Gas and Particle Dynamics in Transdermal Powdered Drug Delivery

Nathan J. Quinlan

The Queen's College, Oxford

Hilary Term, 1999

A thesis submitted for the degree of
Doctor of Philosophy at the University of Oxford

Abstract

Gas and Particle Dynamics in Transdermal Powdered Drug Delivery

A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford

Nathan J. Quinlan • The Queen's College • Hilary Term, 1999

Transdermal powdered drug delivery is an emerging technology for the injection of drugs through human skin, in which particles of solid drug are entrained in a high-speed gas flow and directed towards the skin at a high enough velocity to penetrate the outer layer of dead cells. Hand-held devices based on this idea offer a means of safe, painless and effective delivery of many drugs and vaccines. This thesis describes a programme of research into the fluid dynamics which determine the particle velocity distribution, the most important mechanical characteristic of the system, in prototype drug delivery devices.

Pressure measurements are described which enable characterisation of the gas flow in the drug delivery devices. These are complemented by optical particle detection experiments, which provide a record of the timing of drug particle delivery with respect to the gas flow. Doppler Global Velocimetry (DGV) has been used to measure the velocity field of drug particles. Various tasks involved in the application of DGV to these flow-fields are described. In particular, the use of time-integrated DGV for measurements of unsteady, short-duration flows is discussed. Time-integrated DGV, applied at a range of operating conditions, has provided information on the variation of particle delivery velocity with particle size and with total mass of particles. Time-resolved DGV measurements reveal that particles first emerge in a slow-moving cloud which is driven by transient starting process in the gas flow, followed by a faster stream of particles entrained in a quasi-steady gas flow. The experimental results are complemented by numerical computations of certain aspects of the drug delivery flows. These computations are compared with experimental results, and used to gain additional information on the functioning of the system as a whole.

Acknowledgements

Many people have given invaluable practical and moral support during the time I have spent on this research. Without all of them, it would have been much more difficult, if not impossible, and less enjoyable. Warmest thanks go to my joint supervisors, Roger Ainsworth and Brian Bellhouse. They offered me the opportunity to take part in this unusual collaborative project, and gave their continuous support throughout. Paul Drayson made it possible by offering the financial support of PowderJect. Thanks to John Allen, one of the people who made the Osney lab what it is in every aspect, who crafted instrumentation for many of the experiments described here. Recently, sadly, he passed away before his time. Nigel Brett, John Greenford, Roger Lewis and Gerald Walker built many outstanding rigs and test devices. A substantial part of the work in this thesis would not have been possible without the achievements of Steve Thorpe on DGV. His support, advice and ideas have been helpful in all aspects of this research. Almost everybody who has worked in the Medical Engineering Unit or in Osney over the past few years has given direct practical help at some point, but there a few I would like to thank for their specific and essential contribution. The image processing software which I used was developed by Rob Manners. Mark Roberts carried out pressure measurements which have fed directly into this thesis. Thanks to Mark Kendall and Linda Porter for advice and thought-provoking discussions. Thanks also for the support provided by the staff of PowderJect, particularly Paul Faulkner and Charles Potter. Finally, thanks to everybody who has made this the enjoyable experience that it was. Too many to mention, but they include Maggie, who kept things in perspective and put up with a lot; again, everybody in Osney and the MEU; Jamie and everybody at OUCKC for fine paddling days; Piotr, Evan and all at Queen's; friends at home; and my parents and sisters, distant, but a source of constant support.

Contents

Nomenclature

a	speed of sound *or* constant
A	area
A^*	critical area
b	constant
c	speed of light
c_p	constant pressure specific heat capacity
c_v	constant volume specific heat capacity
C_D	drag coefficient
d	diameter
D	drag
D_h	hydraulic diameter
E	electron energy
f	friction factor, $f = 2\tau_w/\rho u^2$
h	Planck's constant, 6.626×10^{-34} Js *or* heat transfer coefficient
i, j	dummy indices
$\hat{\imath}$	unit vector in direction of incident light
I	image intensity
k	constant light attenuation factor in optical system *or* thermal conductivity
L	length
m	integer index *or* mass
M	Mach number *or* number of time partitions
n	integer index *or* number of particles
N	normalisation image intensity *or* number of particles
N_{fw}	full well electron capacity of a CCD pixel
Nu	Nusselt number
$\hat{o}$	unit vector in direction of observed light
p	pressure
Pr	Prandtl number
Q	heat transfer rate
r	radius of curvature of mirror
R	ideal gas constant or measured image intensity ratio
s	length of measurement region corresponding to one pixel
t	time
T	temperature *or* iodine cell transmission ratio
u	axial velocity

V	volume
$\mathbf{V}$	velocity
x,y,z	Cartesian coordinates

γ	ratio of specific heats = c_p/c_v
ε	root mean square uncertainty
λ	wavelength
μ	viscosity
ν	frequency
ρ	density *or* radius of ring in Fabry-Perot interference pattern
τ	integration period
τ_w	wall shear stress
θ	viewing angle
ξ	non-dimensionalised relative frequency

Subscripts

0	total conditions *or* incident light conditions
a	atmospheric
c	cylinder
d	discriminated image
r	reference image *or* rupture chamber
e	exit
$e+$	supersonic exit condition
$e-$	subsonic exit condition
n	normalisation
p	particle
s	Doppler-shifted *or* surface conditions
t	throat
v	valve

Chapter 1 Introduction

Drug delivery — the transfer of a therapeutic compound to the interior of the body — is an intrinsic element of much medical treatment, and is most familiar to us in the formats of oral ingestion and injection by needle and syringe. Over the past five years, a new technology for the injection of drugs through the skin has emerged. The underlying principle is simple: if particles of a powdered drug are accelerated to sufficiently high velocity in a gas jet and directed towards the skin, they can pass through the outer skin and lodge in deeper layers of tissue. Transdermal powdered drug delivery, as this technique is known, causes no pain or injury, and as described by Bellhouse *et al.* [1], it has been shown to be a viable alternative to injection by needle and syringe for many applications. Moreover, it can enhance the potential of a variety of drugs and vaccines which are limited by available delivery methods.

The subject of this thesis is an investigation into the fluid mechanics of prototype devices which realise this concept of high-velocity particle injection. These prototypes have already been shown to be clinically effective, and are approaching commercial release. Nonetheless, the design and manipulation of the gas-particle jet by which the drug material is delivered to the skin remains a significant engineering challenge. The main objective of the research documented in this thesis is to build a fundamental understanding of the important mechanisms of gas and particle dynamics in this drug delivery technology, through a detailed study of particular prototype devices.

1.1 Transdermal Powdered Drug Delivery

The central idea in this drug delivery technique is the use of high-speed microscopic projectiles as a medium for the introduction of foreign material into biological tissue. This concept originated in the 1990 invention by Sanford *et al.* [2] of a so-called biolistic gene gun. In this device, DNA-coated gold or tungsten beads of 1 – 3 μm in diameter were attached to the front face of a plastic bullet which was accelerated along a tube by explosive means. The bullet would sharply come to a standstill on buffers at the end of the tube, catapulting the DNA-bearing particles onwards into the target plant or mammalian tissue. The particles came to rest inside cells in the target, so that the foreign DNA could be incorporated to transform the cell genetically. A variant gene gun was reported by Sarphie [3], a researcher in the Medical Engineering Unit at Oxford University. In the Oxford device, DNA-coated gold particles were entrained and accelerated in a high-speed flow of helium which was directed onto the biological target. The gas was deflected from the surface of the target tissue, but the particles had sufficient momentum to penetrate the tissue. Using this device, Sarphie achieved successful expression of a delivered gene in plant seeds, leaves and roots.

In 1993, Bellhouse *et al.* [4] patented a development of the gas gene gun for the delivery of drugs in dry powdered form into human skin. Here, the aim was not to deliver drug directly into cells, but to deliver particles into the regions between cells for subsequent dispersion of the drug by dissolution and diffusion. To compensate for the reduced density of typical drugs (in comparison with the metal carrier particles used for DNA delivery), particle size was increased to 10 – 50 μm. Drug particles of this size were found to acquire sufficient momentum in the gas flow to penetrate the outer layers of the skin, and caused no pain or bleeding.

The technology of transdermal powdered drug delivery has emerged from that early research, and is being developed for commercial application [5] with the support of ongoing research within Oxford University. An integrated overview of the technology is given by Bellhouse *et al.* [1], covering engineering, pharmacological and medical facets of the research and development work to date. The general configuration of a current device for transdermal powdered drug delivery is shown in ***Figure 1.1***. The key components of the device are a gas cylinder, in which compressed helium is stored, typically at a pressure of tens of atmospheres; a drug cassette, in which the powdered drug is retained between a pair of plastic membranes; and a convergent-divergent nozzle. Basic compressible flow theory provides a qualitative description of this system's expected mode of operation. When gas is released from the cylinder, a large pressure difference builds up across the drug cassette. The membranes burst, leading to the formation of a shock wave, which propagates down the nozzle and initiates an unsteady high-speed gas flow, as in a classical shock tube. Later, a sustained bulk flow of gas from the cylinder is established, and under certain conditions, the device's convergent-divergent nozzle functions as a supersonic nozzle. In the course of these processes, particles are entrained in the gas flow and accelerate towards the nozzle exit. As the particle-laden flow impinges on the skin, gas is deflected away to the side and vents to the atmosphere through the silencer. The particles, with their relatively large inertia, maintain a high axial velocity and enter the skin. They penetrate the tough outer layer of dead cells, the stratum corneum. Then they can either act locally, or diffuse into the bloodstream in the dermis, an inner layer of skin which is rich in capillary vessels. The devices, and the basic theory of their operation, are described more fully in Chapter 2.

[redacted]

Figure 1.1 A diagram of a device for transdermal delivery of dry powder drugs, reproduced from PowderJect [5].

This technology is one recent development in the growth of drug delivery over the past 20 years as an important aspect of pharmaceutical science, which has been highlighted in reviews by Ganderton [6] and Ranade and Hollinger [7]. This trend has reflected an increasing awareness that the means by which a drug is delivered into the body and/or to a particular target site, in addition to its essential chemical properties, can be an important contributor to the drug's overall value. For example, sophisticated chemical techniques and transdermal patches have both received attention as methods of achieving sustained, controlled release of a drug into the body over a prolonged period. Liquid jet injectors and various refined or semi-automated needle devices, as well as the transdermal powder method described here, are among alternative mechanical approaches.

Transdermal powder drug delivery has many advantages. Most obviously, it is painless and causes no injury, unlike liquid jet injectors and the time-honoured needle and syringe. It does not leave dangerous sharps for disposal, create a contamination hazard, or require a skilled operator. It shares some of the advantages of the needle and syringe in offering rapid action and localised delivery, when required. The direct use of drugs in powder (rather than liquid) form is attractive in itself, since drugs are often manufactured and distributed as powder for compact storage and improved chemical stability, and must be reformulated as a liquid prior to use in a needle and syringe.

In addition to these aesthetic and logistical advantages, transdermal powder delivery can facilitate and enhance the use of new and existing drugs. DNA vaccines, for example, offer a powerful and general new approach to vaccination, by which immune response can be induced without introducing a modified disease-producing organism into the body. They are currently the subject of extensive research and investment. However, DNA vaccines are most effective when delivered into the epidermis, the skin layer immediately beneath the outer layer of dead cells, which is typically 50 μm thick and therefore difficult to target by conventional means. Similarly, some conventional vaccines and anti-viral drugs benefit from delivery to a thin layer called the stratum basale, underneath the epidermis. The new technology offers the potential to target these specific layers through control of drug particle momentum (and hence, penetration depth), or at least to achieve blanket delivery into the shallow regions of the skin. Drugs which require self-administration, or local delivery to sensitive sites such as the genitals, can be made more attractive and accessible for the patient with this technology.

The first commercially available version of the device will be used to administer lidocaine (a local anaesthetic), with the purpose of numbing the site where a major intravenous line (typically for a general anaesthetic) is to be inserted. Clinical trials of this application have been highly successful [1], and preliminary clinical trials with alprostadil (a treatment for male erectile dysfunction) have been promising. Many other applications are planned or in development, including insulin for diabetes sufferers, and treatments for osteoporosis, migraine, melanoma and anaemia. Successful results have been achieved in pre-clinical trials to support some of these developments. It is anticipated that powder injection devices for a variety of drugs and vaccines will become a familiar sight in clinics and homes within 5 to 10 years.

1.2 Motivation and Objectives

The idea of transdermal delivery of drugs in dry powder form, as described in the previous section, is an elegant and simple concept. However, its successful implementation is not trivial. Success, in this context, is defined by several groups of criteria. First and foremost are the medical criteria specific to each application of the system to a particular treatment — the device must fulfil its primary purpose by enabling the drug or vaccine of interest to produce the therapeutic effects of which it is capable. The efficiency of drug delivery, defined in terms of the fraction of drug loaded into the device which becomes available in the body for biological action, must be maximised. In particular, the release of potentially harmful residual powder, whether vented to the atmosphere or deposited on the skin or in the device, should be avoided. The capability to adjust the device to suit different skin types or deliver different quantities of drug is desirable. Aesthetic, ergonomic and economic properties are secondary, but important, since comfort and ease of use of the technology are a large part of its appeal. The device must be acceptably compact, simple, inexpensive, silent, and comfortable for the operator and the patient.

The development of the technology towards these goals is a challenging multi-disciplinary project, which draws on fluid dynamics, particle mechanics, mechanical design and manufacturing engineering, as well as various fields in biology and chemistry. A prerequisite for advancement of the technology is a thorough understanding of both the mechanical and biological processes which are involved in drug delivery. This thesis is concerned solely with mechanical aspects of transdermal powdered drug delivery. The most important mechanical attribute of the system is the distribution (in space and time) of particle velocity on arrival at the skin, since it determines the penetration of the particles into the skin, and their final distribution within the skin. For the purposes of this research, therefore, the drug delivery device can be viewed purely as a means for producing a stream of particles at some high velocity.

The primary goal of this research, then, is a fundamental understanding of the processes which govern particle acceleration and determine particle velocity distribution. The main route towards this objective is the measurement and description of the particle velocity distribution achieved by prototype drug delivery devices. This measurement of particle delivery velocity is in itself an objective of this research, but its greatest importance is as a source of insight into the fluid mechanics. The development of numerical methods for device performance prediction is a third objective. Such models are an exacting test of the physical understanding of device operation, and a useful aid to the exploration of potential new device designs.

A parallel objective relates to the experimental technique used to measure particle velocity. This method is Doppler Global Velocimetry (DGV), a powerful new laser-based tool for non-intrusive whole-field velocity measurement in fluid dynamics, which has been reviewed by Meyers [8] and more recently by Ainsworth *et al.* [9]. The latter authors, members of the Department of Engineering Science in Oxford University, have been at the forefront of the development of DGV. One of the aims of the current research has been to build on existing experience in DGV by applying it in a new application area, and, wherever possible, to contribute to research into the technique itself.

1.3 Methodology

The main thrust of this research has been experimental, largely because of the opportunity to utilise the DGV facility in Oxford. In any case, an entirely computational investigation would have been questionable in the absence of any directly relevant experimental data for guidance and validation. The experiments described in this thesis have consisted of measurements of the flows generated by two types of prototype device discharging a gas or gas-particle jet into the atmosphere. One of the device types tested was an early design which displayed clearly identifiable regimes of classical gas dynamic behaviour, and provided a useful arena in which to characterise the overall structure of the gas flow. Two variants of this type, with similar overall configurations but slightly different nozzle geometries, were tested. The majority of the research effort was concentrated on a more recent device with a substantially different nozzle which more closely resembles the likely design of the first production devices, and which presented greater challenges in interpretation of experimental results. A detailed description of the test devices is given in Chapter 2, along with some theoretical background to the fluid dynamics of the drug delivery devices.

The experiments carried out are summarised in a schematic diagram presented as ***Figure 1.2***. Measurements of unsteady wall pressure were carried out at up to three locations in the device using wall-mounted transducers. These experiments yielded information on the flow of gas in the device, both with and without particles. Pressure measurements were complemented by optical particle detection experiments, in which particle concentration in the flow from the device exit was recorded by a simple non-quantitative technique. The aims of these experiments were to examine the time scales of particle delivery and to identify the periods in the gas flow which are important to particle delivery. Pressure measurements and optical particle detection are classed together as point experiments. These methods and their results are documented in Chapters 3 and 4, respectively.

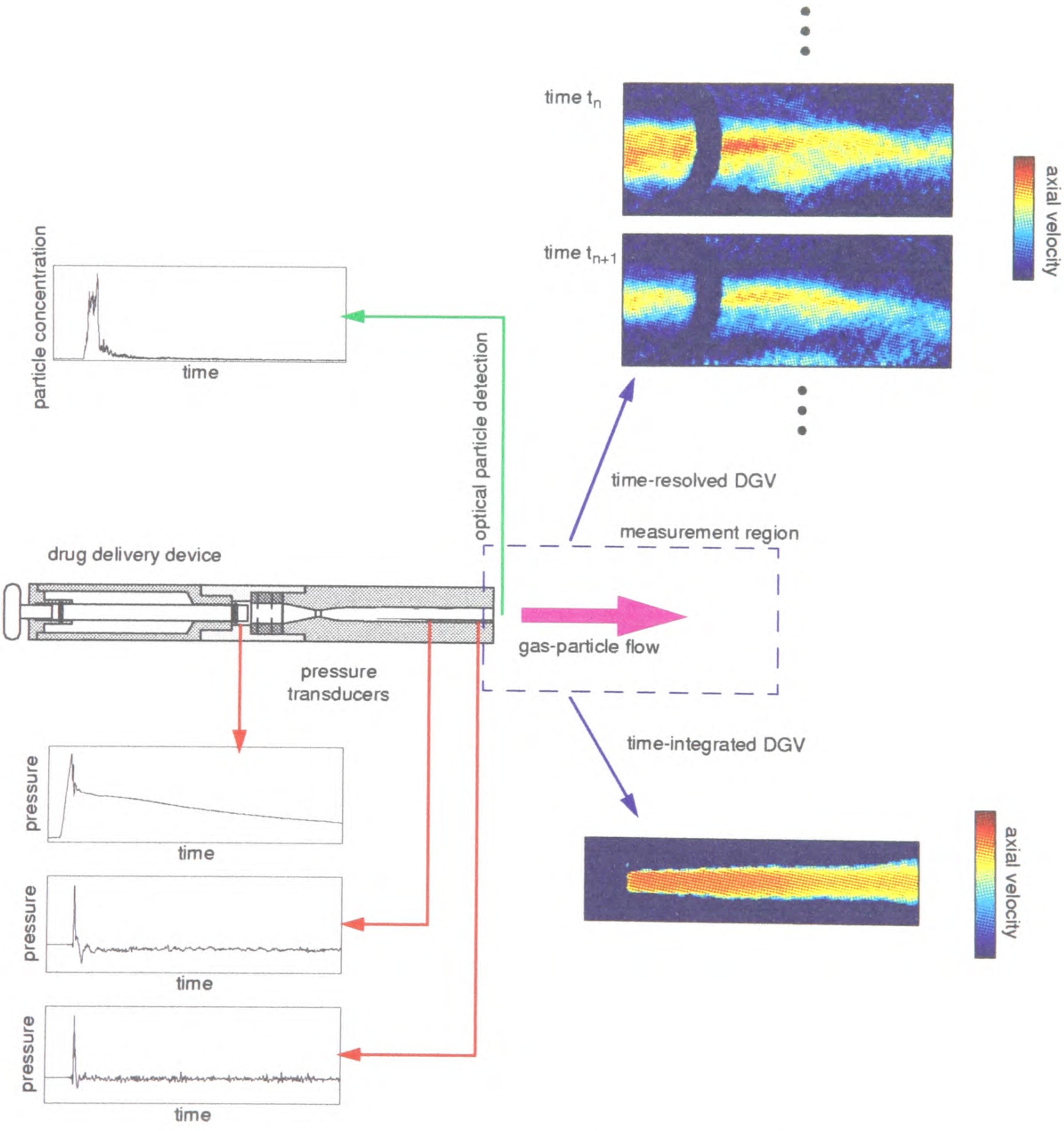


Figure 1.2 A schematic diagram illustrating the various experimental techniques used in the course of this research.

The DGV experiments delivered two-dimensional distributions of axial particle velocity on a plane containing the nozzle centreline. In some cases, the measurement region extended downstream from the nozzle exit, while in others, velocity data inside the nozzle were obtained. Two versions of the technique were used. In time-integrated experiments, the velocity distributions obtained were weighted averages over the entire lifetime of a drug delivery flow-field. With time-resolved DGV, measurements of instantaneous velocity were made, enabling a time history of the unsteady particle velocity field to be assembled. DGV methods and their application in this research are discussed in Chapter 5, and DGV results are presented in Chapter 6.

The task of modelling the fluid dynamics of the drug delivery system was directed towards the development of relatively simple numerical prediction codes which would make light demands on computational resources, and produce (at least) qualitatively realistic and informative predictions. This work was intended primarily as a means of adding to understanding of the flow-fields of interest, through the exercise of developing the models as well as through the analysis of their output. The development and application of the numerical models is described in Chapter 7.

Chapter 2

Devices for Transdermal Powdered Drug Delivery

This chapter fills in some detail in the background to the current research by expanding at several levels on the basic description of the drug delivery system given in Chapter 1. In §2.1, the classical theories of gas dynamics are brought to bear to describe the basic kinds of behaviour to be expected from the device. This exercise is the basis of a simple but informative picture of compressible flow phenomena in the drug delivery device, and provides terminology and reference points, in the form of well-understood idealised flows, which are useful throughout this research. Some of the complex fluid flow phenomena which are important in the drug delivery device, though beyond the reach of basic theoretical approaches, have been investigated before in other engineering applications. The literature dealing with these areas is reviewed in §2.2. Finally, the drug delivery devices and drug particles which were specifically studied in the course of this research are described in detail in §2.3.

2.1 Basic Theoretical Descriptions

The rationale for the basic configuration of the drug delivery device — the combination of a high-pressure reservoir, bursting membranes and a convergent-divergent nozzle — stems from well-established theory in gas dynamics. In this section, some important results from this body of theory are presented to support and extend the outline qualitative descriptions of expected flow in the device which have already been given. The theoretical concepts, equations and terms of reference introduced in this section are an important element of the background to discussions throughout this thesis.

Two classes of gas flow which are widely studied and well understood (at least in idealised forms) are of value in the study of these drug delivery devices. They are shock tube flow, which is dealt with in §2.1.1, and convergent-divergent nozzle flow, discussed in §2.1.2. Mathematical descriptions of these flows are presented here without proof. They can be derived through analyses which are given by Anderson [10], Shapiro [11] and many other authors.

2.1.1 Shock Tube Flow

The classical shock tube, depicted in ***Figure 2.1***, consists of a constant-area tube with a membrane which separates regions of differing gas properties. When the membrane is removed suddenly (in practice, by bursting), a shock wave propagates into the low-pressure or *driven* gas and an expansion wave propagates into the high-pressure *driver* gas. A bulk flow of gas follows the shock at a lower speed, and includes the *contact surface*, the boundary between the driver and driven gases.

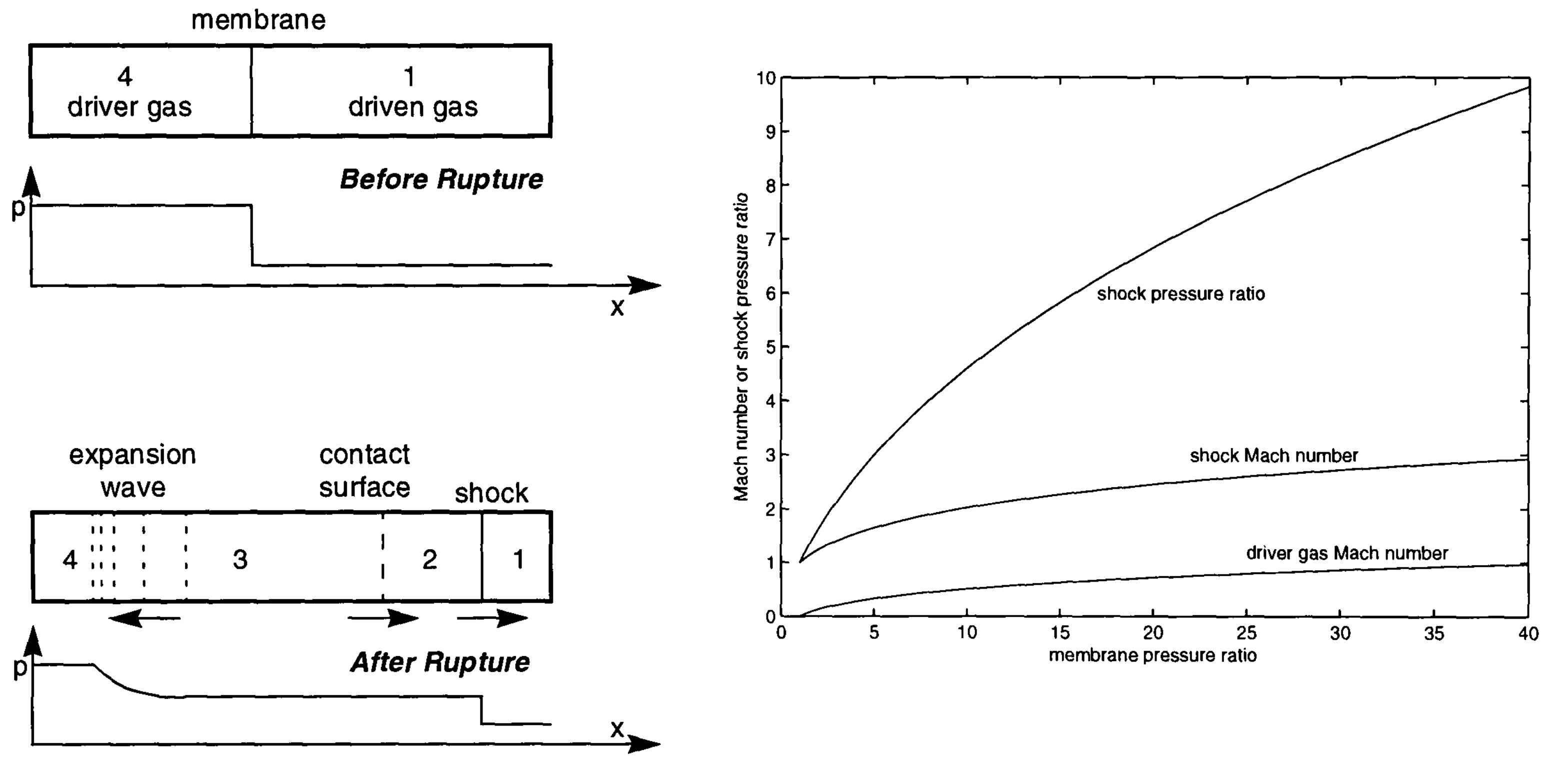


Figure 2.1 Schematic diagram of wave structures in a shock tube, and the variation of flow parameters as functions of the membrane pressure ratio for helium driving air.

