed multidirectional interferometric data. *Appl. Opt.*, 29(2):251–258, 1989.

[276] E.B. Champagne. Nonparaxial image magnification and aberration properties in holography. *JOSA A*, 57:51, 1967.