The theoretical dependence of various flow parameters on the initial membrane pressure ratio is plotted in ***Figure 2.1*** for helium driving air at equal initial temperatures. The pressure ratio across the shock, p_2/p_1, is related to the initial pressure ratio across the membrane, p_4/p_1, by:

$$\frac{p_4}{p_1} = \frac{p_2}{p_1}\left(1 - \frac{(\gamma_4 - 1)\frac{a_1}{a_4}\left(\frac{p_2}{p_1} - 1\right)}{\sqrt{2\gamma_1\left[2\gamma_1 + (\gamma_1 + 1)\left(\frac{p_2}{p_1} - 1\right)\right]}}\right)^{\frac{-2\gamma_4}{\gamma_4 - 1}}, \qquad (2.1)$$

where the subscript 1 denotes conditions in the undisturbed driven gas, 4 denotes the driver gas and 2 denotes accelerated driven gas between the shock and the contact surface, as indicated in ***Figure 2.1***. In other equations, the subscript 3 will refer to the region of accelerated driver gas between the contact surface and the expansion wave. a is the speed of sound and γ is the ratio of constant-pressure specific heat to constant-volume specific heat. The Mach number of the shock wave, M_s, and the

Mach number of flow behind the contact surface, M_3, are given as functions of shock pressure ratio by ***Eq. 2.2*** and ***Eq. 2.3***, respectively.

$$M_s = \sqrt{\frac{\gamma_1 + 1}{2\gamma_1}\left(\frac{p_2}{p_1} - 1\right) + 1}\ . \tag{2.2}$$

$$M_3 = \frac{2}{\gamma_{4} - 1}\left(\frac{p_2}{p_4}\right)^{\frac{-(\gamma_4 - 1)}{2\gamma_4}}\left[1 - \left(\frac{p_2}{p_4}\right)^{\frac{\gamma_4 - 1}{2\gamma_4}}\right]\ . \tag{2.3}$$

The drug delivery device resembles a shock tube in the very early stages of flow, just after the downstream membrane has burst. A shock propagates through the nozzle, compressing and accelerating the air through which it passes, and eventually emerges from the nozzle exit in some modified state, followed by a transient flow in which the contact surface is embedded. This shock wave will be referred to as the *starting* or *primary* shock. An expansion wave propagates upstream into the rupture chamber, which at first contains a driver gas consisting of air mixed with helium which has flowed in from the open cylinder valve. The expansion wave causes the driver gas to accelerate from rest, through the drug cassette (accelerating and entraining drug particles) and into the nozzle. The analogy to the simple shock tube is weakened as the starting shock encounters varying duct area in the nozzle, and the expansion wave is reflected off the upstream wall of the rupture chamber. Nonetheless, classical shock tube theory offers a useful framework for discussion of the early stages of flow in the drug delivery device, as well as a realistic description of the events immediately following membrane rupture.

2.1.2 Convergent-Divergent Nozzle Flow

The convergent-divergent (or Laval) nozzle appears in many engineering applications, including rocket nozzles, guide vanes in gas turbines, supersonic wind tunnels and gas dynamic lasers. The textbook analysis of the convergent-divergent nozzle rests on the theory of quasi-one-dimensional isentropic flow and normal shock waves. The key points of this theory will be outlined here.

Static pressure p in any compressible flow is related to total pressure, p_0, and Mach number, M, by ***Eq. 2.4*** below.

$$\frac{p_0}{p} = \left(1 + \frac{\gamma - 1}{2}M^2\right)^{\frac{\gamma}{\gamma - 1}} \tag{2.4}$$

The total pressure can be defined as the pressure to which the gas would be compressed, hypothetically, if the flow was isentropically decelerated to rest. Total pressure is therefore uniform throughout an isentropic flow. Similarly, total temperature is defined as the temperature at which the gas would come to rest if it was decelerated isentropically. It is related to static temperature by ***Eq. 2.5***.

$$\frac{T_0}{T} = 1 + \frac{\gamma - 1}{2}M^2 \tag{2.5}$$

T_0 is also constant in an isentropic flow; more generally, it is constant if the fluid does not exchange heat or work with its surroundings. In the context of an isentropic convergent-divergent nozzle flow, it is useful to imagine total temperature and total pressure as the static temperature and pressure in a large stagnant reservoir, from which gas flows isentropically through the nozzle.

The Mach number at any location in the nozzle is related to the local nozzle cross-sectional area by:

$$\left(\frac{A}{A^*}\right)^2 = \frac{1}{M^2}\left[\frac{2}{\gamma + 1}\left(1 + \frac{\gamma - 1}{2}M^2\right)\right]^{\frac{\gamma + 1}{\gamma - 1}} . \tag{2.6}$$

In this equation, A^* is the cross-sectional area required to (hypothetically) accelerate or decelerate the flow isentropically to the speed of sound. In a truly isentropic flow, it corresponds to the physical section area at which M=1. It can be shown that A/A^* has a minimum value of 1, occurring at M=1. The well-known implications of this are that a steady isentropic flow cannot pass through an area smaller than that at which M=1, and that for given total pressure and total temperature, the mass flow rate through a nozzle is at its maximum, or *choked*, when M=1 at the smallest section, the *throat*. Also, for any area ratio, ***Eq. 2.6*** is satisfied by a subsonic and a supersonic Mach number. Therefore, it is possible for a gas to accelerate through a contracting duct, reaching the speed of sound at the narrowest point, and then accelerate beyond $M = 1$ through an expanding duct. This is the principle underlying a convergent-divergent nozzle.

With this information, it is possible to predict the behaviour of steady flow through the transdermal powdered drug delivery nozzle. Some of the flow regimes discussed below are illustrated for an imaginary nozzle in ***Figure 2.2***. If pressure outside the nozzle exit is fixed at atmospheric pressure, p_a, the nozzle flow is determined by the total pressure, p_0. At very low values of the pressure ratio p_0/p_a, flow is subsonic throughout the nozzle, with a maximum velocity at the throat. This is depicted in ***Figure 2.2*** as case *A*. Mach number throughout the nozzle rises with p_0 until the throat Mach number is 1 (case *B*). The pressure ratio at this condition can be calculated by requiring that static pressure in the subsonic flow at exit equals atmospheric pressure. From ***Eq. 2.4***, this implies:

$$\left(\frac{p_0}{p_a}\right)_1 = \left(1+\frac{\gamma-1}{2}M_{e-}^2\right)^{\frac{\gamma}{\gamma-1}} . \qquad (2.7)$$

Here, M_{e-} is the exit Mach number in subsonic isentropic flow through the nozzle when flow in the throat is sonic. It is one of the two solutions to ***Eq. 2.6*** for area ratio $A/A^* = A_e/A_t$, the ratio of exit to throat areas.

As pressure ratio is pushed above this threshold, Mach number in the throat cannot rise further. Instead, the flow accelerates through a zone of supersonic flow downstream of the throat, which terminates in a standing normal shock wave (*C, D* and *E*). The shock initially stands at the throat and is infinitesimally weak, but it moves downstream and strengthens as the pressure ratio is raised. In the limit, it stands at the nozzle exit; the Mach number just upstream of the shock is then M_{e+}, the supersonic solution to ***Eq. 2.6***, and the nozzle pressure ratio is given by ***Eq. 2.8***.

$$\left(\frac{p_0}{p_a}\right)_2 = \frac{\left(1+\frac{\gamma-1}{2}M_{e+}^2\right)^{\frac{\gamma}{\gamma-1}}}{1+\frac{2\gamma}{\gamma+1}(M_{e+}^2-1)} \qquad (2.8)$$

This is the minimum pressure ratio for fully supersonic flow in the divergent part of the nozzle. The denominator in the right hand side of ***Eq. 2.8*** represents the pressure ratio across the exit shock wave, and can be derived from ***Eq. 2.2***. With further increases in nozzle pressure ratio p_0/p_a, the shock pressure ratio must be reduced to match the flow pressure downstream of the shock to atmospheric pressure. To achieve this, the normal shock opens out into a pair of oblique shocks, extending from the lip of the nozzle, and interacting with each other in a complex jet structure. This state (illustrated as *F*) is described as *overexpanded*, and it remains in effect until the pressure just upstream of the shocks equals atmospheric pressure, at which point the shocks vanish. At this threshold condition, the jet from the nozzle exit is *neutrally expanded*, supersonic and shock-free, and the pressure ratio is given by another variation of ***Eq. 2.4***:

$$\left(\frac{p_0}{p_a}\right)_3 = \left(1 + \frac{\gamma - 1}{2} M_{e+}^2\right)^{\frac{\gamma}{\gamma - 1}} . \qquad (2.9)$$

For higher total pressures, the jet from the nozzle exit is *underexpanded*, with exit pressure higher than atmospheric pressure, and expansion waves are required to drop the exit pressure to the atmospheric level.

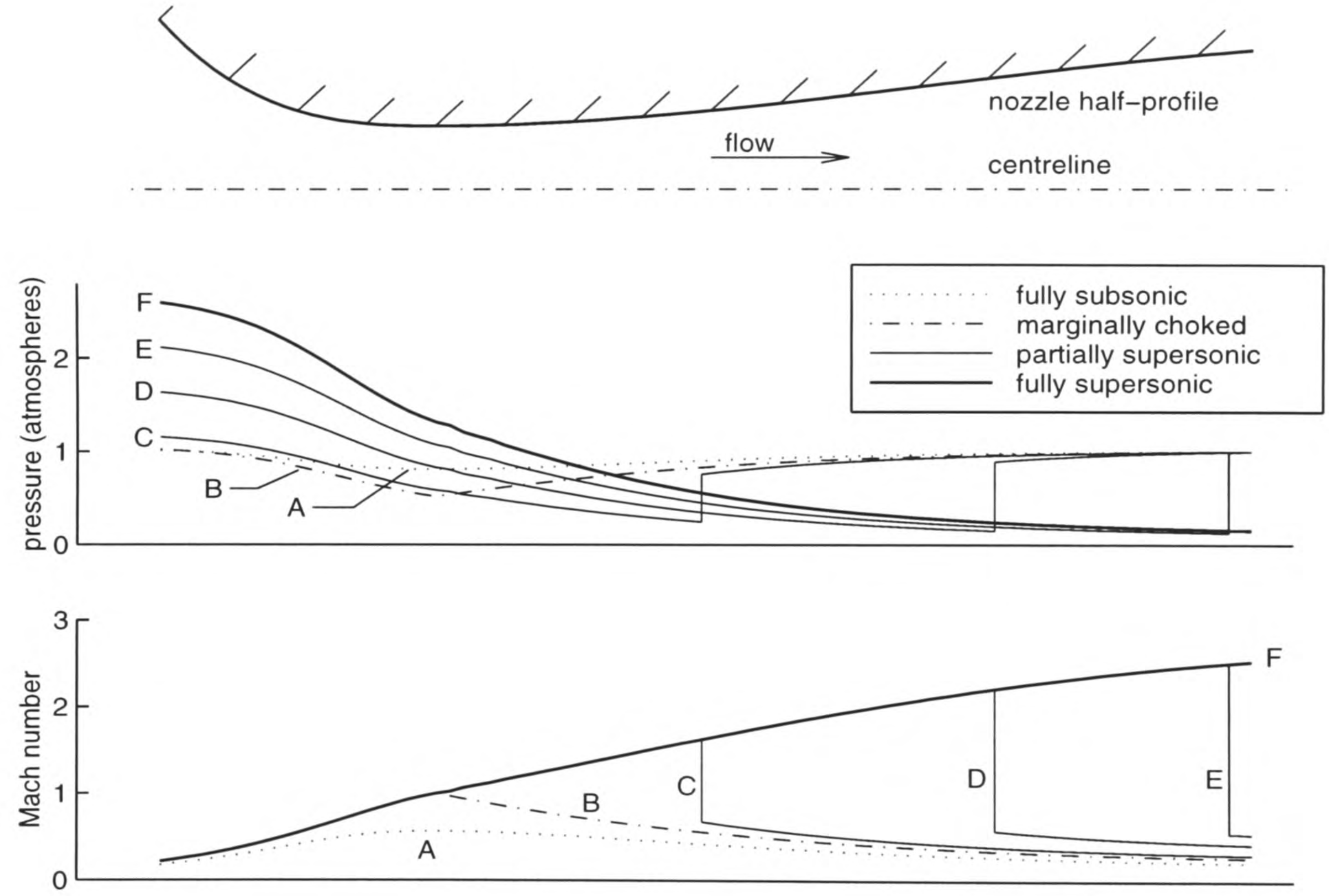


Figure 2.2 Nozzle geometry, pressure profiles and Mach number profiles in a generic convergent-divergent nozzle with constant back pressure of 1 atmosphere and varying total pressure. Curves are labelled alphabetically in order of increasing total pressure.

Figure 2.3 is a map of nozzle design (represented by M_{e+}) and operating condition (represented by p_0/p_a), in which the various flow regimes are summarised. The boundaries between the regimes are defined by the critical pressure ratios specified in ***Eqs. 2.7*** to ***2.9***. The chart shows that the pressure ratio required for shock-free supersonic flow increases very rapidly with exit Mach number. On the other hand, the pressure required to choke the throat falls off with increasing Mach number, and tends to 1. This is because a nozzle with high supersonic exit Mach number has a low subsonic exit Mach number, allowing a substantial pressure rise from throat to exit in a marginally choked flow.

Interestingly, there is a minimum pressure ratio at which supersonic flow to the exit can be achieved. From ***Eq. 2.8***, it can be shown that this minimum is 1.58 (for helium, $\gamma = 5/3$), and is obtained in a nozzle with an exit Mach number of 1.53.

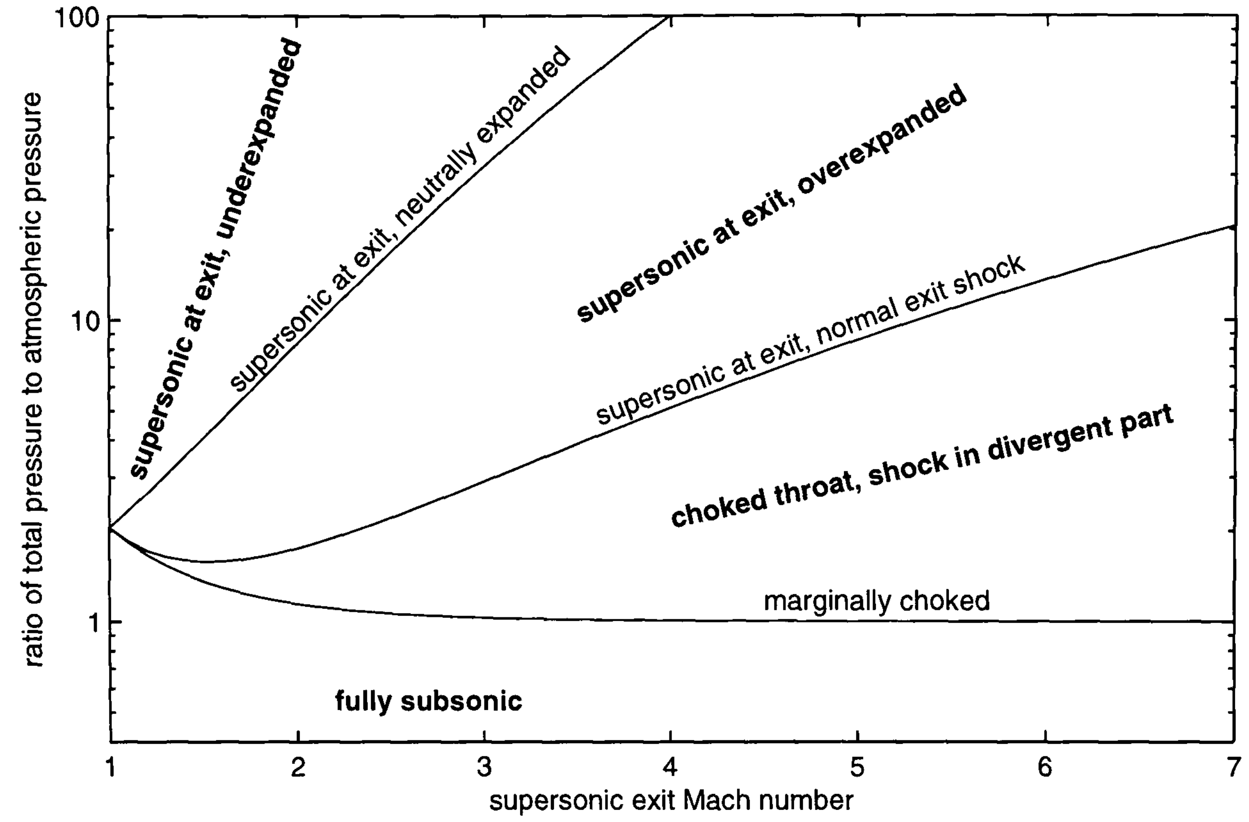


Figure 2.3 A map of possible steady nozzle flow regimes, expressed in terms of supersonic exit Mach number and the ratio of total pressure to atmospheric pressure.

The drug delivery system has much in common with the idealised convergent-divergent nozzle. After the initial pressure difference across the bursting membrane is resolved through the transient shock tube-like process, a relatively steady flow through the nozzle is expected, driven by the large pressure ratio between the cylinder and the atmosphere. The pressure in the cylinder, which decays as the cylinder drains, is closely linked to the total pressure which determines the nozzle flow structure. In fact, the reservoir system, consisting of the combined cylinder and rupture chamber, may be closely analogous to the ideal of a stagnant reservoir feeding an isentropic flow. The major discrepancy between the real nozzle and the theoretical nozzle is viscosity, which, to some extent, will cause a departure from isentropic flow in the real nozzle. In an extreme scenario, boundary layers could have a significant impact on the large-scale structure of the flow, particularly through their interaction with shock waves, as discussed in §2.2.2.

2.2 Survey of Literature on Relevant Aspects of Fluid Dynamics

The drug delivery devices studied in this research, characterised by a short-duration, strongly unsteady flow in which a finite particle load is entrained, appear to be unique among engineering systems. This thesis represents the first substantial investigation into the mechanics of these devices, and consequently, there is very little literature which deals specifically with the problems approached in the current research. However, there is a substantial body of work on various relevant topics in fluid dynamics, which is reviewed in this section. To begin, §2.2.1 outlines the findings of an investigation into a precursor of the current drug delivery technology.

2.2.1 A Gas-Particle Gene Delivery Device

The concept of transdermal injection of powdered drug emerged from the technology of biolistic gene delivery, a method for the insertion of genetic material into living cells by bombardment with DNA-coated microprojectiles at high velocity. The history of this technique was outlined in Chapter 1. Dr. David Sarphie [3] of the Department of Engineering Science in Oxford University developed and investigated the biolistic gene gun which was the immediate predecessor to the first transdermal powdered drug delivery devices. The system was similar in overall configuration to the current drug delivery devices, but larger by a factor of about 5. The use of a gas flow to impart momentum to the DNA-coated metallic beads was a novel development, motivated by the desire for improvements in repeatability and control over existing methods. Sarphie reported biological testing and a range of aerodynamic measurements of the Oxford system, including measurements of the average velocity of the particle cloud by an optical time-of-flight technique. Results showed that velocity increased with upstream total pressure. The measured interval between the starting shock and the beginning of particle delivery, and the measured particle velocities, indicated that the initial transient flow, rather than quasi-steady nozzle flow, was the dominant flow mechanism in particle acceleration.

2.2.2 Convergent-Divergent Nozzle Flow

§2.1.2 illustrates the power and simplicity of quasi-one-dimensional isentropic flow theory and normal shock theory in modelling flow in a convergent-divergent nozzle. However, when nozzle operating conditions call for a shock to compress the exit flow, the simple theories can be unrealistic.

Shapiro [11, p. 135] makes the general observation that "when... a compression shock appears in the duct, the latter is found to be radically different from a plane discontinuity. This difference arises as the result of an interaction between the shock wave and the boundary layer on the walls of the duct". The departure of nozzle flow from the quasi-one-dimensional inviscid ideal is the main topic of this section.

From among the many experimental illustrations in Shapiro's text which support the statement above, a schlieren photograph of flow in a plane convergent-divergent nozzle is shown here in ***Figure 2.4***. Supersonic flow is terminated not by a normal shock, but by a series of successively weaker oblique shocks and expansion zones. Upstream of the first shock, the boundary layer is just thick enough to be discernible in this image. Immediately downstream of the first shock, it thickens drastically or separates. In a classic experiment carried out in 1903, Stodola [12] measured centreline pressure profiles in a conical nozzle resembling one of the nozzles tested in the current drug delivery research, operating under a range of exit pressures. In Stodola's results, shown in the right hand side of ***Figure 2.4***, the shock appears as a sharp discontinuity only when it is located close to the throat. When deceleration from supersonic to subsonic flow occurs at any substantial distance from the throat, compression is effected through a gradual, uneven pressure rise.

[redacted]

Figure 2.4 Examples of extended shock systems in convergent-divergent nozzles: (*left*) a schlieren photograph by Prandtl, reproduced from reference [11], and (*right*) pressure profiles measured by Stodola [12].

The origin of these phenomena is the shock wave–boundary layer interaction, which is a major area of research in itself. A generic interaction of a normal shock with a flat plate boundary layer is illustrated in ***Figure 2.5***, a diagram by Adamson and Messiter [13]. There must always be a subsonic layer adjacent to the wall, embedded in the boundary layer. The adverse pressure gradient associated with the shock tends to cause a thickening, and often a separation, of the boundary layer. Because pressure information can propagate upstream through the subsonic layer, the thickening effect extends upstream of the shock, resulting in an outward deflection of flow. In the supersonic outer region, this leads to the formation of oblique compression waves which coalesce into an oblique shock and intersect the main shock. A third shock is required to prevent a mismatch of flow direction downstream of the system; it extends obliquely downstream of the intersection point and into the boundary layer. This structure is often called a λ-*shock,* after its characteristic shape.

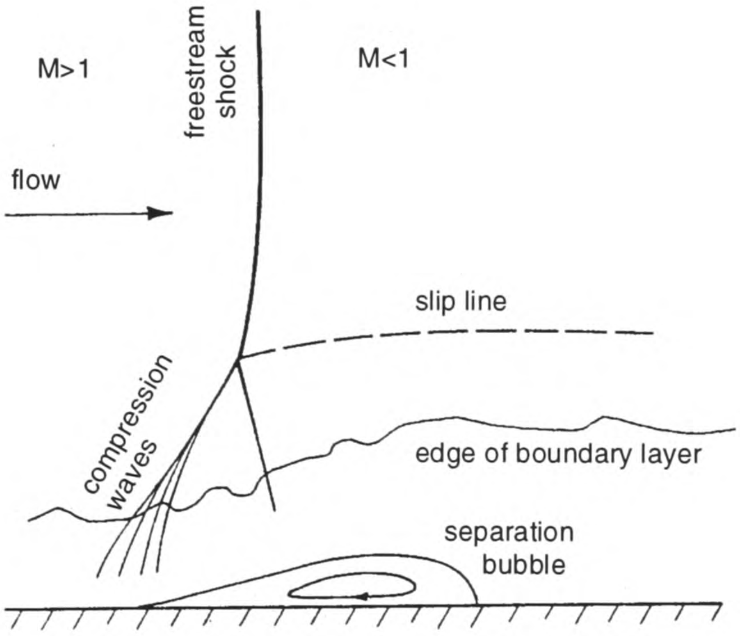


Figure 2.5 Sketch of a shock wave–boundary layer interaction, after Adamson and Messiter [13],with labels added by the present author.

The length scale of this flow structure is determined by the boundary layer thickness, and it is easy to see how the oblique shocks at the base of the system can extend as far as the centreline in a duct with a thick boundary layer, so that the normal shock vanishes. Additional shocks may then be required downstream of these first shocks to complete the compression of the nozzle exit flow, as in the Schlieren image in ***Figure 2.4***.

When a shock triggers separation in the divergent part of a nozzle, self-sustaining flow oscillations can occur, even if the nozzle operating conditions are nominally steady. Using Mach-Zehner interferometry, Meier *et al.* [14] observed and investigated a behaviour in which a shock periodically moved upstream past the throat and out of the nozzle, leaving wholly subsonic flow. As supersonic flow was re-established in the divergent part, the shock re-appeared, and the cycle began again. Szumowski *et al.* [15] confirmed this and identified two additional modes of oscillation — one in which the shock and the separation point oscillated together, and one in which they drifted apart for part of the cycle, with the shock remaining downstream of the throat at all times in both cases. The behaviour of a single test nozzle switched between these three modes as the ratio of total to exit pressure was varied.

Biedron and Adamson [16] give the clearest explanation of these oscillations by considering a shock, with an associated boundary layer separation, which is perturbed downstream from its equilibrium position. In the absence of shock-induced separation, the shock would drift upstream to its correct location. However, the further downstream the shock is perturbed, the stronger the shock and the larger the separated flow region. The backflow in the separation bubble (according to Meier *et al.* and Szumowski *et al.*) or the effective constriction of the nozzle area downstream of the shock due to the separation bubble (according to Biedron and Adamson), or both, raise the back pressure and give an additional impetus to the shock to accelerate upstream. Consequently, the shock overshoots the equilibrium position, until these mechanisms are reversed and it is pushed downstream to repeat the process. Though not explicitly mentioned by the authors cited above, the overshoot may be due to delay in the response of the separated flow to changes in the strength and position of the shock.

2.2.3 Flow Establishment in Convergent-Divergent Nozzles

Steady nozzle flow, discussed in the previous section, is not a useful model for the very early stages of flow in the nozzle of a drug delivery device, when a shock wave is formed at membrane bursting and propagates through the nozzle. A transitional unsteady flow of some kind must take place before steady nozzle flow can be established. This type of starting process has been the subject of some research, motivated mainly by interest in short-duration supersonic wind tunnels consisting of a divergent or convergent-divergent nozzle located downstream of a shock tube. In these tunnels, a strongly unsteady flow is set in motion by propagation of a shock wave through the nozzle. This gives way after some time to the desired steady test flow.

The fundamental question of shock propagation through a duct of varying area has been studied as a mathematical problem by several authors. In 1957, for example, Chisnell [17] derived a useful approximate theoretical relationship between shock strength and changes in duct area, showing that a shock weakens on passing through an increase in area. This method was analysed in more detail by Whitham [18]. The form of Chisnell's expression is $Af(p_2/p_1, \gamma)$ = constant, where p_2/p_1 is the pressure ratio across the shock, and A is the varying duct area. The function f is defined by a lengthy expression which will not be repeated here. Chisnell showed good agreement between this approximation and exact solutions for cylindrical and spherical shocks. His formula has been used in this research to estimate shock propagation speeds, as described in Chapter 4.

Smith carried out quasi-one-dimensional numerical simulations [19] and experimental investigations [20] of the unsteady flow in the scenario shown schematically in ***Figure 2.6***. A divergent nozzle is placed in the end wall of a shock tube, with a diaphragm positioned in the opening which forms the inlet and throat of the nozzle. The nozzle flow is initiated when a shock in the main tube arrives at the end wall and is reflected. The diaphragm bursts, and a *primary shock* propagates into the nozzle, becoming weaker and slower as it moves through the diverging duct. Compression waves travel upstream from the weakening shock, and coalesce to form a *secondary shock*, which propagates upstream with respect to the flow, but is swept downstream in the developing supersonic nozzle flow. Meanwhile, an unsteady expansion wave follows the shock system, also propagating upstream relative to the local flow, but downstream in the nozzle frame of reference. Gas moving through this expansion wave accelerates from sonic flow at the throat and approaches the supersonic steady state. Smith's major finding is that it is usually the unsteady expansion which completes the starting process, thus setting its duration. However, as ***Figure 2.6*** hints, the secondary shock propagates across the expansion wave, and can cross it almost completely under certain conditions. Smith notes that this occurs when the initial pressure in the nozzle (that is, the pressure downstream of the primary shock) is significantly higher than the pressure in steady flow. When that is the case, the secondary shock is swept through the nozzle behind the expansion wave and marks the close of the starting process.

[redacted]

Figure 2.6 Smith's sketch [20] of wave propagation in starting of a supersonic divergent nozzle driven by a shock tube.

Amann [21] and Amann and Reichenbach [22] used shadowgraphs to visualise a flow of this kind. One of their images is shown here in ***Figure 2.7***. The primary shock, the secondary shock (which displays a λ structure and induces boundary layer separation) and the contact surface are clearly visible. Other images revealed complex shock reflection systems in the nozzle throat, from which the secondary shock originates. More recently, Tocarcik-Polsky and Cambier [23] carried out computational simulations of the flow investigated experimentally by Amann and Reichenbach, enabling the two-dimensional shock systems to be examined in great detail.

When total pressure is not sufficient to drive a supersonic flow to the nozzle exit in steady state, a stationary shock system must be formed in the nozzle. The establishment of this shock system can be observed directly in schlieren photographs obtained by Gvozdeva and Zhilin [24]. In the image shown in the right of ***Figure 2.7***, the unsteady secondary shock has split into a pair of steady intersecting oblique shocks and an unsteady normal shock. In subsequent images, the transient element of the structure was absorbed into the steady flow structure as a second pair of oblique shocks. The steady shock system induces boundary layer separation without downstream re-attachment, effectively setting up a parallel jet inside the nozzle.

[redacted]

Figure 2.7 Images by Amman and Reichenbach [22] (*left*) and Gvozdeva and Zhilin [24] (*right*) of wave structures in the starting flow in a convergent-divergent nozzle, with labels added by the present author.

Whereas the other researchers cited above considered only nozzles with convergent sections of small (or zero) length and area change, Britan and Vasil'ev [25] computed and measured the starting process in a complete convergent-divergent nozzle, attached directly to the downstream end wall of a shock tube, without any secondary diaphragm in the nozzle. This experimental configuration is similar to the arrangement of a drug delivery device. Britan and Vasil'ev found that the greatest influence of the convergent part of the nozzle was to cause an elevation of pressure and temperature in the inlet during the early stages of flow, due to a partial reflection of the incident shock. Results presented in the form of instantaneous axial profiles of density suggest that steady flow is established immediately after the passage of the secondary shock; there is no evidence of an unsteady expansion following the shock. This is consistent with Smith's observation that the secondary shock moves upstream of the expansion wave only when the initial nozzle pressure is greater than the steady pressure. In the results presented by Britan and Vasil'ev, pressure data are not given, but the initial (downstream) density in the nozzle is slightly less than the computed density in the steady flow.

2.2.4 Gas-Particle Flow

Interest in gas-particle flow stems from its importance in several applications, including rocket nozzles, pipeline transport of powdered materials, and combustion. Comprehensive reviews of gas-particle flow research have been published by Rudinger [26] and (as part of a broader text on multiphase flow) Soo [27]. The starting point taken by both authors is Newton's Second Law for a single spherical particle in a flow, expressed in this form:

$$\rho_p \frac{\pi d_p^3}{6} \frac{du_p}{dt} = \frac{1}{2} C_D \rho |u - u_p| (u - u_p) \frac{\pi d_p^2}{4} \qquad (2.10)$$

where C_D is the drag coefficient, u and u_p are the velocity of gas and particle, respectively, ρ and ρ_p are the gas and particle density, and d_p is the diameter of the particle. The term $(u - u_p)$ is called the *slip velocity*. Generally, C_D depends on Reynolds number $Re = \frac{\rho |u - u_p| d_p}{\mu}$ and Mach number $M = \frac{|u - u_p|}{a}$, where a is the local speed of sound in the gas. ***Eq. 2.10*** implies the basic result that particle acceleration in response to a given slip velocity is inversely proportional to particle diameter (in the absence of large variations in drag coefficient). Various methods for the determination of C_D are available, as discussed in §7.2.2 of this thesis.

If the mass fraction of particles in the flow is small, the gas flow is unaffected by their presence. For large mass fractions, the transfer of momentum (and heat) between the particles and the gas can have a significant effect, and the particle and gas flow-fields are strongly coupled. If the volume fraction is high, further complications appear, since the area through which the gas can flow is reduced. The effects on the gas flow of momentum exchange between gas and particles can be understood in precisely the same way as friction, for which a standard one-dimensional analysis is given by (for example) Shapiro [11]. When the particles lag the gas, they cause Mach number to tend towards 1 and total pressure to decrease. In subsonic flow, pressure and temperature increase, and in supersonic flow, they decrease, as a consequence of the particles' action. One of the interesting consequences of this fact for flow in a convergent-divergent nozzle is that Mach number 1 does not occur in the throat, but downstream of it.

The literature on gas-particle flow is concerned almost exclusively with steady flow. It offers little guidance as to the mechanisms which determine the rate of particle uptake or entrainment when a mass of resting particles is suddenly exposed to flow, as in the drug delivery system. At least two groups of researchers have conducted relevant work in this area. Anilkumar *et al.* [28], investigating flow in volcanoes, placed a bed of particles on a horizontal plate or mesh support in the driver section of a vertical shock tube. On rupture of the shock tube diaphragm, an expansion wave would propagate upstream (downwards) through the bed, causing the particles to accelerate. Some photographs of an accelerating cloud of uniform particles are shown in ***Figure 2.8***. Anilkumar *et al.* speculated that the bed initially develops horizontal fractures at regions of lower packing density, which are then partitioned by particles falling across the fractures. The resulting bubbles are elongated in the flow direction, eventually leading to the filament-like structures seen in the later images. Pressure measurements taken by Anilkumar *et al.* showed increasing pressure as the particle cloud approached and passed over the transducer location, due to the pressure drop required to accelerate the cloud. They found that the particle cloud approached homogeneity only when particle volume fraction fell below 1%, very late in the flow.

[redacted]

Figure 2.8 Photographs by Anilkumar *et al.* [28] of particle bed acceleration. The first four frames on the left were taken in the first 5 ms from rest, and the others span 30 to 40 ms after initiation.

Rogue *et al.* [29 and 30] studied a variation in which the particle bed was supported on a mesh or a light membrane, above and downstream of the diaphragm in a vertical shock tube. In this case, a shock wave passed through the particle support with minimal disruption, impacted on the bed, and was partially transmitted and partially reflected. Horizontal fractures (that is, normal to the flow direction) were again observed at the leading edge of the bed, but unlike Anilkumar *et al.*, Rogue *et al.* asserted that they were due to particle collisions in the flow direction. A strong pressure gradient across the cloud was again measured.

In an interesting series of experiments described in reference [29], particle beds of various configurations were studied. First, a few isolated particles were placed on the support; then a single dense layer of particles, then a double layer, and finally, a substantial bed of about 15 layers were tested.

Photographs of the particle cloud showed that the single and double layers accelerated more rapidly than any other beds. The experimenters attributed this to the large pressure difference which can be maintained across beds of one and two layers, and the high gas velocity due to constricted flow area within a dense bed. This trend of increasing acceleration is reversed as more layers are added to the bed because of the greater resistance offered to the flow. These experiments illustrate the complex effects of particle loading in non-equilibrium gas-particle flow.

2.3 Test Devices and Particles

This section describes the specific drug delivery systems which have been tested in the course of this research. The physical characteristics of each system are determined primarily by the device (described in §2.3.1), but the drug (or model drug) particles are also a vital element of the system. Test particles are described in §2.3.2.

2.3.1 Drug Delivery Devices

As mentioned in Chapter 1, this research into drug delivery flow-fields has been approached by making measurements on the flow-fields created in existing prototype drug devices of three designs, which have been used extensively in pre-clinical and clinical trials. The geometries of the major components, and the terminology used to described them, are specified in detail in ***Figure 2.9***.

The major components of the device are the cylinder and the rupture chamber (together making up the reservoir system), the drug cassette, and the nozzle. Three different nozzle geometries have been used, but the internal geometries of the cylinder, rupture chamber and drug cassette were not varied throughout this work.

Compressed helium is initially stored in the *cylinder*, which is in fact only approximately cylindrical in its internal shape. It has a useful volume of 4.6 ml. At the downstream end of the cylinder, a valve is formed where the 4.7 mm diameter plunger passes through a 5 mm diameter opening. When the plunger is retracted, the gap between the plunger and the valve wall is sealed by an O-ring fixed to the plunger. When the plunger is pushed forwards, the O-ring clears the cylinder to open an annular gap of (nominal) area 2.36 mm^2. The other end of the plunger is exposed at the opposite end of the cylinder, through a gap which is permanently sealed by a second O-ring. The cylinder is pressurised by detaching it from the rest of the device, coupling the outlet valve to a filling system fed from a helium bottle, depressing the plunger to open the cylinder outlet valve, and opening a valve in the filling system. Fill pressure is set by a regulator on the gas bottle.

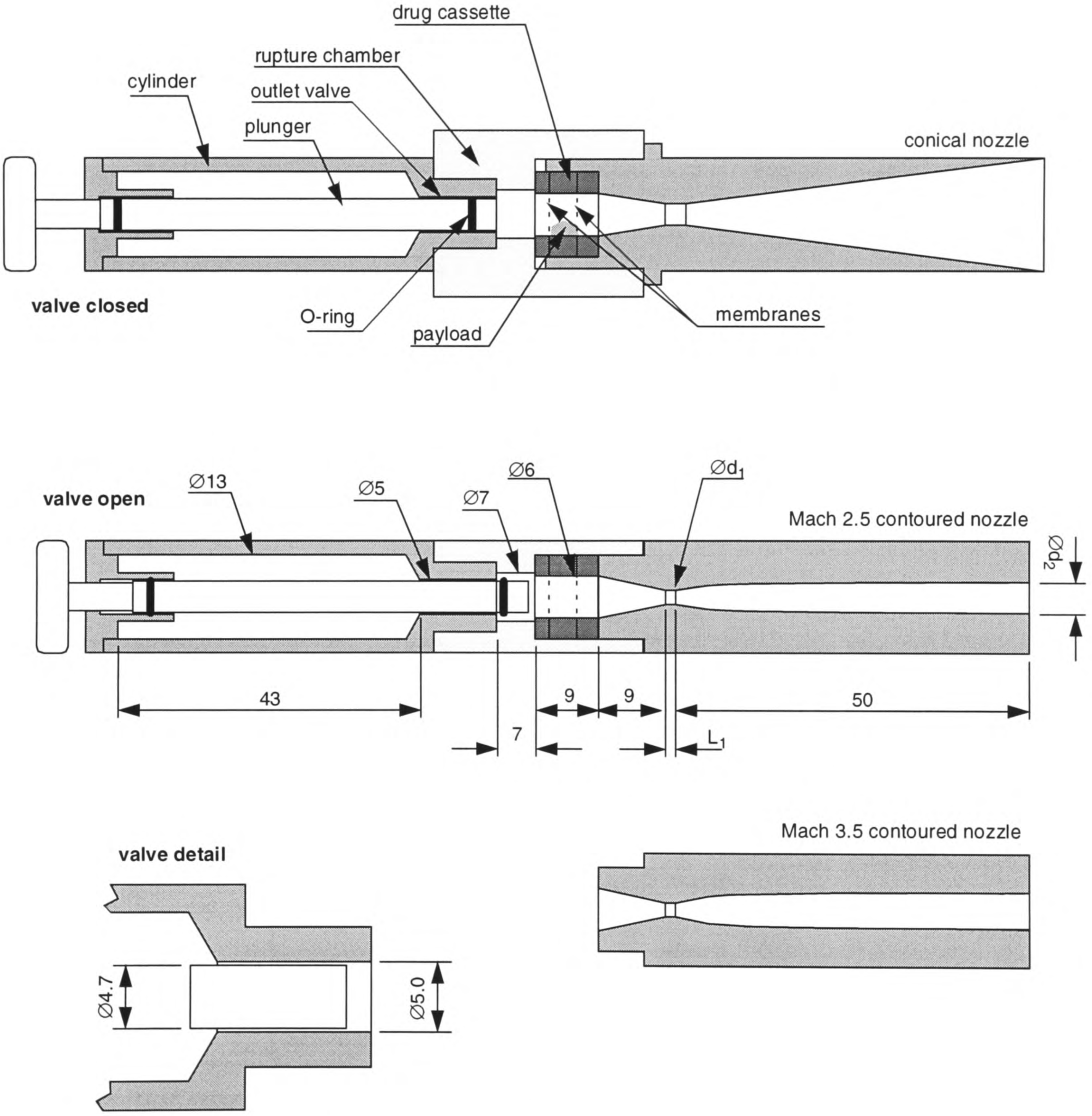


Figure 2.9 Geometry of the test devices with dimensions in mm. The dimensions L_1, d_1 and d_2 are specified in ***Table 2.1***.

The valve opens into the *rupture chamber*, a small empty volume whose primary purpose is to isolate the high-pressure gas store from the drug cassette. Its volume (with the plunger protruding) is 0.134 ml. The *drug cassette* is a rigid structure of three plastic rings, on which the two bursting membranes are mounted. A chamber formed by the membranes and the central ring contains the powder drug.

Various mylar membranes from 12 to 75 μm in thickness, and polycarbonate membranes of 20 μm thickness, have been used in this research. The cassette is assembled in advance of a shot, placing the *payload* (a collective term for the drug or other particles to be delivered in a shot) in the space between the membranes. The cassette is then clamped between the nozzle and the rupture chamber when the nozzle is threaded into place. The volume between the membranes is 0.254 ml.

Sketches of the three nozzles tested are included in ***Figure 2.9***. The Mach 2.5 and Mach 3.5 contoured nozzles possess contoured divergent profiles, which were designed for shock-free supersonic flow using the method of characteristics [11]. They are very similar in their overall dimensions, and their convergent sections and throat sections are identical. They differ only in their design exit Mach number (the specified exit Mach numbers of 2.5 and 3.5 are nominal values, based on quasi-one-dimensional isentropic flow theory). The conical nozzle differs substantially from the contoured nozzles. It was designed with the aim of distributing drug over a large area, and features an enlarged throat diameter and a greatly enlarged exit diameter, with a simple conical divergence from throat to exit. It closely resembles the nozzles to be used in the first commercial drug delivery devices.

nozzle designation:		**Mach 2.5 contoured**	**Mach 3.5 contoured**	**conical**
geometry:	convergent section	conical	conical	conical
	throat diameter $\mathbf{d_1}$ (mm)	1.5	1.5	3
	throat length $\mathbf{L_1}$ (mm)	1.5	1.5	3
	divergent section	MOC design	MOC design	conical 6° half-angle
	exit diameter $\mathbf{d_2}$ (mm)	2.23	3.10	13.5
	exit to throat area ratio	2.21	4.27	20.3
conditions in supersonic exit flow	exit Mach number	2.54	3.53	6.52
	exit velocity (m/s)	1450	1580	1700
minimum ratio of total pressure to back pressure for:	choking at throat	1.06	1.02	1.00
	supersonic exit flow	2.20	3.96	17.3

Table 2.1 Geometric parameters of the three nozzles tested, with some operating parameters based on isentropic quasi-one-dimensional flow theory. MOC: method of characteristics.

Table 2.1 lists some major dimensions and gas dynamic parameters which describe each nozzle. The exit Mach number, exit velocity and critical pressure ratios given in the table are based on equations from frictionless quasi-one-dimensional flow theory (summarised in §2.2.2), with $\gamma = 5/3$ for helium. These parameters are given as an indication of expected nozzle performance, rather than as an accurate prediction. The exit velocities specified in the table are based on a total temperature of 293 K. The last two rows of the table indicate the minimum ratio of total pressure to exit static pressure required for two of the nozzle flow regimes identified in §2.1.2, namely choked flow and fully supersonic flow. It can be seen that the large exit area of the conical nozzle, in comparison with the contoured nozzles, brings a very large increase in Mach number, and a correspondingly large jump in the pressure ratio required to drive supersonic flow. The gain in exit velocity, however, is slight.

2.3.2 Particles

Lidocaine hydrochloride (usually referred to simply as *lidocaine*) is widely used as a local anaesthetic, and is planned as the first commercial application of the new drug delivery technology. Since it is also inexpensive and non-toxic, it was an obvious choice of test particle with which to evaluate device performance under realistic operating conditions. Lidocaine is a crystalline compound which is supplied by the manufacturers as a powder with a broad unspecified particle size range. In the current work, the raw powder was graded by sieving into nominal size categories 20–38 μm, 38-53 μm, 53-75 μm and 75-106 μm. The relative density of lidocaine is 1.20.

Many experiments with lidocaine have been carried out in order to investigate the drug delivery flow-fields under operational conditions. However, for more general investigations, lidocaine was found to be unsuitable as a representative particle, because of batch-to-batch variations, moisture absorption, the lack of a reliable means of sizing the particles, and the suspicion that particles were breaking up in device actuation. These factors contributed to uncertainty as to the properties of the particles. Therefore, polystyrene microspheres, supplied in a range of well-specified size distributions by Duke Scientific, were adopted as a model drug particle. With these particles, it was possible to be confident that variability in particle properties was not introducing unwanted variations to experimental results, and (for example) to test the drug delivery system for sensitivity to particle size.

The sizes of the particles used (specified as mean diameter ± standard deviation) are 4.7 ± 0.6 μm, 15.5 ± 2.0 μm and 26.1 ± 2.5 μm. Their relative density is 1.05.

The third and final particle material used in drug delivery payloads was a hydrophobic silica-based powder, Aerosil R972, manufactured by Degussa [31]. The manufacturers state that the "average primary particle size" is 16 nm, and that the particles tend to form loosely bound agglomerates. Large agglomerates are certainly visible to the naked eye in samples of the powder. Given that particle sizes of the order of microns are normally used successfully as flow tracers in laser anemometry, 16 nm particles suspended in a gas dynamic flow field would be expected to follow the gas velocity extremely closely. This was the motivation for using these particles.

2.4 Summary

Flow phenomena in gas-particle flows for delivery of powdered drugs have been discussed in this chapter, and the prototype drug delivery devices tested in the course of this research have been described in detail. The resemblance of the real flow in the drug delivery device to two idealised gas flow systems, convergent-divergent nozzle flow and shock tube flow, has been discussed, and key results from the mathematical theory of these flows have been presented. These theoretical descriptions have been expanded with a review of some relevant experimental and computational results from the literature. In particular, the impact of viscous effects on convergent-divergent nozzle flow, and the unsteady flows which result when a shock tube flow is used to initiate a nozzle flow, have been discussed. Finally, the prototype drug delivery devices tested in this research were specified in detail.

Chapter 3 Point Experiments: Methods

Two classes of point experimental technique have been used in this research. They are pressure measurements, carried out in various modes inside drug delivery devices, and optical particle detection, used to record the timing of particle flow at a location just outside the nozzle exit. The instrumentation, signal processing, data acquisition, data processing and physical procedures which make up these experimental techniques are outlined in this chapter.

3.1 Pressure Measurements

In order to study the gas flows in the prototype drug delivery devices, pressure was measured at two locations in both the Mach 3.5 contoured nozzle and the conical nozzle, and in the rupture chamber fitted to the conical nozzle. The transducers used in this work, and their installation in the test devices, are described in §3.1.1. Procedures for the recording and manipulation of the transducer outputs are specified in §3.1.2.

3.1.1 Apparatus

Semiconductor pressure transducers of a type described by Ainsworth *et al.* [32] have been used for nozzle applications. Each transducer consists of a Wheatstone bridge circuit laid down on a silicon diaphragm, which is mounted over an evacuated cavity. The Wheatstone bridge acts as a strain gauge — the diaphragm deforms under loading, and the value of each resistance is altered. A constant input voltage is applied across two opposite terminals of the bridge, and the relatively small output voltage is measured across the other pair of opposite terminals. The output voltage varies with the resistance of the bridge elements, and can be calibrated for its dependence on pressure. These transducers achieve a flat frequency response up to the diaphragm's natural frequency, which is typically of the order of 400 kHz.

Specifically, Kulite XCQ-150A semiconductor transducers were used in this work. This model was chosen primarily for its compact design — the diaphragm measures 1.2 × 1.2 mm. Two transducers were flush-mounted in the walls of each nozzle. One was centred 3.25 mm upstream of the nozzle exit, and the other a further 15 mm upstream. As shown schematically in ***Figure 3.1***, each transducer was mounted on a ceramic rod and recessed in a hole drilled though the nozzle wall. The remaining gap between the diaphragm and the original nozzle wall surface was filled with RTV, an elastomeric material, so that disruption to the nozzle wall geometry was minimised, and the diaphragm was damped and protected. Installation and calibration were carried out by the late Mr. John Allen in the Department of Engineering Science, University of Oxford. Previous research [32] has established that filling material tends to damp the resonant oscillations of the transducer, but does not affect the value of the measured pressure.

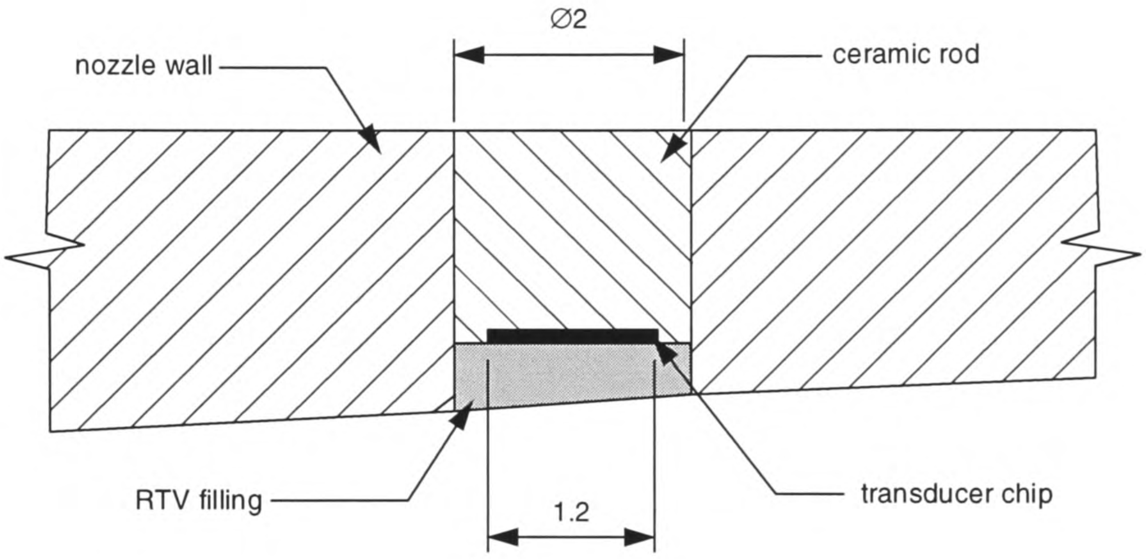


Figure 3.1 A longitudinal section of a nozzle wall, showing the transducer installation schematically. Dimensions in mm.

The nozzle transducers were calibrated after installation by placing each nozzle inside a pressure vessel and recording the output voltage under steady loading at absolute pressures from 0.2 bar to 3.4 bar. Three of the four transducers displayed excellent linearity, but the conical nozzle's downstream transducer displayed hysteretic behaviour during the calibration which caused its response to depart from linearity by up to 0.38 bar, or 3.7% of nominal full scale (150 psi).

A Kistler type 601A piezoelectric transducer was used for pressure measurements in the rupture chamber. Pressure on this sensor's steel diaphragm is transmitted to a quartz crystal underneath, which becomes electrically charged in response to the loading, with a sensitivity of approximately 16 pC/bar. The charge in the crystal is converted to a voltage signal by a dedicated amplifier. The Kistler transducer is cylindrical, with a diameter of 5.5 mm on the sensor face, while the internal diameter of the rupture chamber is 7 mm. Because of the transducer's relatively large size, it partly protruded into the rupture chamber and was partly recessed in its mount, as shown in ***Figure 3.2***. However, the relatively low gas velocities expected in the rupture chamber meant that the possibility of flow disruption by the instrumentation was less of a concern than in the nozzle. The Kistler transducer was calibrated by Mr. Mark Roberts under steady pressures in a small custom-built pressure vessel.

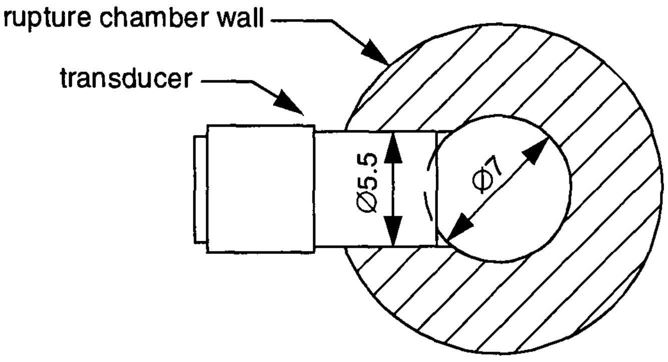


Figure 3.2 Cross-section of the rupture chamber, with pressure transducer fitted. Dimensions in mm, not to scale.

3.1.2 Signal Processing and Data Acquisition

The nozzle pressure transducer signals were amplified to protect against corruption by electrical noise. The outputs of transducers in the Mach 3.5 nozzle were usually amplified at a gain of 50 using AD524 amplifiers. Signals from the conical nozzle transducers were passed through amplifier modules based on AMP-05 chips, operating at a gain of 100. The AMP-05 systems were chosen for this task because of their built-in filters, as explained below. The output of the piezoelectric transducer in the rupture chamber was processed by a dedicated charge amplifier.

Two data acquisition systems were used. In the majority of experiments, a Datalab DL12 waveform recorder was used with all instrumentation. It provided 12-bit analogue-to-digital conversion and could record 8192 samples at a rate of up to 500 kHz. In some experiments, a lower sampling rate of 200 kHz was selected to enable data acquisition over a longer period. When higher sampling rates were required to study rapid events in the early stages of nozzle flow, an Iwatsu digital oscilloscope was used at a sampling rate of up to 20 MHz. The data recorded by both of these systems were downloaded to a personal computer via an IEEE-488 interface for storage. Subsequent analysis and manipulation of the data was carried out using the numerical utilities software package Matlab.

In both nozzles, the magnitude of resonant oscillations of the transducers (at 320–400 kHz) was sometimes large enough to degrade the transducer output signals. Consequently, filtering was required. In the case of the Mach 3.5 contoured nozzle, the influence of oscillations was slight over most of the duration of a run, and filtering was not necessary when studying pressure variations over the longer time scales of interest. It was only in measurements of the very early stages of flow, captured on the oscilloscope at high sampling rates, that resonance presented a problem for work with the Mach 3.5 contoured nozzle. Software filtering was employed to suppress resonant oscillations when required.

In the conical nozzle, on the other hand, the transducers were subjected to fluctuating pressures which excited resonance to a greater extent over much of the lifetime of nozzle flow, and the signals were often seriously degraded. Transducer signals from the conical nozzle were usually sampled at 500 kHz using the Datalab recorder, with the result that the resonant oscillations would appear at an aliased frequency. Therefore, the signals were routinely passed through a 150 kHz low-pass filter before sampling. These filters were built into the AMP-05 amplifier modules.

Occasionally, software filtering has been applied to pressure measurements presented in this thesis, purely to improve the clarity in certain presentations of results (for example, when multiple signal plots are overlaid in a graph). Processing of this kind is noted in the text whenever it has taken place.

3.2 Optical Particle Detection

The purpose of the Optical Particle Detection (OPD) experiments is to complement the pressure measurements by identifying the periods of the gas flow history in which the particles are delivered. This knowledge is desirable as a basic and important element of the understanding of the mechanics of powder drug delivery devices.

The principle of the OPD technique is simple. As particles emerge from the drug delivery device, they pass through a laser beam just outside the nozzle exit and scatter light onto an optical sensor. A temporal record of particle delivery is contained in the time-varying intensity of the sensor's output signal, which depends on the number of particles in the illuminated volume at any time. The apparatus is described in full in §3.2.1. Procedures for the manipulation of the signal to obtain the final form of the data are outlined in §3.2.2.

3.2.1 Apparatus

The simple apparatus sketched in ***Figure 3.3*** provided a means of detecting drug particles as they emerged from the device. The beam from the Argon ion laser described in §5.2.4 was focused to a diameter of approximately 2 mm and directed across the nozzle exit, passing through the nozzle centreline at right angles approximately 2 mm downstream of the nozzle exit plane. The drug delivery device was positioned vertically, firing downwards. A photodiode was positioned just off the beam, with its 1 mm^2 sensing element facing towards the intersection of the laser beam and the nozzle centreline. Photodiode type BPX65 was chosen for its rapid response time, which is specified as 3.5 ns under 20 V reverse bias. The photodiode current, typically of the order of 5 μA, was amplified by the simple circuit shown in ***Figure 3.4***, which was built around a TL071 amplifier. The amplifier circuit also maintained a reverse bias of 9V across the photodiode to minimise its response time. The amplified photodiode signal was recorded on a Datalab DL12, as described in §3.1.2, and then passed to a computer.

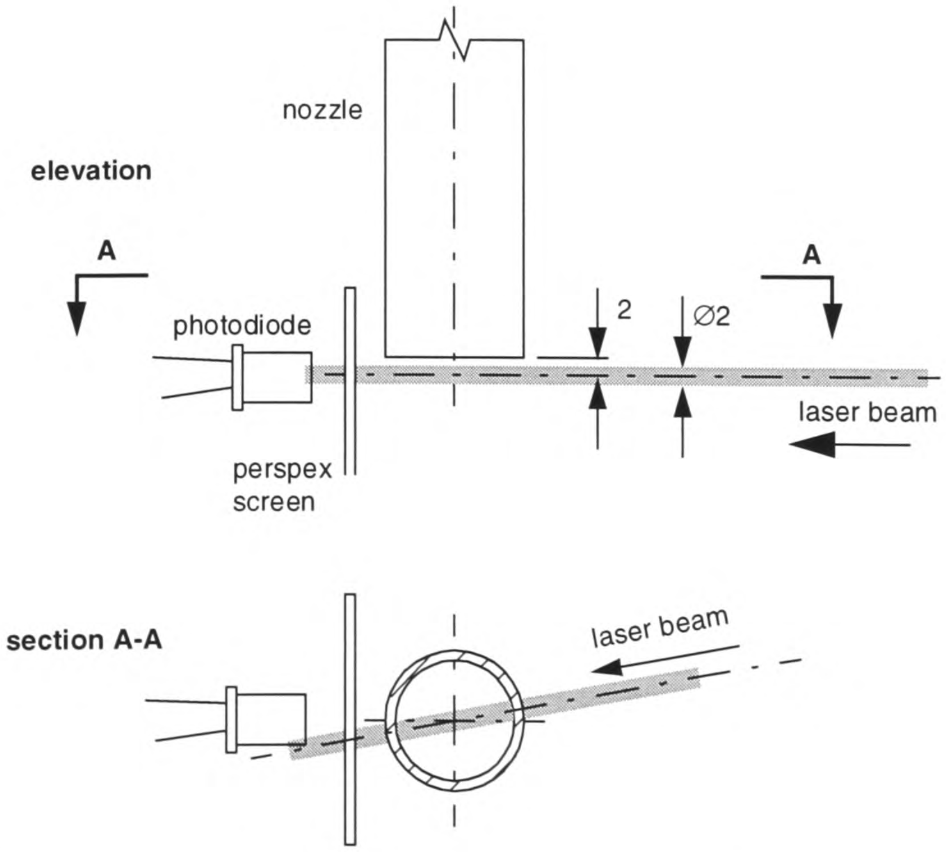


Figure 3.3 Arrangement of the optical particle detection apparatus and the device under test. Dimensions in mm, not to scale.

A perspex screen was located just in front of the photodiode, mounted independently of it, to shield it from mechanical disturbances due to the flow. In early versions of this experiment, the photodiode registered a spurious signal, even in the absence of light, which was thought to be due to currents induced through movement of the photodiode or its leads. This movement was probably caused by the impact of the starting shock wave on the photodiode.

The intensity of light scattered onto the photodiode face by any one particle is a function of the particle's size, the scattering direction, and the distance between the particle and the sensor. The latter two factors depend on the particle's position. Therefore, if the particle size distribution was narrow (as it is for polystyrene microspheres) and the influence of particle position was slight (for example, if the sensor was very far from the probe volume), then the total light received by the photodiode would be proportional to the mass of particles in the probe volume. Conceivably, the OPD signal could be calibrated as a measure of particle concentration. The current research does not extend to such a quantitative interpretation of the data, since the main aim is knowledge of the timing of particle delivery. However, it is reasonable to regard the OPD signal as an indicator (perhaps a non-linear indicator) of the mass of payload in the flow at any time.

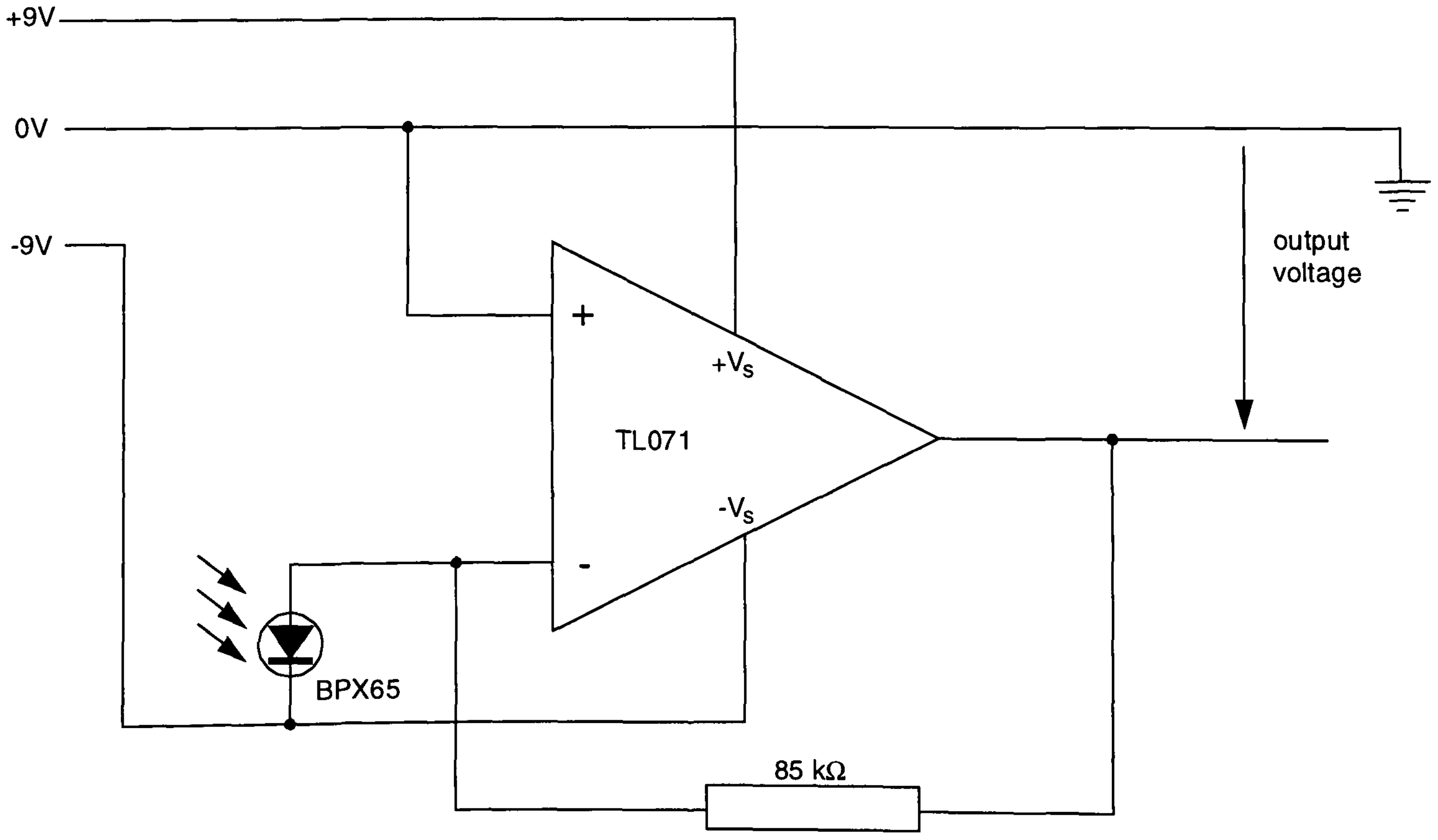


Figure 3.4 The photodiode amplifier circuit used in the optical particle detection apparatus

3.2.2 Data Processing

The data from each OPD run were reduced to a few meaningful scalar parameters to facilitate comparison of individual results and concise presentation of the entire data set. Before computing these parameters, it was necessary to correct each OPD signal history by subtracting out a background level so that a value of zero in the corrected signal would consistently signify the absence of particles. This was essential because any interpretation of the signal magnitude could be skewed by variations in the background level, and calculations which involve integration of the signal over time are particularly sensitive to non-zero base level.

Non-zero background level can be caused by stray laser light, background light and contributions from electronics. Two typical OPD signal histories, recorded on different time scales when the device was fired without a particle payload, are shown in ***Figure 3.5***. Both signals have a substantial constant offset along with high-frequency noise at a level of approximately 0.2 mV (root mean square). Both signals begin with a spike just after time $t = 0$ (when the upstream pressure transducer triggers data acquisition). Since this feature does not appear when the illumination is switched off, it may be due to refraction of the beam through density gradients in the gas flow.

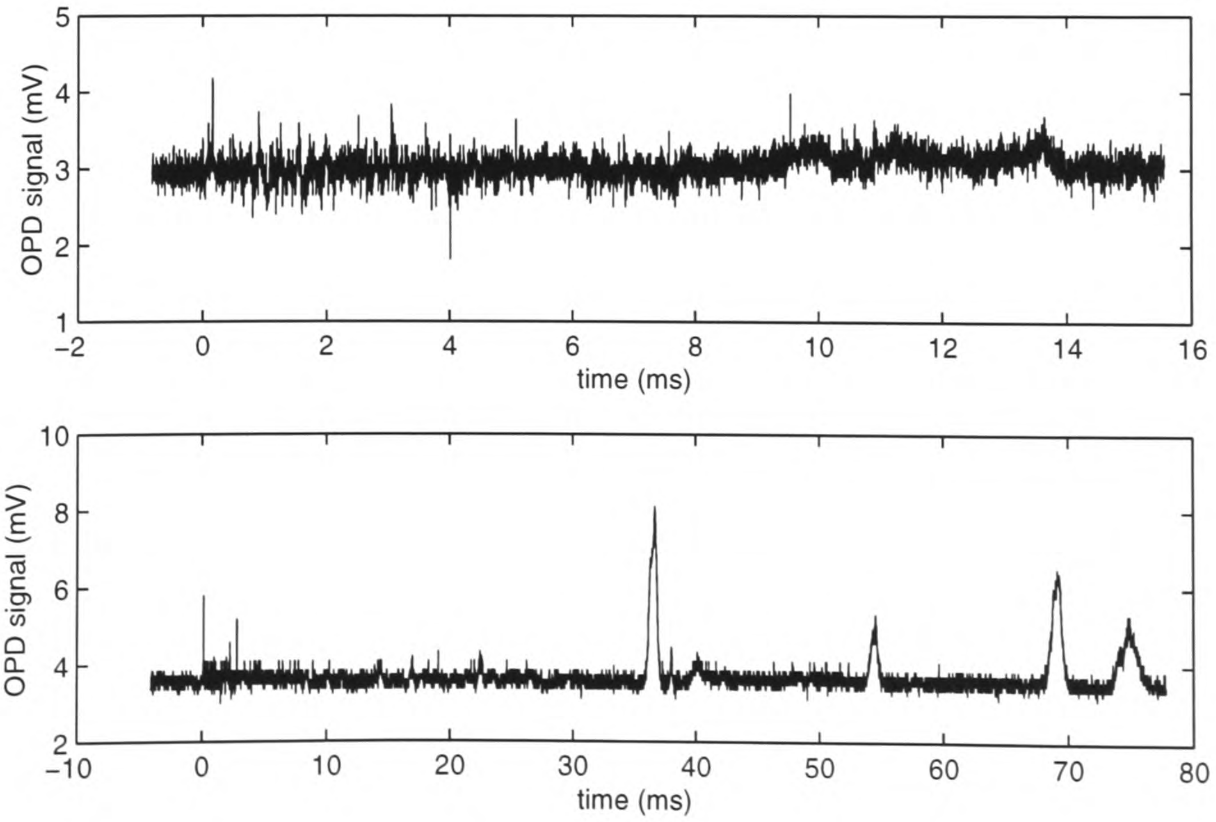


Figure 3.5 OPD signals recorded in shots fired without a drug payload. Note different time scales.

However, a greater problem is posed by spurious unsteady components of the OPD signal which are not due to payload particles. The signal in the upper graph in ***Figure 3.5*** rises very slightly at $t \approx 10$ ms, and the lower plot includes bursts of activity beginning at $t \approx 14$ ms, some of which are very large. These examples are typical of the features which have often been recorded in the absence of payload particles. They can also be recognised in the results of shots fired with drug payloads, after the main activity in the OPD signal ceases. These bursts are very inconsistent, but they can be managed by virtue of the fact that they always appear relatively late in the run. After studying the results of all OPD experiments, a cutoff point at $t = 5.5$ ms was identified. Features like the bursts observed in the lower graph in ***Figure 3.5***, which are not associated with payload particles, have never been observed before $t = 5.5$ ms, and signal due to light scattering by payload particles invariably ceased before that time.

The presence of light-scattering material other than drug in the jet, which could cause these spurious features, has been studied in the context of DGV, and is discussed fully in §5.5.1. That study suggested that the signal was most likely to be due to condensate or debris from the atmosphere which becomes entrained in the jet.

In the light of the issues discussed above, a procedure of three stages was devised for correction of non-zero background level in the raw OPD signals. First, a reference "no-drug" signal history is compiled as an average of three OPD results captured in the absence of payload. This is smoothed and subtracted from every OPD signal history to account for repeatable background contributions to the OPD signal. Then, a contribution specific to each run is calculated by averaging the (already partially corrected) OPD signal over the first 0.6 ms of the run, before the onset of particle delivery. This average value is subtracted from the signal as a whole to compensate for changes in background light levels from shot to shot. Finally, data recorded after $t = 5.5$ ms are discarded.

The parameters used to characterise OPD signals are illustrated in ***Figure 3.6*** for a typical set of results. They are all defined with respect to the background-corrected signal history, as described above. The *time of first delivery* is defined as the time when the background-corrected OPD signal first rises above 2 mV. The *20% rise time* in a run is the time at which the signal first exceeds 20%

of the maximum value attained on that particular run. Both of these parameters are means of identifying the beginning of particle delivery. The third and final parameter is defined with reference to the integration of the processed OPD signal over time. It is the *90% delivery time*, defined as the time at which the integrated signal reaches 90% of its maximum value, and it is a measure of the time required to complete particle delivery. Measurements of the timing parameters — the time of first delivery, the 20% rise time and the 90% delivery time — will be expressed in §4.3.2 in terms of delays from the instant at which the starting shock is detected at the downstream nozzle pressure transducer.

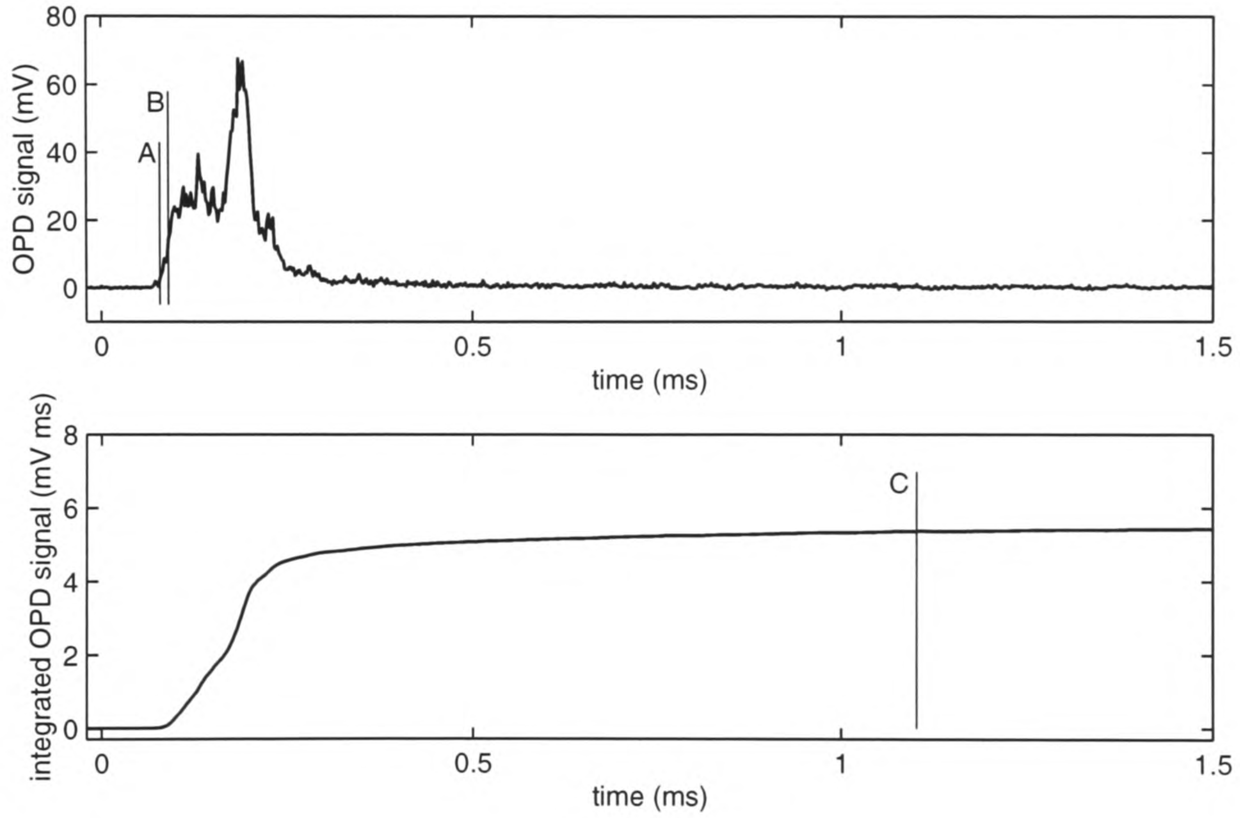


Figure 3.6 Illustration of the scalar parameters used to characterise the OPD time history. A: time of first delivery; B: 20% rise time; C: 90% delivery time.

3.3 Summary

In this chapter, the apparatus and procedures used to measure unsteady pressure in drug delivery devices, and to record the timing of drug particle delivery, have been described. Time-varying pressure has been measured at locations 3.25 mm and 18.25 mm upstream of the exit plane in both the Mach 3.5 contoured nozzle and the conical nozzle. Pressure has also been measured in a rupture chamber fitted to the conical nozzle.

An apparatus for optical particle detection (OPD) has been devised, which enables the temporal variation of particle concentration at a particular location to be recorded in a qualitative manner. The OPD apparatus has been configured to detect particles passing through a probe volume 2 mm downstream of the exit of the conical nozzle. Scalar parameters have been defined which characterise the timings of the beginning and end of particle delivery as a means of summarising the OPD signal data. The results of pressure measurements and OPD experiments are presented in the next chapter.

Chapter 4 Point Experiments: Results

Understanding of the gas flow is an essential element in any useful understanding of the particle delivery system as a whole, since it is through the gas flow that the energy initially stored in the compressed helium is converted to a kinetic form and transmitted to the particles. This chapter is devoted to an experimental study of the gas flow, both with and without particle payloads. The complementary approaches of unsteady pressure measurement and optical particle detection (both described in the previous chapter) have provided a means of direct observation of the gas flow in drug delivery devices, and of the relationship between the gas and particle flows. The pressure measurements, confirming and building on the general theoretical discussions of Chapter 2, support a description of the gas flow. The optical particle detection results enhance the pressure data by enabling identification of the parts of the flow history which are important for particle delivery. They also allow the influence of the particles on the gas flow to be assessed.

Prototype drug delivery devices were studied in various configurations using these point methods. First, pressure measurements were carried out in a device with a Mach 3.5 contoured nozzle and in a device with a conical nozzle, in the absence of any particle payload, to develop a description of the gas flow. In the Mach 3.5 contoured nozzle, pressure was measured at two locations on the nozzle wall for various combinations of fill pressure and membrane strength. Results of these tests are presented and discussed in §4.1. In a device with a conical nozzle, pressure was again measured at two nozzle locations, and also in the rupture chamber, across a range of test conditions. These measurements are described in §4.2.

Optical Particle Detection (OPD) and nozzle pressure measurements were carried out in tandem on a conical nozzle device loaded with various payloads. The last two sections of the chapter deal with the results of these experiments, examining the timing of particle delivery (§4.3) and the influence of the particle phase on the gas flow (§4.4).

4.1 Pressure in the Mach 3.5 Contoured Nozzle

As described in Chapter 3, the Mach 3.5 nozzle was instrumented with pressure transducers at positions 3.25 and 18.25 mm upstream of the nozzle exit. The pressure measurements taken in this nozzle, operating without any model drug, are presented in this section. A representative set of results are presented and discussed in §4.1.1 to establish the character of the flow and illustrate the nature of the measurement. The repeatability of these basic results is demonstrated in §4.1.2. In §4.1.3, data for the very early stages of the flow are examined. In §4.1.4, the effects of device configuration (fill pressure and membrane burst pressure) on the gas flow are assessed. A brief overall description of the flow, based on the results of all these experiments, is given in §4.1.5.

4.1.1 A Representative Result

The time-varying pressure measured at the upstream and downstream transducers in the Mach 3.5 contoured nozzle on a typical run, with 23 μm mylar membranes and a fill pressure of 60 bar, is shown on different scales in ***Figure 4.1*** and ***Figure 4.2***. These data are presented as a typical sample of the results obtained in this work, and are used to support a description of the general structure of flow in the nozzle.

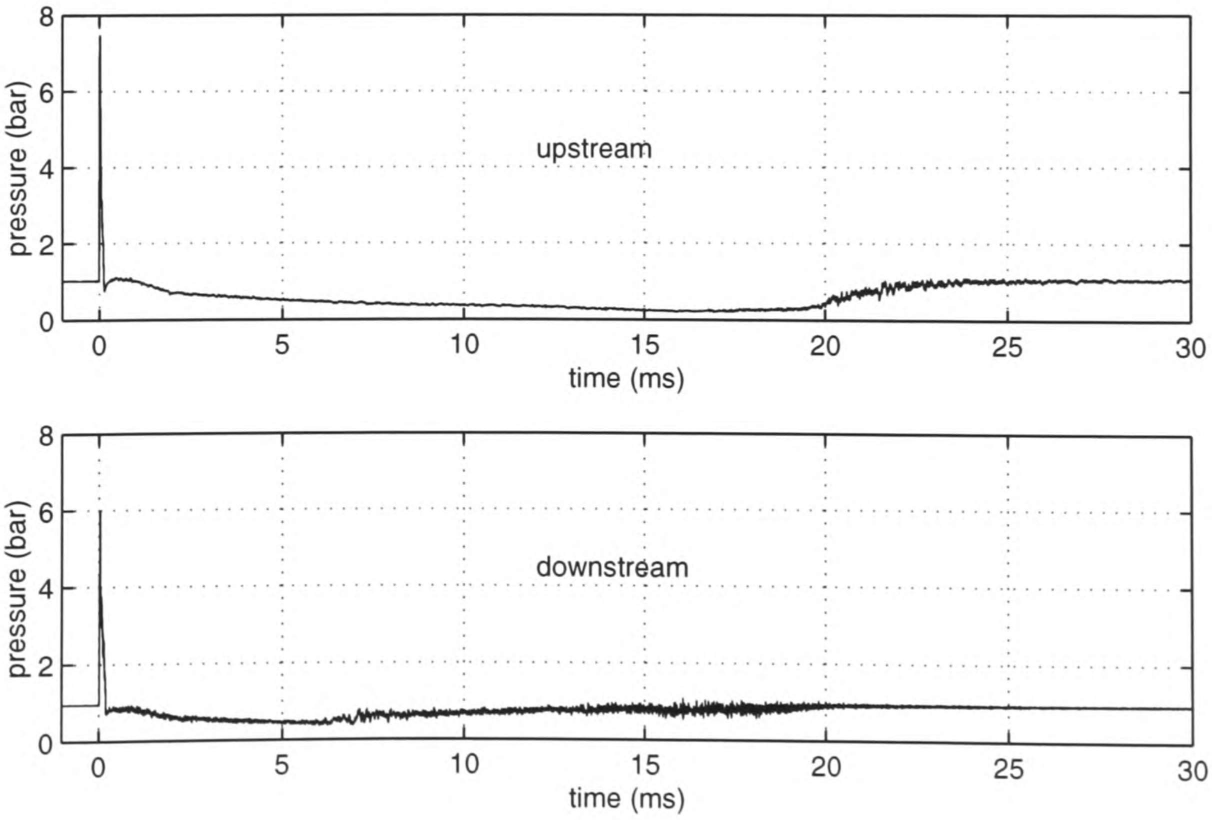


Figure 4.1 Unsteady pressure measured at the upstream and downstream transducers in the Mach 3.5 contoured nozzle, with 60 bar fill pressure and 23 μm membranes.

Initially, measured pressure at both transducers is approximately 1 bar, indicating steady atmospheric pressure inside the nozzle. A sharp rise in pressure to over 7.5 bar is recorded at the upstream transducer at time t = 0. This choice of time zero will be maintained as a convention throughout this chapter. The initial pressure rise is detected slightly later at the downstream transducer, and is followed by a rapid fall in pressure to approximately 1 bar at both transducers. Pressure then rises briefly and begins to fall slowly below atmospheric pressure. This decay comes to an end at the upstream transducer after 19 ms, when pressure begins to rise again and, over a period of about 4 ms, returns to a level just over 1 bar. At the downstream transducer, pressure begins to rise at 7 ms and takes at least 13 ms to return to atmospheric pressure.

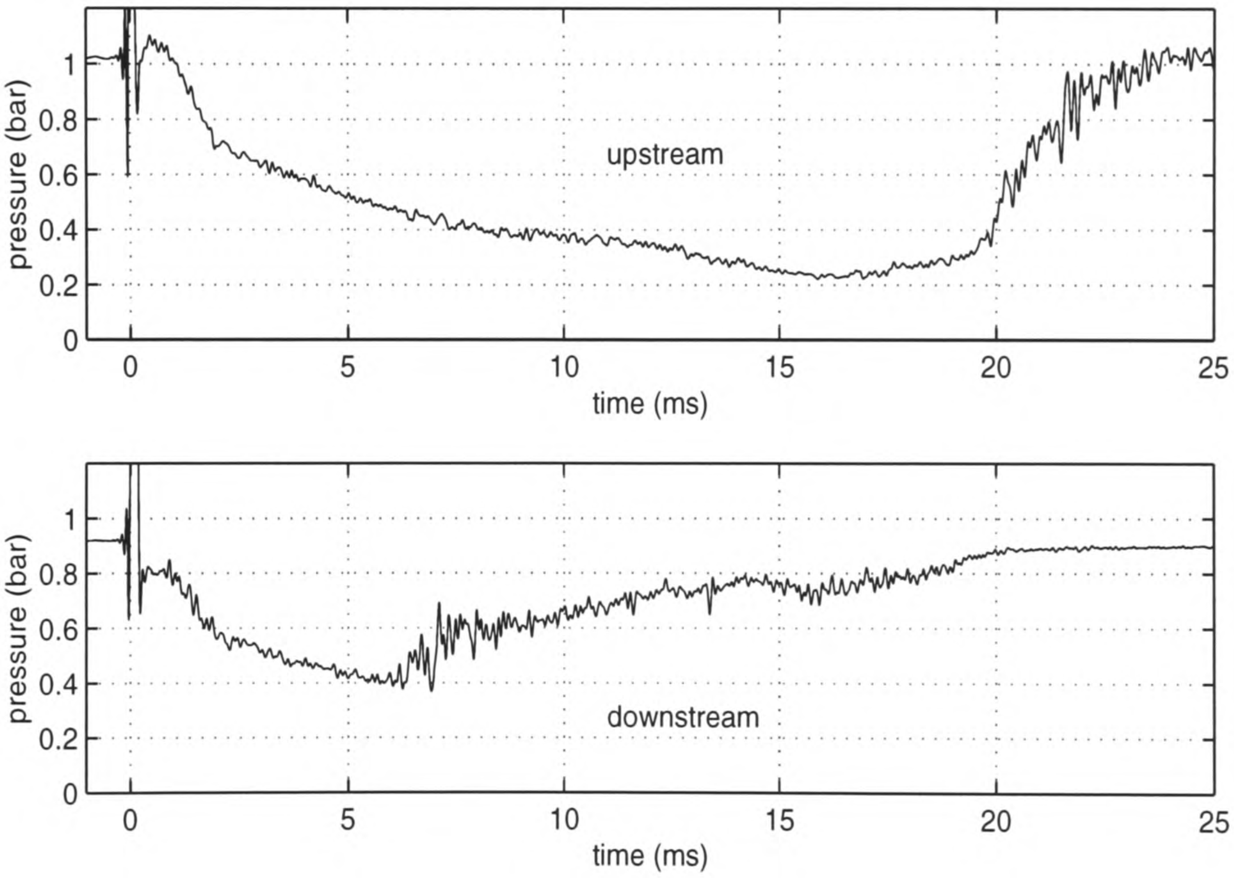


Figure 4.2 Unsteady pressure measured at the upstream and downstream transducers in the Mach 3.5 contoured nozzle, with 60 bar fill pressure and 23 µm membranes The plots are based on the same data as ***Figure 4.1***, shown on a different scale and subjected to a 10 kHz low-pass filter for clarity.

These observations provide the basis for a description of the flow in the nozzle. As the theoretical discussions and literature survey in Chapter 2 suggested, the flow history consists of a strongly unsteady starting process followed by a period of relatively slowly-varying flow. The initial sharp pressure rise is the starting or primary shock wave, which originates at bursting of the downstream membrane and propagates down the nozzle. This shock initiates the flow, and it is followed by the system of waves which constitute the starting process, bringing about a transition from the flow behind the starting shock to the subsequent slowly-varying flow. The starting process itself is examined more closely in §4.1.3. After the starting process, the nozzle pressure displays the behaviour expected of a classical convergent-divergent nozzle flow. Flow is subsonic in the convergent part of the nozzle, sonic in the throat, and supersonic in the divergent region. This decay in static pressure which can be observed at both transducers reflects the decay of total pressure as the reservoir system drains.

The nozzle flow is overexpanded (that is, at a lower pressure than the atmosphere) throughout the duration of quasi-steady flow, and must be compressed to atmospheric pressure. This requires a shock wave or a system of shocks, which may be located in the nozzle, or outside, in the structure of an overexpanded jet. As total pressure falls, the shock system migrates upstream through the nozzle, arriving at the downstream transducer at $t = 7$ ms. It traverses the downstream transducer slowly, causing local static pressure to rise there. When the data are filtered and viewed on the scale of ***Figure 4.2***, it is apparent that this process takes place in at least three gradual stages, separated by periods of approximately constant pressure. This suggests that at the downstream transducer, compression is effected by an extended system of multiple shocks of the type seen in ***Figure 2.4***. The shock structure arrives at the upstream transducer at 19 ms and traverses it relatively quickly. Subsonic flow at approximately atmospheric pressure is established downstream of the shock system as it carries on upstream.

The slowly varying part of the flow history, which persists through most of the lifetime of the flow, will be referred to throughout this thesis as *quasi-steady* flow. In the context of nozzle flow, this term implies that the reservoir pressure varies so slowly that for practical purposes, the flow at any instant is identical to the hypothetical flow which would exist under constant total pressure and back pressure. The fluctuations discussed in the next paragraph, for example, or the oscillating transonic flows described in §2.2.2, are not in conflict with this definition of quasi-steady flow because they are not driven by any unsteady input of external origin.

High-frequency fluctuations, though apparent in the signals from both transducers, are much more pronounced in the case of the downstream transducer, particularly during the final rise to atmospheric pressure. Viscous effects near the nozzle exit can provide a reasonable explanation for these fluctuations. The boundary layer should be substantially thicker at the downstream transducer than at the upstream transducer, as they are separated by 25% of the nozzle's total length. The influence of the boundary layer on the shock structure, and the possibility of flow separation with associated unsteadiness, are therefore greater at the downstream transducer. Also, the proximity of the downstream transducer to the exit would make it more accessible to unsteady viscous phenomena in the external jet, which could exert an influence on flow in the nozzle through disturbances propagating upstream via the subsonic zone within the boundary layer.

4.1.2 Repeatability

To assess repeatability of the experiment, five runs were carried out in an unvarying configuration similar to a typical operating condition. The cylinder was filled to 41.5 bar and fired with a pair of 19 μm membranes and no particle payload.

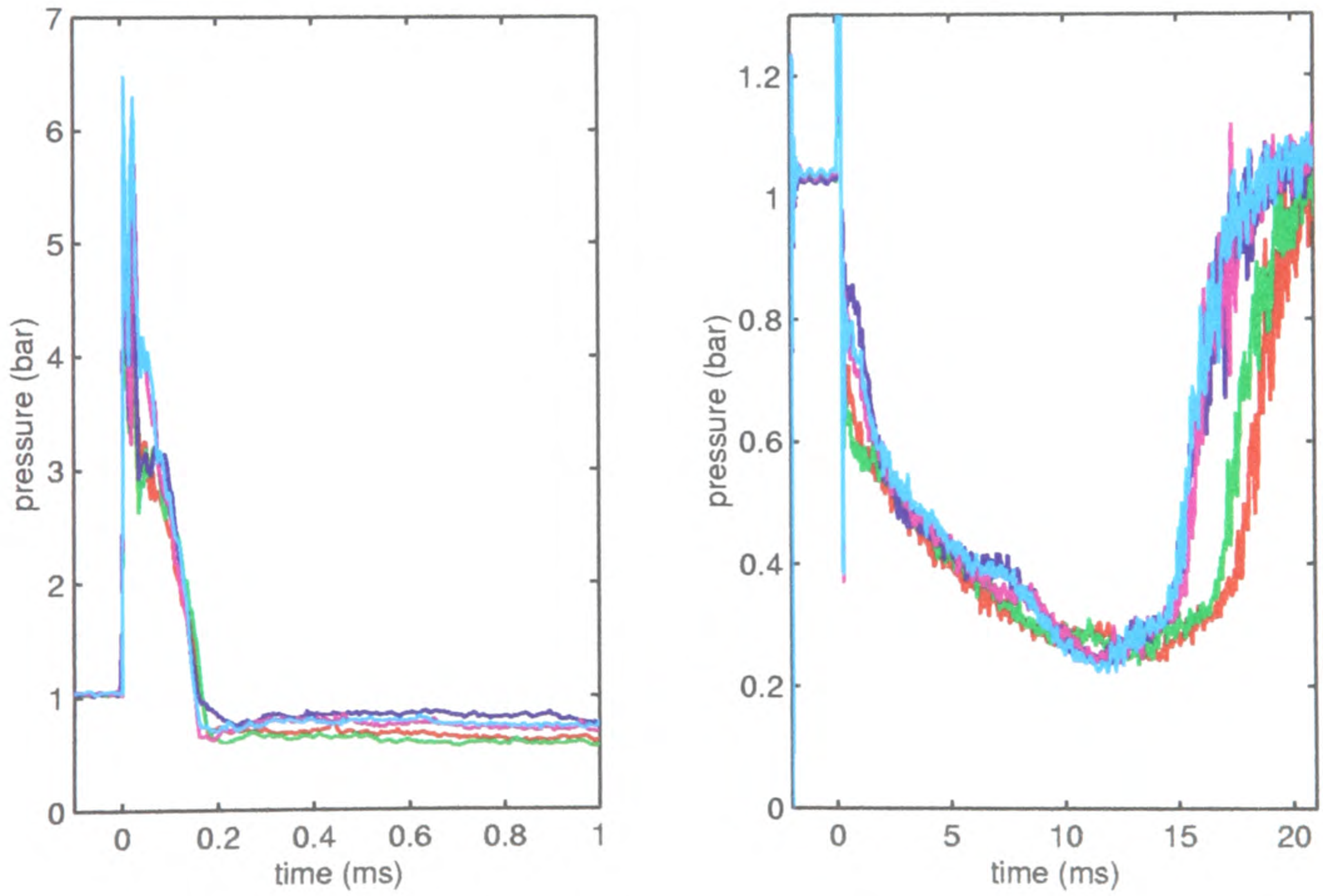


Figure 4.3 Measured pressure at the Mach 3.5 contoured nozzle upstream transducer in 5 runs with a fill pressure of 40 bar and 19 μm membranes, presented on two scales. The data in the right-hand plot have been filtered for clarity.

Figure 4.3 and ***Figure 4.4*** show the superimposed results of measurements at the upstream and downstream transducers, respectively, for the five runs. The pressure histories are qualitatively very similar. Notable variations from run to run occur in the first 0.1 ms of the starting process and in the first 1.5 ms of quasi-steady flow. These early variations may be due to variations in membrane burst strength and burst mechanics. The duration of quasi-steady supersonic flow at the downstream transducer varies by up to 3 ms. Otherwise, the time history of pressure is consistent from shot to shot.

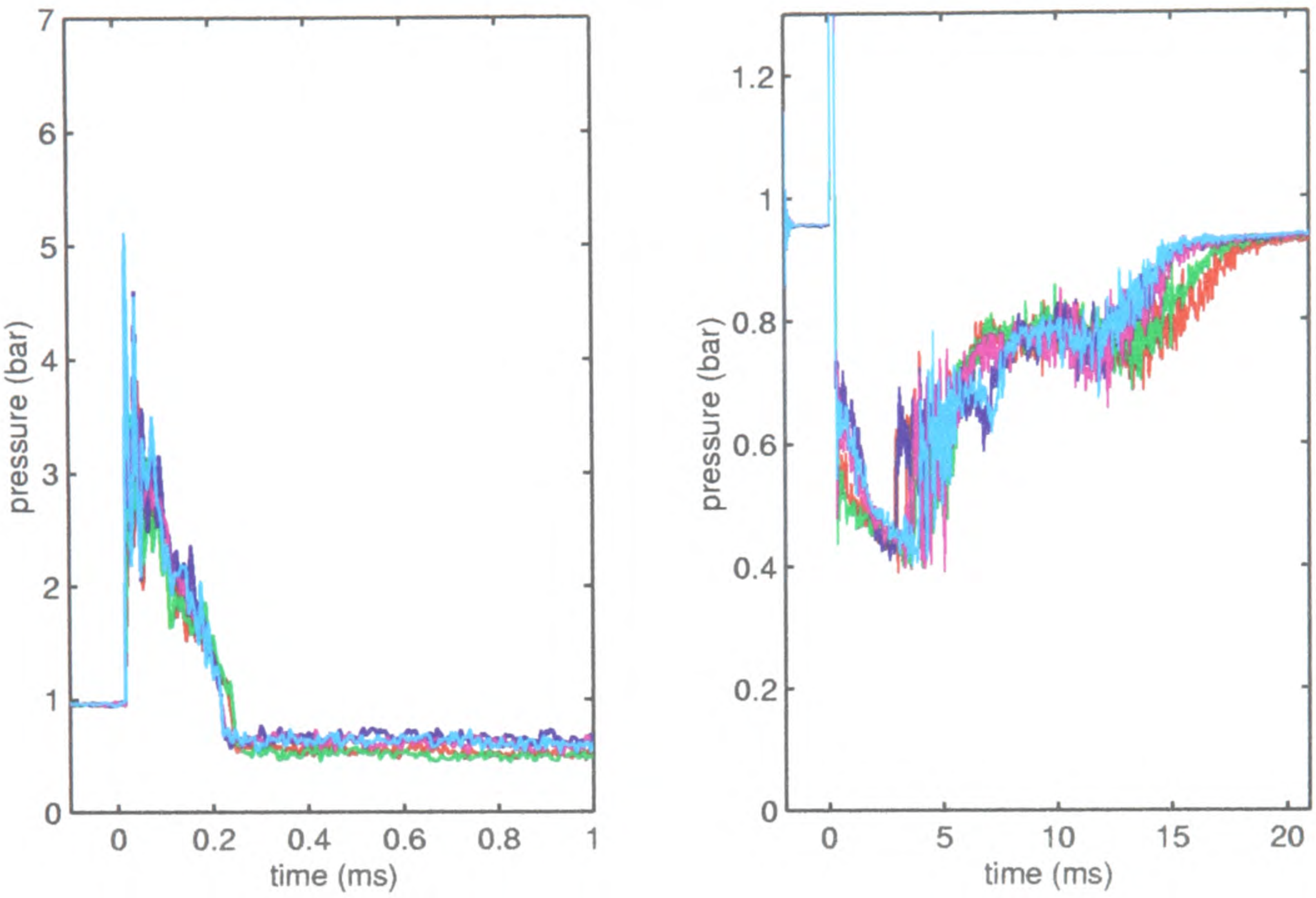


Figure 4.4 Measured pressure at the Mach 3.5 contoured nozzle downstream transducer in 5 runs with a fill pressure of 40 bar and 19 μm membranes, presented on two scales. Data in the right-hand plot have been filtered for clarity.

4.1.3 Starting Process

In this section, the strongly unsteady early stages of flow in the Mach 3.5 device are discussed with reference to a typical set of pressure measurements for that period. The propagation speed of the starting shock wave is analysed and used to evaluate the flow velocity in the starting process, which will be referred to in later discussions of particle velocity in Chapter 6. ***Figure 4.5*** shows the nozzle pressure measurements in the first 0.3 ms of flow, in a run with a fill pressure of 30 bar and 12 μm mylar membranes. These data were recorded on an oscilloscope at a sampling rate of 4 MHz, and processed in software with a 300 kHz low-pass filter. At both transducers, flow is initiated by a shock wave across which pressure increases from 1 bar to approximately 6 bar. Pressure begins to fall immediately after the passage of the shock, and finally reaches a level which is effectively steady on the time scale of this experiment, marking a well-defined end point to the starting process. The duration of the process is 160 μs at the upstream transducer, increasing to 210 μs at the downstream transducer.

The final stage of the fall in pressure is distinctly more rapid than the drop which precedes it, and it terminates abruptly on reaching the final steady pressure. This rapid pressure drop is probably the signature of a secondary shock of the type described by Smith [19,20] and others, as discussed in §2.2.3 of this thesis. The delay from the starting shock to the final pressure fall is longer at the downstream transducer than at the upstream transducer. This is consistent with the expected behaviour of the secondary shock, which would propagate upstream with respect to the primary (starting) shock. The apparent secondary shock leads immediately into quasi-steady flow, rather than giving way to a further gradual pressure decay. This suggests that the shock has completely traversed the unsteady expansion (illustrated in ***Figure 2.6***) through which gas accelerates up to its supersonic quasi-steady condition. Smith [20] observed that this only occurs when the pressure downstream of the primary shock (in this case, atmospheric pressure) is greater than the final quasi-steady pressure. This condition is satisfied in the results shown here. These results suggest that the secondary shock is indeed the closing feature of the starting process in this nozzle.

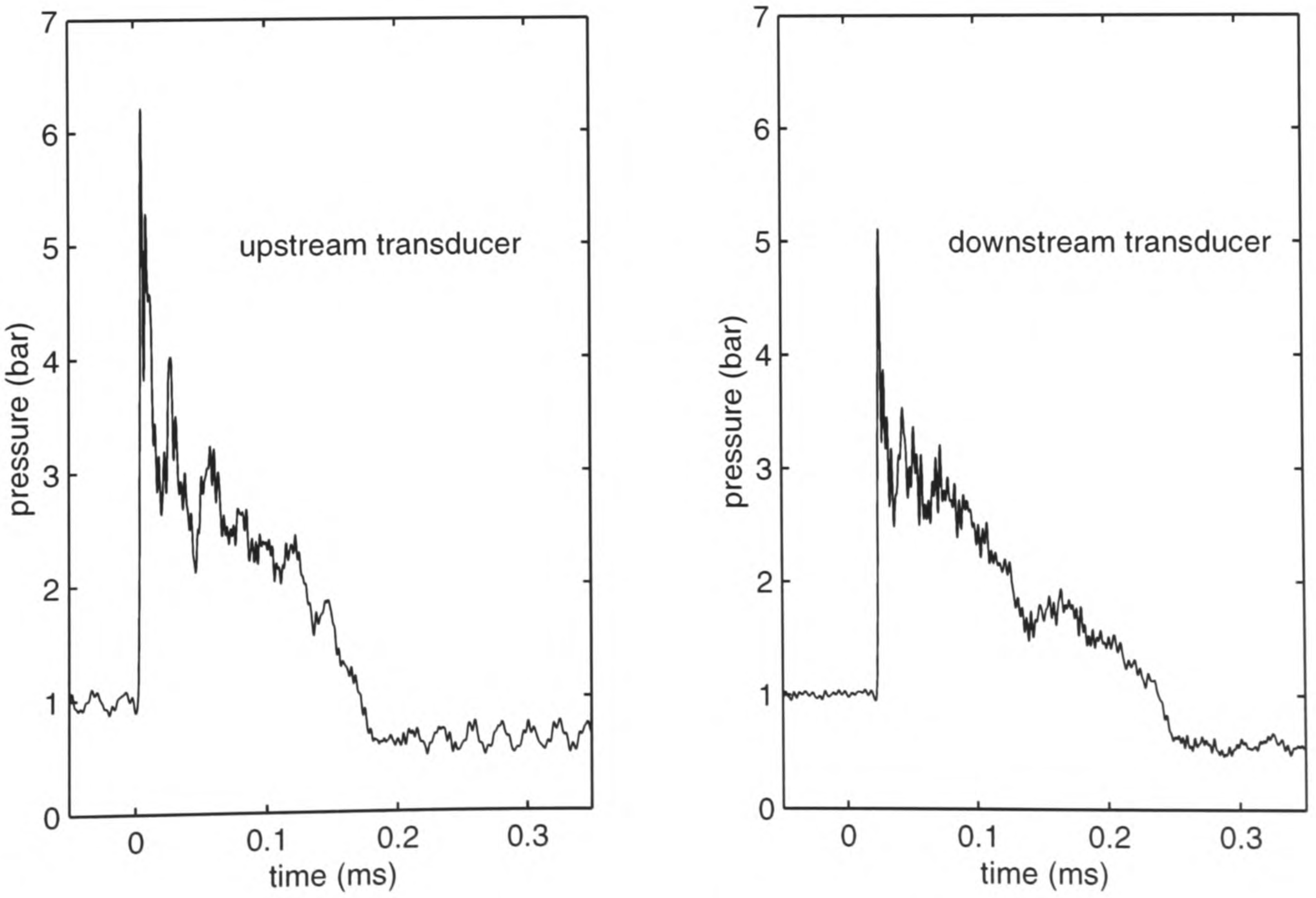


Figure 4.5 Pressure measurements in early stages of flow in the Mach 3.5 nozzle, operating with a fill pressure of 30 bar and 12 μm mylar membranes.

The speed and strength of the primary shock were studied in nozzle pressure measurements taken with mylar membranes of various thicknesses. Shock speed was calculated in three ways. It was first calculated directly from the transit time required for the shock to travel the 15 mm distance between the transducers. The results of this calculation are shown in ***Figure 4.6*** as a function of membrane burst pressure ratio, which was assessed in experiments described in §4.2.4. The finite data acquisition sampling rate, which was 200 kHz in most cases, imposes uncertainties of up to ±300 m/s on the shock speeds calculated from the transit time. These uncertainties have been calculated and shown on ***Figure 4.6*** as error bars, and are consistent with the observed scatter in measured shock speed.

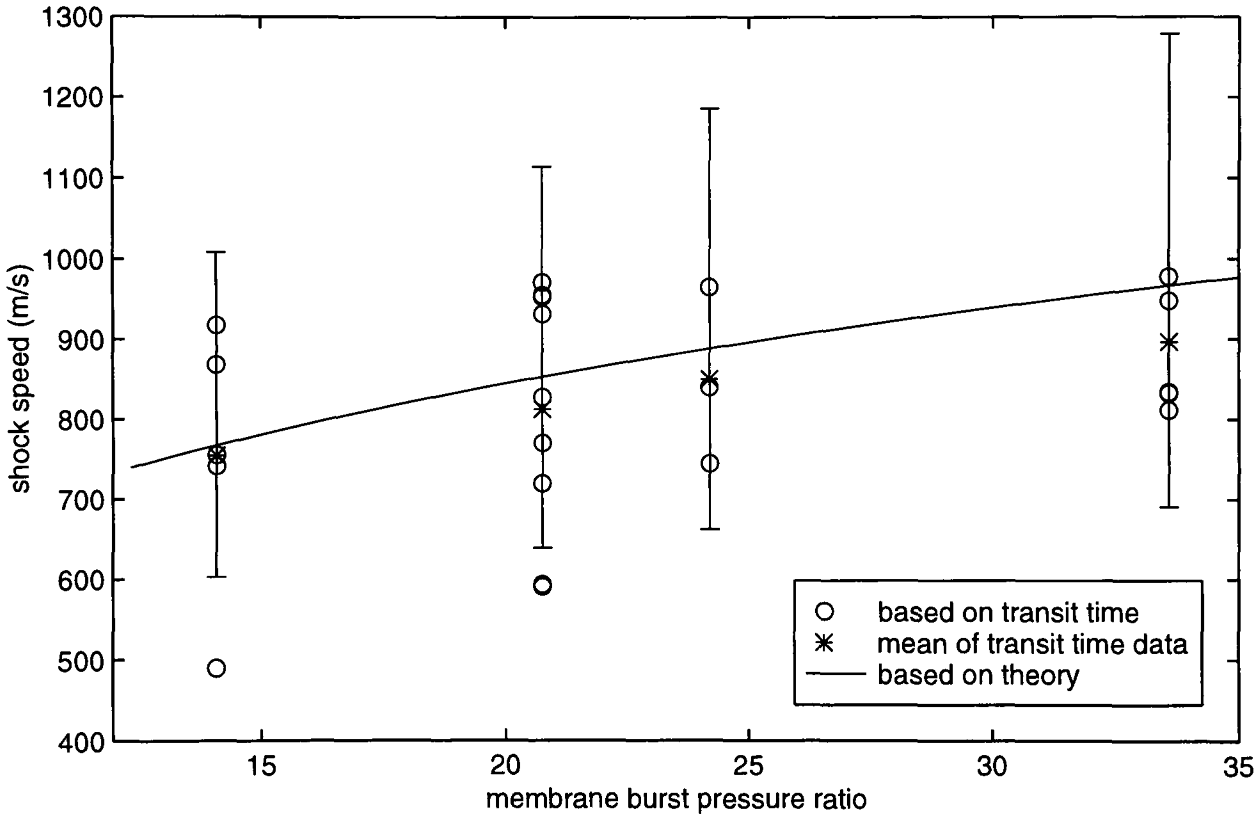


Figure 4.6 Shock speed in the Mach 3.5 nozzle, based on the transit time of the shock between the two transducers. Error bars are based on uncertainty due to finite sampling rate in data acquisition. The speed predicted by shock tube theory is also shown.

The shock speed was also calculated from the measured pressure ratio across the shock, using ***Eq. 2.2***, which relates shock Mach number to shock pressure ratio. For this purpose, the highest pressure measured at each transducer was interpreted as the pressure immediately behind the shock. The average of the shock pressure ratios at the two transducers was used to calculate an average propagation speed between the transducers. Shock speeds based on measured shock pressure ratio are shown in ***Figure 4.7***, along with mean values based on transit time, for comparison.

Finally, basic shock tube theory (specifically, ***Eqs. 2.1*** and ***2.2***) was used to predict the shock speed on the basis of membrane burst pressure ratio. These theoretical predictions are shown in both ***Figure 4.6*** and ***Figure 4.7***.

The transit time data indicate a shock speed of 755 m/s for the 12 μm membranes, increasing to 897 m/s for the 36 μm membranes. The shock speed values based on pressure ratio are scattered over a smaller range, and are 3 – 7% lower than the values based on transit time, probably because pressure behind the shock decays too quickly for the peak pressure to be accurately captured. However, the two analyses are in reasonable agreement, confirming the validity of the pressure measurements. The theoretical predictions overestimate the mean speeds based on transit time by 2 – 8%, suggesting that the influence of the nozzle profile on shock propagation is slight.

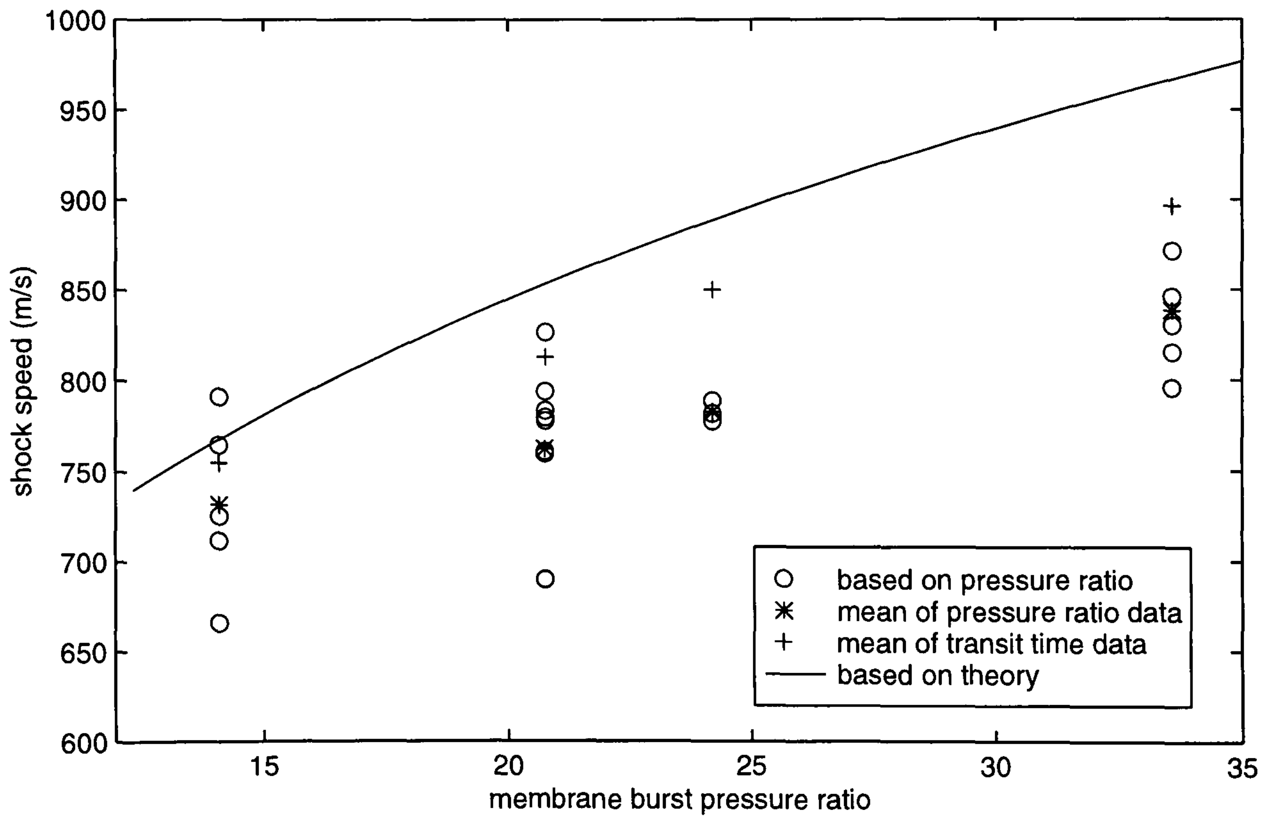


Figure 4.7 Propagation speed of the starting shock in the Mach 3.5 nozzle, based on the average pressure ratio across the shock at the two transducers. The speed predicted by shock tube theory is also shown. Speed based on transit time (averaged for each membrane type) is included for comparison.

20 μm polycarbonate membranes were used in all particle velocity measurements reported in Chapter 6. These membranes were not tested in this set of experiments, but their burst pressure ratio is known as 19.8 from measurements described in §4.2.4. With this information, it is possible to interpolate on the data presented here to predict the propagation speed of the primary shock when this type of membrane is used with the Mach 3.5 contoured nozzle. This calculation yields a shock prop-

agation speed of 810 m/s for the 20 μm polycarbonate membranes, corresponding to a bulk gas speed behind the shock of 554 m/s. This finding will be referred to in Chapter 6, in a discussion of measured particle velocities.

4.1.4 Effect of Fill Pressure and Membrane Thickness

A series of pressure measurements with various values of fill pressure and various membrane types was carried out to investigate the effects of the two most readily controllable device operating parameters. The pressure signals measured at the upstream transducer for fill pressures of 40, 50 and 60 bar, with 19 μm membranes, are shown in ***Figure 4.8***. They suggest that fill pressure has little or no effect on the starting process. However, the static pressure throughout the period of quasi-steady supersonic flow tends to increase with fill pressure. A comparison of pressures measured at the downstream transducer for various fill pressures, though not shown here, leads to the same observation. This simply reflects an increase in nozzle total pressure with fill pressure (initial cylinder pressure). The cylinder pressure would be precisely the same as the nozzle total pressure, by definition, if flow from the cylinder to the nozzle inlet was isentropic. Rupture chamber pressure measurements described in §4.2.1 and numerical modelling (§7.1.5) provide more detailed and direct information on the flow from the cylinder into the nozzle.

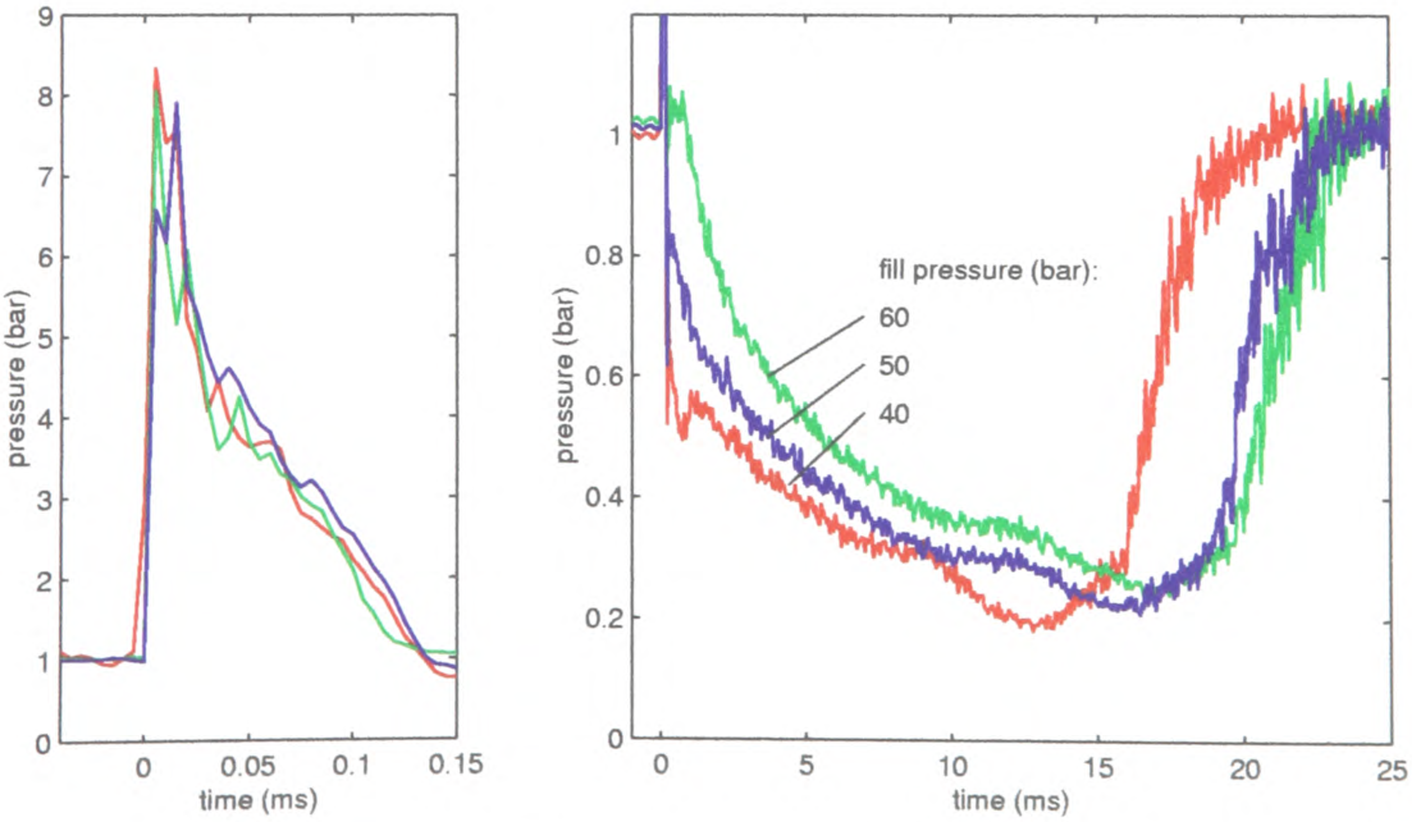


Figure 4.8 Pressure measurements at the upstream transducer in the Mach 3.5 nozzle, with 19 μm membranes and fill pressures of 40, 50 and 60 bar.

The results of upstream transducer measurements taken in three runs with different membrane types, and a common fill pressure of 50 bar, are shown in ***Figure 4.9***. The most pronounced differences occur in the starting process, in which the pressure across the starting shock increases with membrane thickness. The duration of the starting process is not affected by membrane type. However, pressure in the early part of supersonic flow, from $t = 1$ to $t = 3$ ms, appears to increase slightly with membrane thickness. The remainder of the quasi-steady flow is independent of membrane type. Overall, these data indicate that the effects of fill pressure and membrane strength are restricted to the quasi-steady flow and the starting process, respectively. Specifically, fill pressure sets the total pressure for the quasi-steady nozzle flow, while membrane strength dictates the strength (and speed) of the primary shock.

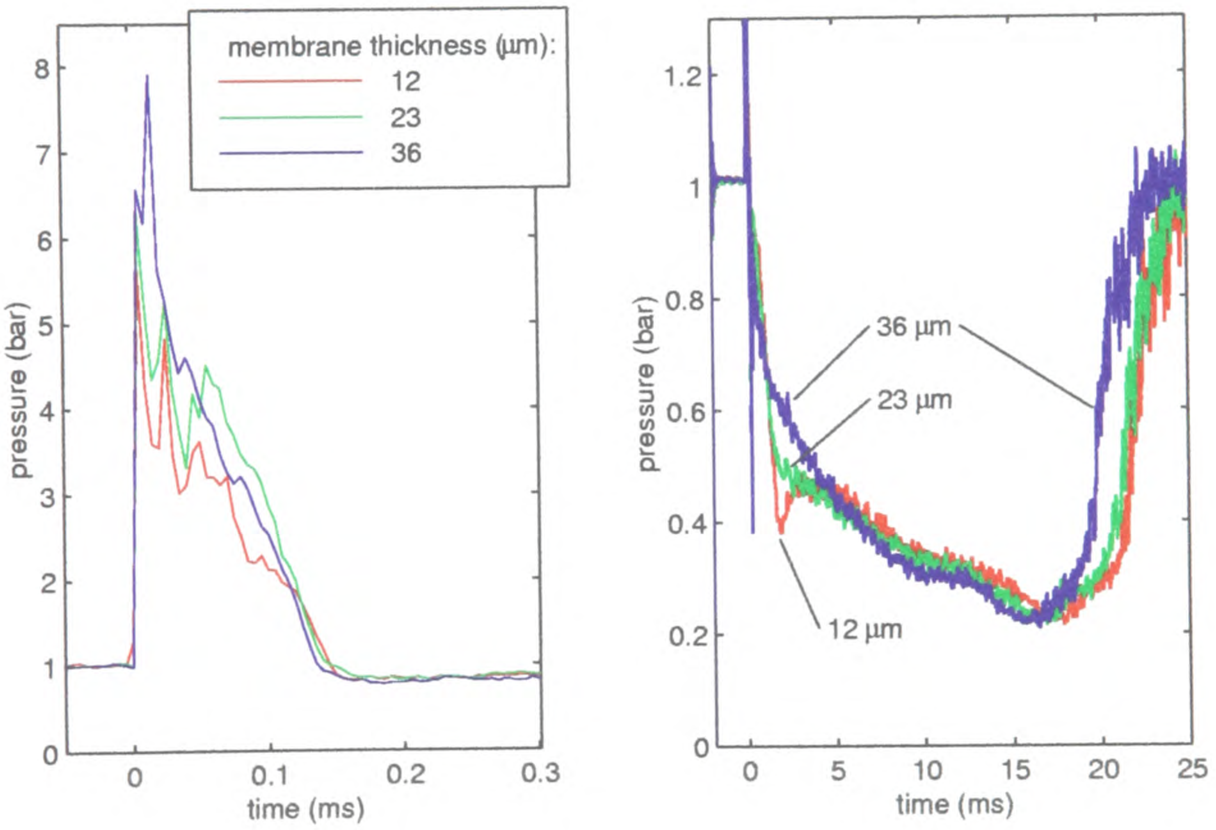


Figure 4.9 Measurements of pressure at the upstream transducer in the Mach 3.5 nozzle, with a fill pressure of 50 bar, and mylar membranes of 12, 23 and 36 μm thickness. The same results are shown on two scales, and data in the right-hand plot have been filtered for clarity.

4.1.5 Overview of Flow in the Mach 3.5 Contoured Nozzle

Flow in the contoured nozzle is opened by an unsteady shock wave, which is formed in membrane bursting and weakens slightly as it propagates through the varying nozzle area. Subsequently, a quasi-steady nozzle flow is established after a starting process of some 200 μs in duration. At first, the quasi-steady flow is supersonic throughout the divergent part of the nozzle. As the reservoir drains, total pressure falls. The static pressure in the nozzle falls, and a shock system eventually begins to move upstream through the nozzle, terminating the supersonic flow after a period of several milliseconds. The total pressure throughout the quasi-steady flow can be varied by selecting the fill pressure of the gas cylinder.

4.2 Pressure in the Conical Nozzle and the Rupture Chamber

Pressure measurements in a device fitted with a conical nozzle are described in this section. The conical nozzle was instrumented in the same way as the Mach 3.5 contoured nozzle, with two flush-mounted semiconductor transducers, located 3.25 mm and 18.25 mm upstream of the exit. In addition, a piezoelectric pressure transducer was mounted in the rupture chamber used with this nozzle. This work yielded descriptions of the flow in the conical nozzle, which is the test nozzle that most closely resembles clinical drug delivery devices, and delivered data on the functioning of the reservoir system. Many of the results described in this section were obtained by Mr. Mark Roberts as part of his final year undergraduate research project [33], working in collaboration with the author.

The arrangement of this section is similar to that of §4.1. It begins with a general description of flow in the conical nozzle, based on a discussion of a typical data set in §4.2.1. The repeatability of the flow is demonstrated in §4.2.2. Specific discussions of the early stages of flow, and the effects of fill pressure and membrane type, are given in the subsequent sections. Finally, the understanding of flow in the conical nozzle gained through these experiments is reviewed in §4.2.5.

4.2.1 A Representative Result

Data obtained from nozzle and rupture chamber transducers, on a typical run, are presented in ***Figure 4.10***. This data set displays the typical features common to all measured flow histories in the conical nozzle, and will be used here to illustrate a discussion of the general nature of the flow in this nozzle. For this particular run, the fill pressure was 60 bar and 20 μm polycarbonate membranes were used. The nozzle transducers' signals were passed through a 150 kHz low-pass filter in hardware before sampling. As mentioned in Chapter 3, values of pressure reported for the downstream transducer are not reliable, because of hysteresis observed in calibration.

The earliest activity in the system is recorded in the rupture chamber when pressure begins to rise at approximately $t = -0.25$ ms. Pressure rises to just over 20 bar, at which point the membranes burst, and gas begins to flow from the rupture chamber through the cassette and nozzle. Pressure in the rupture chamber begins to decay and approaches atmospheric pressure asymptotically. Shortly after membrane bursting, a jump in pressure is registered at both nozzle transducers, as the starting shock propagates down the nozzle. It is followed immediately by a sharp drop in pressure. After this, pressure at the upstream transducer rises gradually towards atmospheric pressure, with large decaying fluctuations. At the downstream transducer, gas appears to flow at or near atmospheric pressure from an early stage, suggesting that the flow is subsonic in that region of the nozzle.

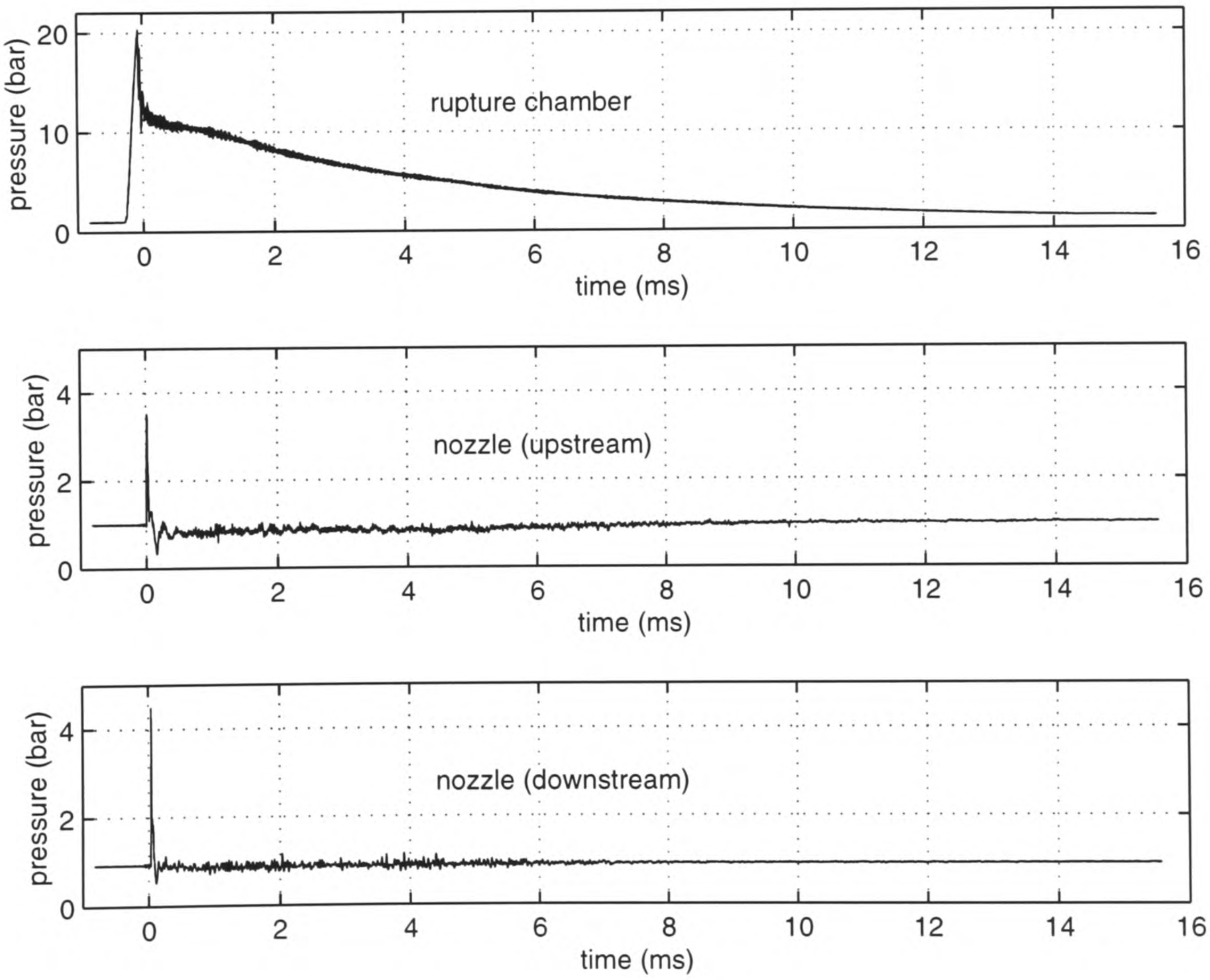


Figure 4.10 Simultaneous measurements of pressure in the rupture chamber and conical nozzle, with a fill pressure of 60 bar and 20 μm membranes.

This differs from the overexpanded supersonic flow observed in the Mach 3.5 contoured nozzle, in which static pressure at the transducers dropped below atmospheric pressure and continued to fall. In contrast, the above pressure measurements taken in the conical nozzle suggest that quasi-steady supersonic flow is never established in the region of the nozzle where the transducers are located.

This is consistent with the measured rupture chamber pressure, which does not exceed 12 bar during quasi-steady nozzle flow, whereas the minimum total pressure required for supersonic exit flow has been calculated as 17.5 bar (as stated in §2.3.1). Consequently, a shock or a shock system must be located inside the nozzle. The nozzle pressure measurements suggest that a shock system extends over a substantial axial region in the nozzle, and that the upstream transducer is located within this region. The shock system migrates upstream and weakens as the total pressure decays, causing the pressure at the upstream transducer to rise gradually towards the condition in the fully subsonic flow which exists downstream of the shock system. Further pressure measurements, presented in §4.2.4, as well as the particle velocity measurements described in Chapter 6, are consistent with this interpretation.

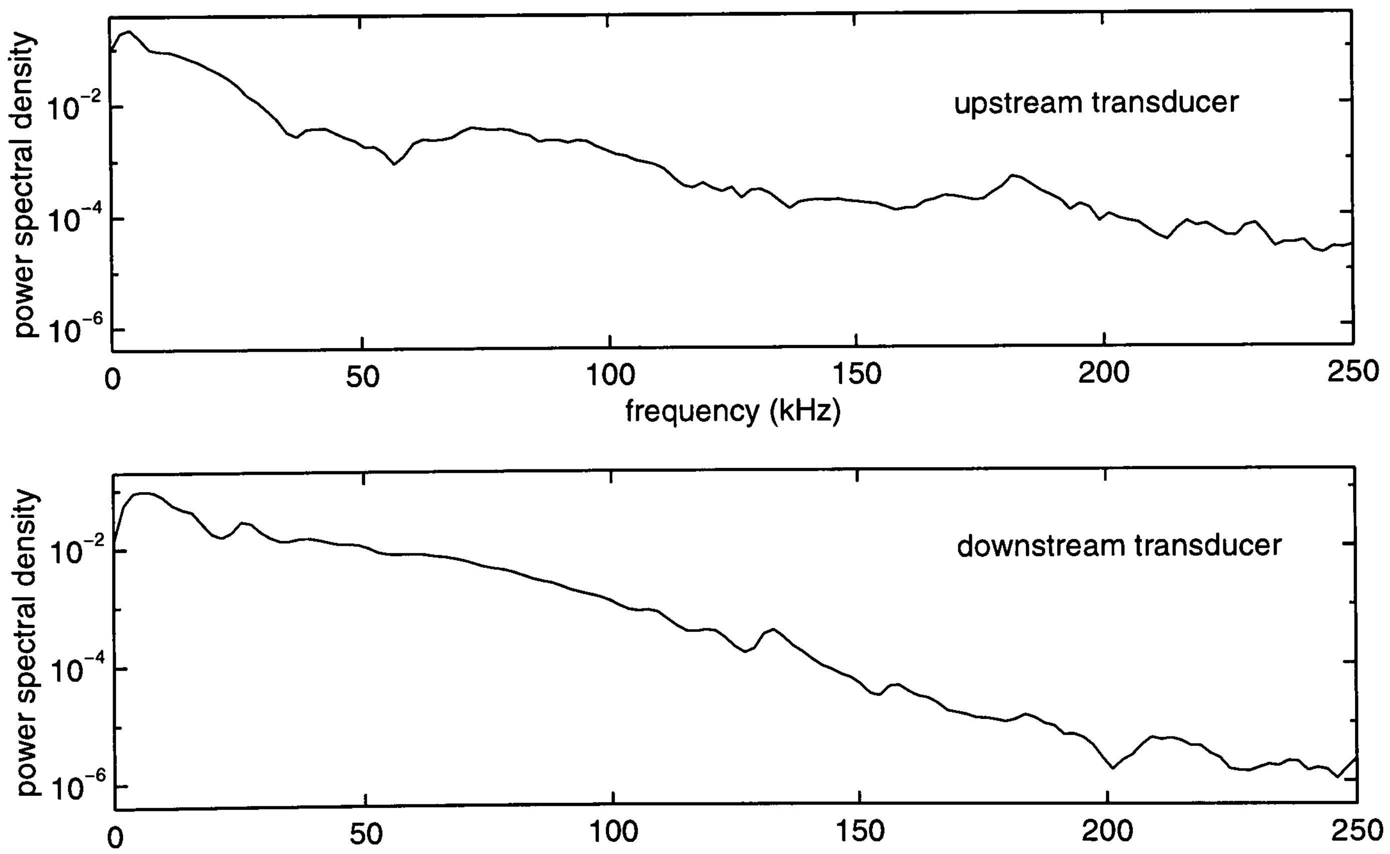


Figure 4.11 Power spectral densities of the pressure data recorded from transducers in the conical nozzle, with a 150 kHz hardware filter and 500 kHz sampling.

Pressure fluctuates rapidly during the quasi-steady flow at both transducers. The power spectral densities of both pressure signals are shown in ***Figure 4.11***, indicating that these fluctuations possess well-defined characteristic frequencies. Pressure at the upstream transducer oscillates at approximately 4 kHz, while pressure at the downstream transducer oscillates at 4 – 6 kHz. As described in Chapter 2, oscillations in transonic flow are a well known consequence of the interaction of shock waves and boundary layers, with associated flow separation. Such oscillation, associated with the shock system discussed above, may well be the cause of the fluctuating pressure in the conical nozzle.

4.2.2 Repeatability

Five runs at nominally identical conditions were carried out to assess repeatability of the conical nozzle flow-field and its measurement. Results for the upstream transducer are shown in ***Figure 4.12***. The pressure history appears to be quite consistent. In particular, the amplitude and timing of oscillations which occur in the first 1 ms of flow are highly repeatable. Measurements at the downstream transducer display a similar degree of repeatability.

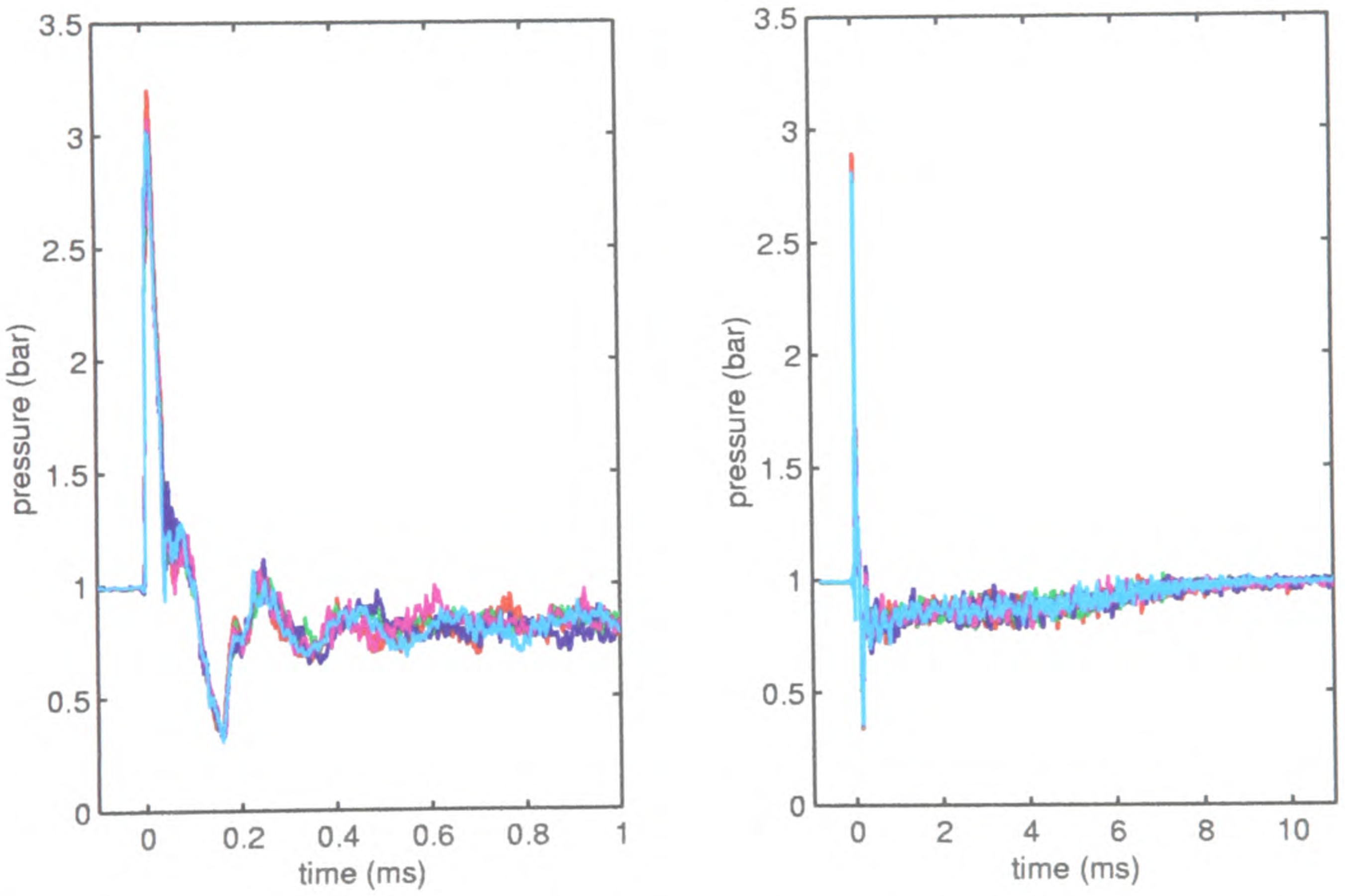


Figure 4.12 Measurements of pressure at the upstream transducer in the conical nozzle in five runs with 60 bar fill pressure and 20 μm polycarbonate membranes. The data shown in the right hand plot have been filtered for clarity.

4.2.3 Starting Process

Pressure measurements in the early stages of flow in a device with a conical nozzle are shown in ***Figure 4.13***. The starting processes in both the rupture chamber and the nozzle are resolved more clearly on these time scales. As the rupture chamber fills, pressure reaches a maximum and then drops sharply when an unsteady expansion wave propagates upstream over the transducer. This wave is formed at bursting of the upstream membrane, and it initiates the expansion of gas in the rupture chamber into the volume which was initially sealed off between the membranes. As this expansion process subsides, inflow from the cylinder begins to dominate and pressure in the (enlarged) rupture chamber rises again. A second sharp drop occurs as the downstream membrane bursts, generating another expansion wave. Pressure then rises briefly while a balance between inflow from the cylinder and outflow through the nozzle is established, and a long-term gradual monotonic decay begins.

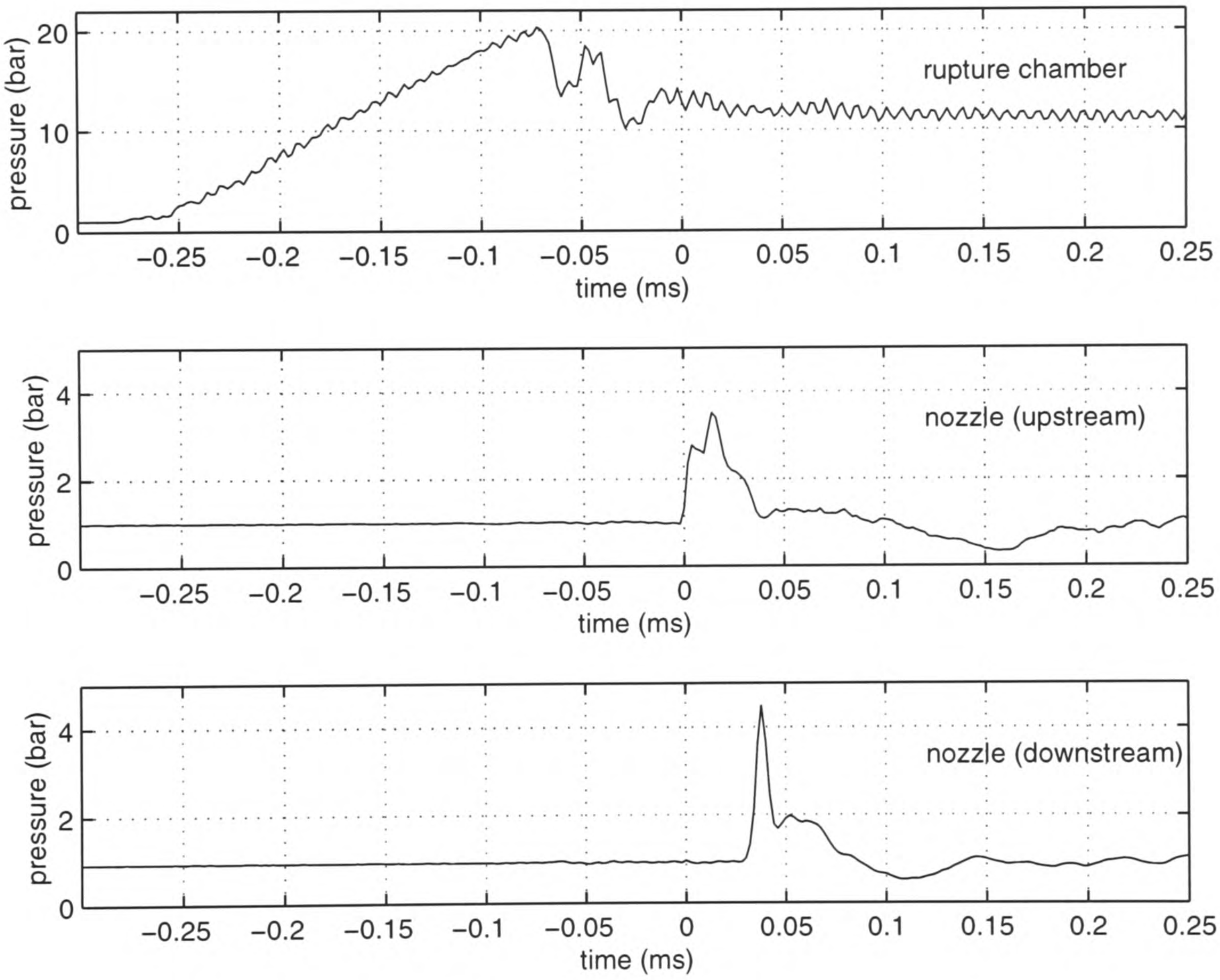


Figure 4.13 Simultaneous measurements of pressure in the rupture chamber and conical nozzle, with a fill pressure of 60 bar and 20 μm membranes.

Flow in the conical nozzle is started by the primary shock wave, followed immediately by the pressure decay required to reach the quasi-steady condition. In contrast with the flow in Mach 3.5 nozzle, there is no clearly identifiable secondary shock; instead, the pressure decay after the passage of the primary shock is erratic, and leads into the fluctuating quasi-steady flow. Thus, the duration of the starting process is not well-defined, but it is of the order of 200 μs.

In order to estimate the flow velocities in the starting process, the speed of the primary shock was calculated from nozzle pressure measurements taken with a mylar membranes of a range of thicknesses. Shock speed was evaluated empirically from the transit time of the shock over the known distance between the transducers, and, using ***Eq. 2.2***, from the pressure ratio measured across the shock at the upstream transducer. For this purpose, the shock was assumed to be represented by the first part of the two-stage pressure rise observed at the upstream transducer. This choice was borne out by the results of the analysis, as explained below. The shock propagation speeds were also calculated from both standard constant-area shock tube theory, and from the standard theory augmented by Chisnell's approximation [17] for shock propagation in a duct of non-uniform area. These calculations predicted the propagation speed as a function of the membrane burst pressure, which was available from the rupture chamber measurements recorded directly in each experiment.

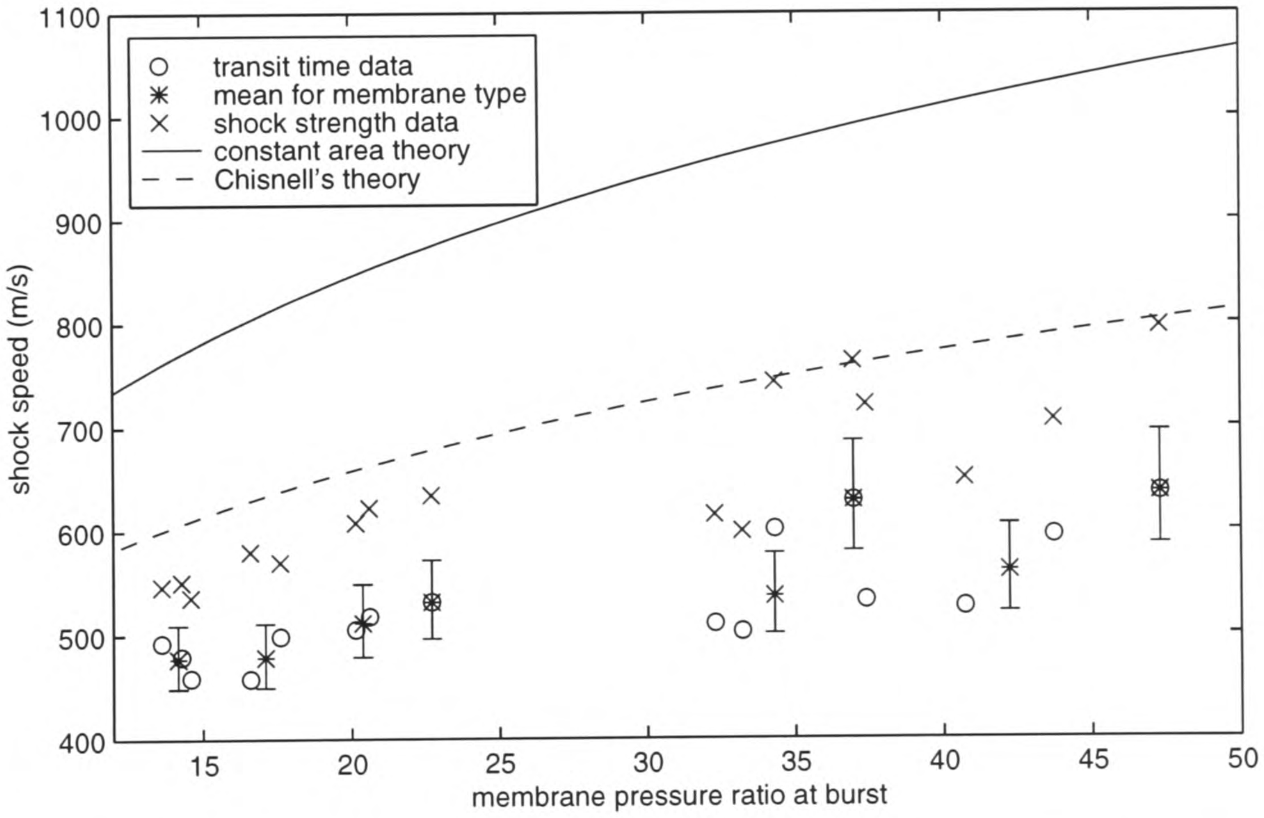


Figure 4.14 Shock propagation speed, based on the shock's transit time between transducers, on the shock strength, and on theory, shown as a function of the measured membrane burst pressure. The mean speed based on transit time is plotted for each type of membrane tested, with an error bar indicating the uncertainty due to finite sampling rate.

The results are shown in ***Figure 4.14***. All four versions of the shock speed increase with membrane burst pressure ratio, as expected. The speeds based on transit time range from 477 m/s for the 12 μm mylar membranes to 636 m/s for the 75 μm membranes. Again, the shock speeds based on transit time are scattered because of the finite sampling rate used in the experiments. The transit time values are 60 – 200 m/s less than the speeds inferred from the shock pressure ratio. This discrepancy is at least partly due to the fact that the latter values are based on measurements at the upstream transducer alone, since shock pressure ratio is accurately known only for that transducer. The speed calculated from pressure information is therefore biased towards the upstream transducer location, where the cross-sectional area is smaller and the shock is stronger and faster. The speed based on transit time, on the other hand, is effectively an average over the distance between the transducers. However, the fact that shock speed based on pressure ratio is in reasonable agreement with calculations based on transit time suggests that it was correct to interpret the initial pressure rise alone as the complete shock wave. If the absolute maximum pressure was taken as the pressure behind the shock, greater discrepancies between the two calculations of shock speed would result.

Conventional theory for constant area shock tubes overpredicts the measured shock speed (based on transit time) by as much as 450 m/s. Chisnell's theory for shock propagation through a duct of variable area gives a considerable improvement, overpredicting by 120 – 200 m/s or 25 – 36%. By contrast, in the Mach 3.5 nozzle (which has a smaller exit area), constant area shock tube theory gave a reasonable estimate of shock speed. Although Chisnell's approximation is not a complete model of the shock propagation process, its relatively close agreement with measurement suggests that the large area changes in this nozzle have a significant influence on the primary shock.

The shock speed associated with 20 μm polycarbonate membranes has been calculated (from the data based on shock transit time) as 482 m/s. The gas velocity immediately behind a shock of this speed is 195 m/s. In §4.4, measurements of shock propagation speed and calculations of the gas velocity behind the shock are presented for flows with particle payloads.

4.2.4 Effect of Fill Pressure and Membrane Thickness

Pressure measurements in the conical nozzle and in the attached rupture chamber have been carried out for a range of combinations of membrane type and fill pressures. This investigation offers an opportunity to characterise the response of the flow in the device to changes in these operating parameters, and provides some further information on the basic characteristics of the flow.

Plots of measured pressure in the rupture chamber over the first few ms of flow, for various membrane thicknesses and a common fill pressure of 60 bar, are shown in ***Figure 4.15***. As expected, the maximum pressure, which occurs immediately before bursting of the upstream membrane, increases as the thickness of the membranes is increased. For the two thickest membranes, the pressure required for bursting approaches the fill pressure, and the time required to complete bursting of both membranes increases dramatically. When the strongest membranes are used, pressure rises very slowly after bursting of the upstream membrane, and falls for a significant period in the case of the 36 μm membrane. This is partly due to the fact that mass flow into the rupture chamber falls with the pressure ratio across the valve as the cylinder drains and the rupture chamber fills. Also, the slow filling might allow time for heat transfer from the gas to the device walls, after the gas temperature was initially raised by the compressive action of gas flowing in from the cylinder. This effect would tend to reduce pressure in the rupture chamber.

In all cases, the downstream membrane bursts at a lower pressure than the upstream membrane. The difference is very large for the two thickest membranes, which undergo prolonged loading at a roughly constant pressure. Apparently, the history of membrane loading, as well as the magnitude of the load, influences the burst pressure.

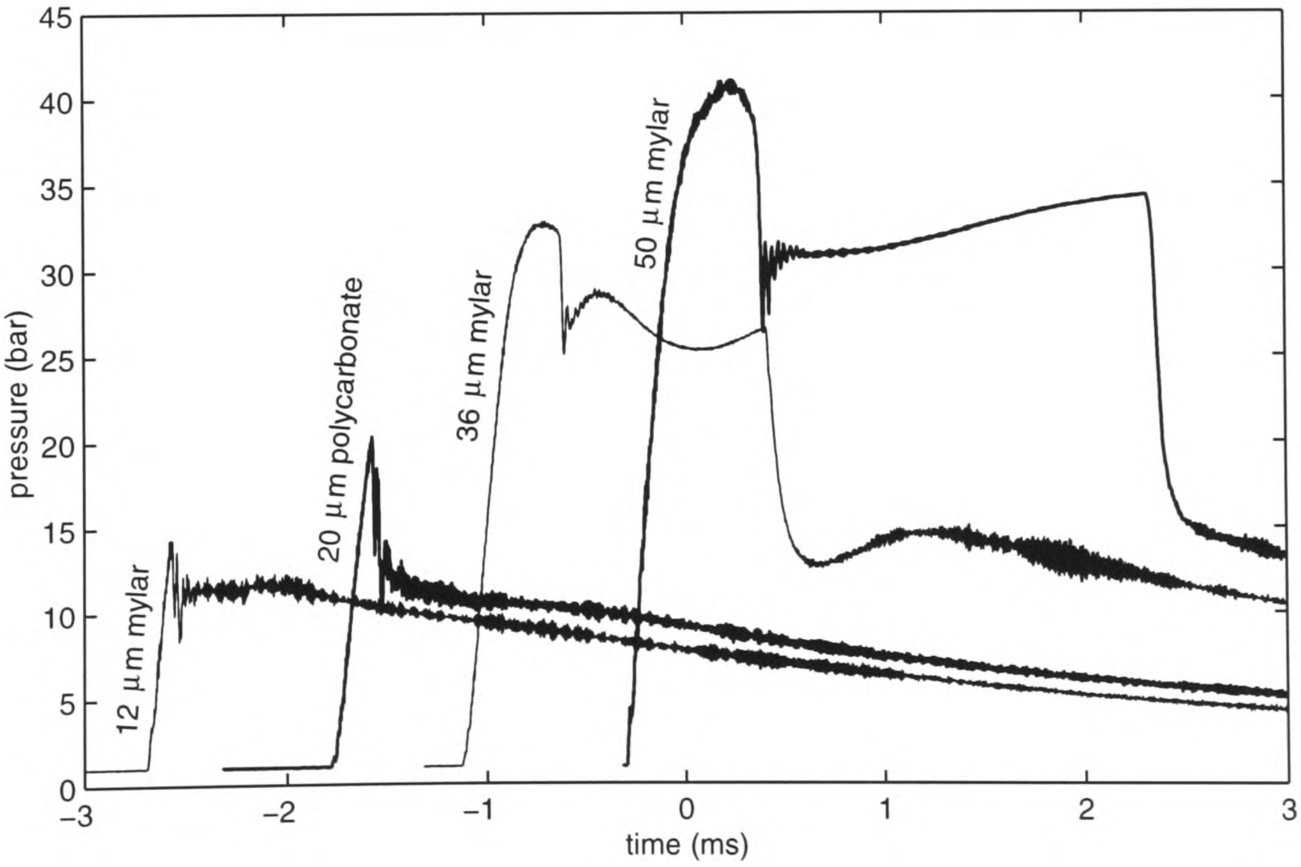


Figure 4.15 Measurements of rupture chamber pressure for fill pressure of 60 bar and four membrane types: 12 μm Mylar (marked M12), 20 μm polycarbonate (PC20), 36 μm mylar (M36) and 50 μm Mylar (M50). The data are arbitrarily shifted in time, for clarity.

The rupture chamber pressure measurements for these four membrane types are shown in full in ***Figure 4.16***. The decay of pressure after bursting follows very similar paths for the 12 μm mylar and 20 μm polycarbonate membranes. Pressure after rupture of the 50 μm mylar membranes tends towards a similar profile after some 3 ms, though it is initially higher. The flow after rupture of the 36 μm mylar membrane appears to be exceptional, with pressure remaining consistently higher than for any other membrane.

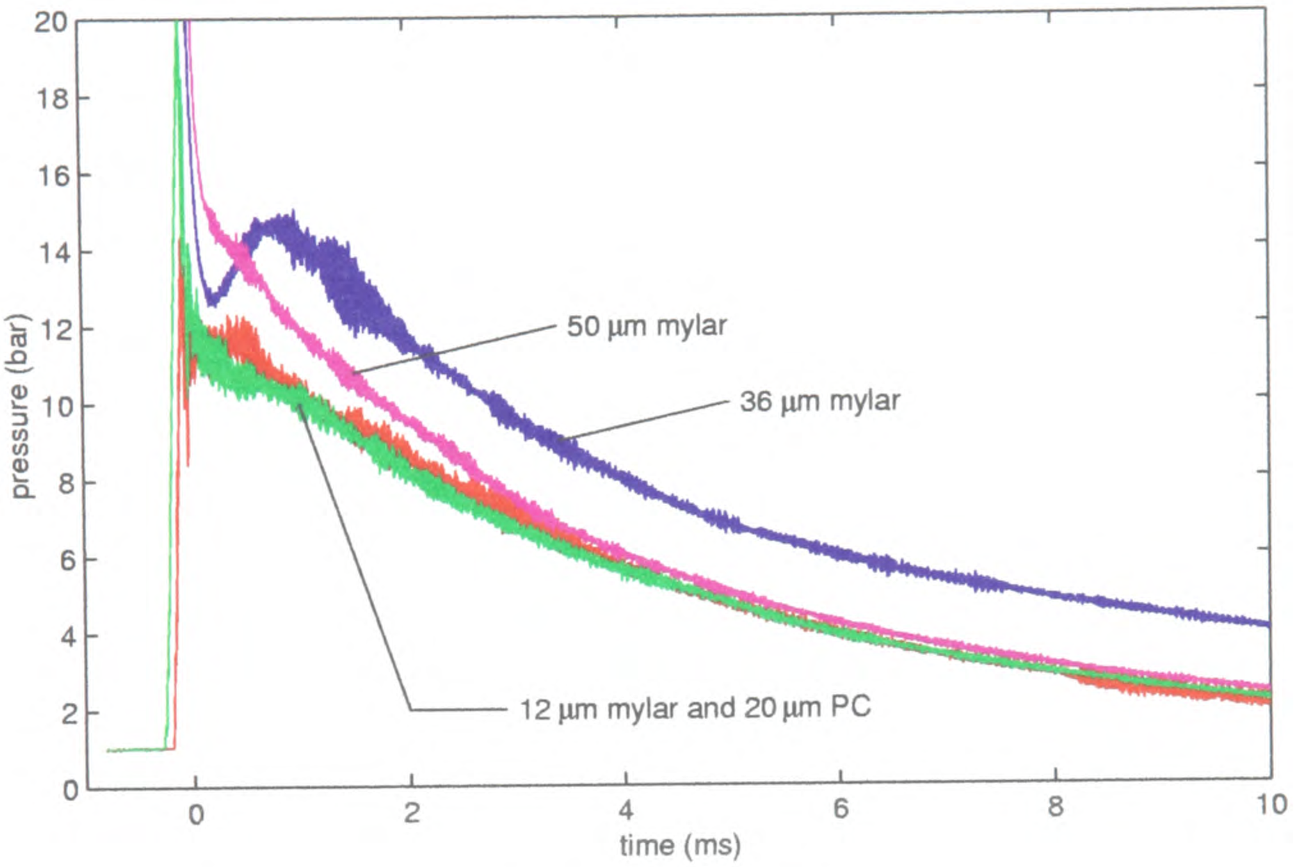


Figure 4.16 Measurements of rupture chamber pressure for fill pressure of 60 bar and four membrane types: 12 μm Mylar (marked M12), 20 μm polycarbonate (PC20), 36 μm mylar (M36) and 50 μm Mylar (M50).

Rupture chamber pressure measurements such as those described above have been used to compile a database of membrane burst pressures. The maximum rupture chamber pressure recorded in typical runs at 60 bar fill pressure (effectively, the burst pressure of the upstream membrane) was taken as the ***operating strength*** of the membrane. In a separate set of experiments, each membrane was clamped and gradually pressurised over a period of several seconds until it burst. The burst pressure in this scenario is referred to as the ***static strength.*** All the data are plotted in ***Figure 4.17***. In static loading, the burst pressure of mylar membranes is approximately proportional to their thickness. In device operation, membranes up to 20 μm thick have a burst pressure which is proportional to thickness and slightly higher than their static strength. The operating strength of the thicker membranes is less than a linear dependence on thickness would predict. This trend cannot entirely be explained by the prolonged loading which the thickest membranes encounter, since they tolerate higher pressures under static loading than in device actuation. It may be that the bursting process is influenced by the combination of dynamic load, prolonged static load and perhaps temperature swings which membranes experience in device operation. The 20 µm polycarbonate membrane, which is the most commonly used in practice, is slightly weaker than mylar membranes of similar thickness.

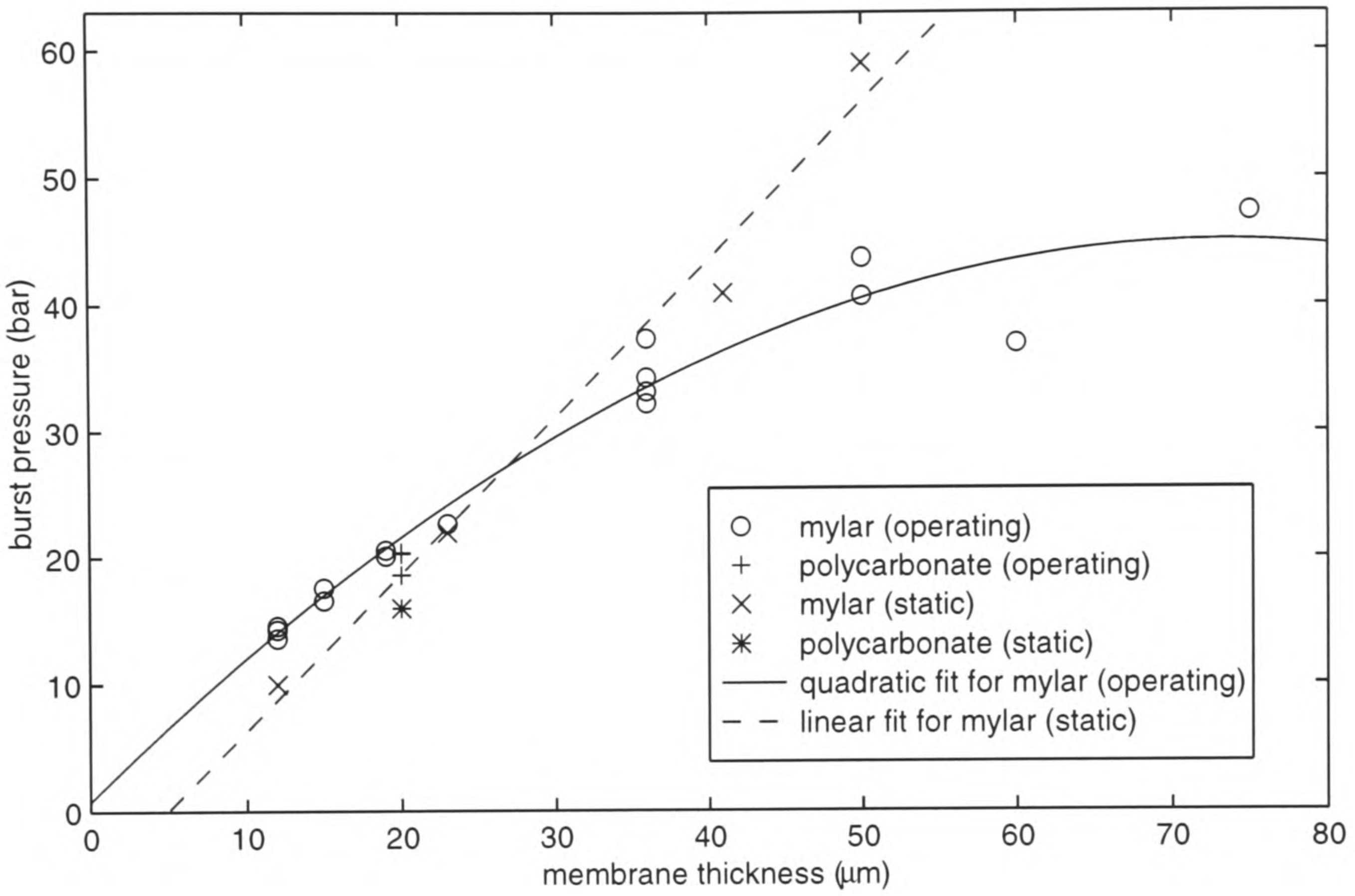


Figure 4.17 Burst pressure of various membrane types, measured in the rupture chamber in operation with a conical nozzle.

The influence of membrane burst strength on flow in the nozzle is evident in the pressure measurements shown in ***Figure 4.18***, which were taken with four different membrane types. The data were acquired in the same experiments as the rupture chamber data shown in ***Figures 4.15*** and ***4.16***. As expected, increase in membrane thickness results in an increase in the strength of the starting shock (taken as the first part of the two-stage compression). In the case of the polycarbonate membranes, the compression wave which follows the starting shock is stronger than that observed with any of the mylar membranes. This suggests that this second compression wave depends on membrane material properties and the detailed mechanics of the bursting process. At times more than 100 µs after the starting shock, the pressure is independent of membrane type. The influence of the membrane, therefore, is confined to the starting process.

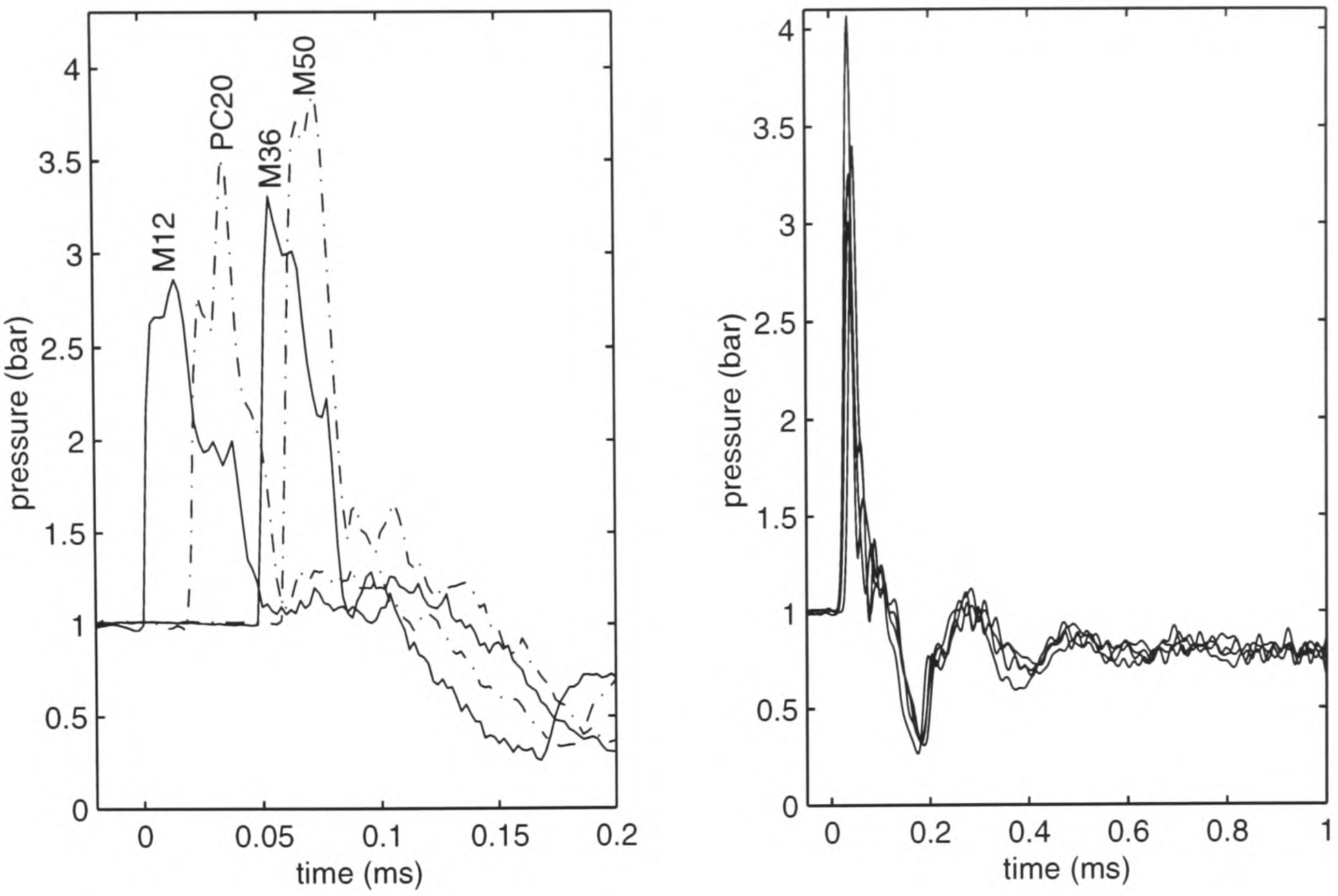


Figure 4.18 Plots on two time scales of pressure measured at the upstream transducer in the conical nozzle for 60 bar fill pressure and four membrane types: 12 μm mylar (marked M12), 20 μm polycarbonate (PC20), 36 μm mylar (M36) and 50 μm mylar (M50). For clarity, the plots in the left-hand graph have been arbitrarily displaced in time, and those on the right have been filtered.

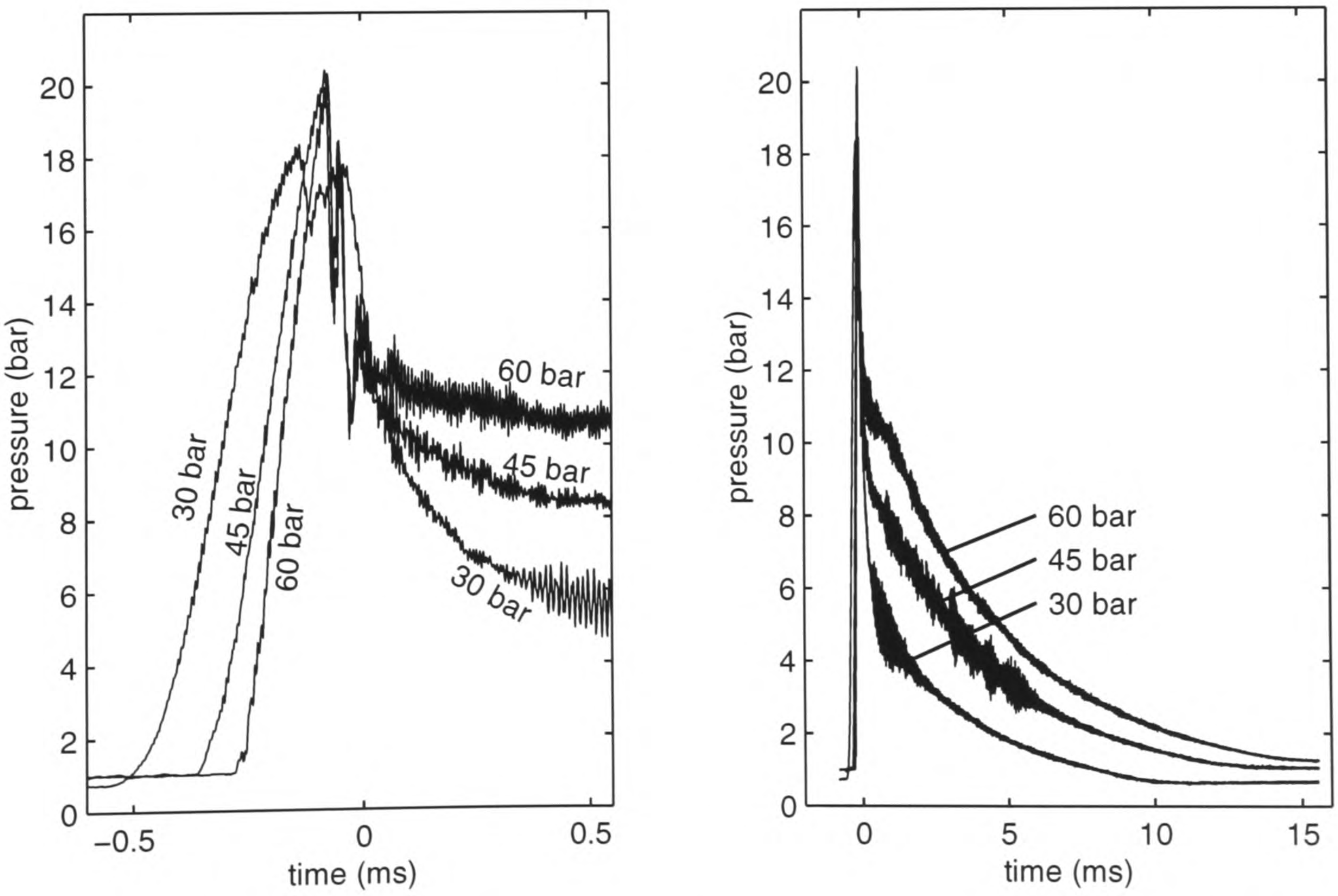


Figure 4.19 Time histories of measured rupture chamber pressure for fill pressures of 30, 45 and 60 bar, with 20 μm polycarbonate membranes and a conical nozzle, shown on two scales. Annotations indicate the fill pressure.

The effects of fill pressure on the flow history will be discussed in the remainder of this section. Measurements taken in the rupture chamber in runs with fill pressures of 30, 45 and 60 bar are shown in ***Figure 4.19***. Fill pressure appears to have a direct and simple influence on the pressure throughout the system. At the higher fill pressures, the rupture chamber pressurises more quickly before membrane bursting (since mass flow into the rupture chamber increases with cylinder pressure) and pressure is maintained at a higher level after bursting. However, the rupture chamber pressure at the start of quasi-steady flow is significantly less than the fill pressure (the initial cylinder pressure) in each case. This points to major losses in total pressure in the flow between the cylinder and the rupture chamber, which represent inefficiency. Numerical modelling predictions, presented in Chapter 7, provide more insight into this aspect of the flow.

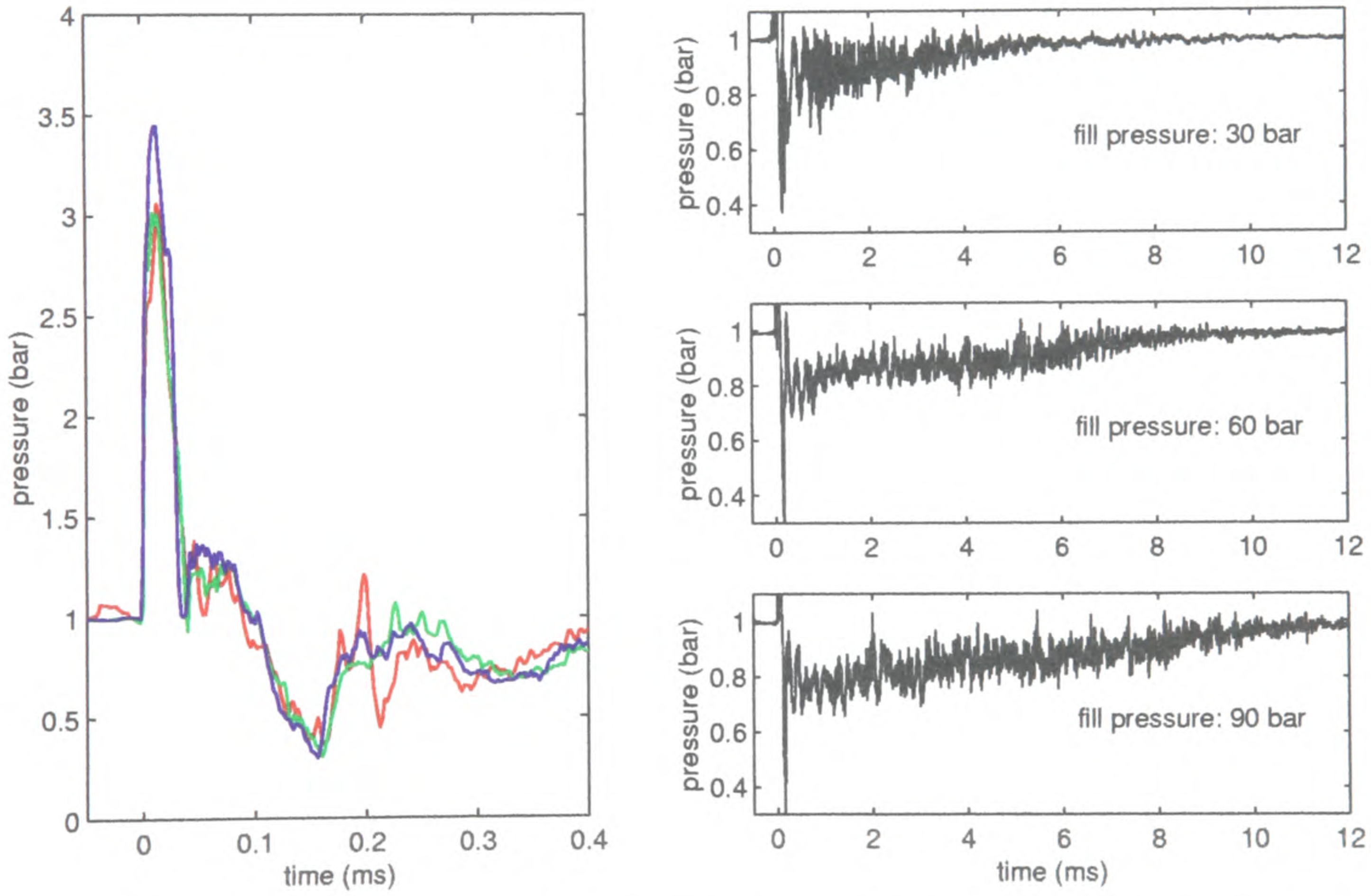


Figure 4.20 Pressure measured at the upstream nozzle transducer for fill pressure values of 30, 60 and 90 bar, with 20 μm polycarbonate membranes. Early segments of each of the three results are overlaid in the graph on the left; the graphs on the right show a portion of each run individually.

Nozzle pressure measurements, at fill pressures of 30, 60 and 90 bar, with 20 μm polycarbonate membranes, are shown in ***Figures 4.20*** and ***4.21***. No effect due to fill pressure is apparent in the early stages of the pressure histories at either nozzle location. After approximately $t = 0.2$ ms, however, pressure at the upstream transducer displays the influence of fill pressure. As the fill pressure is increased, pressure throughout the quasi-steady nozzle stage of flow becomes lower, and the duration of flow becomes longer. At the downstream transducer, the mean level of pressure appears to be unaffected by increases in fill pressure. The duration of the 4-6 kHz oscillatory activity grows longer, reflecting the increased lifetime of the flow.

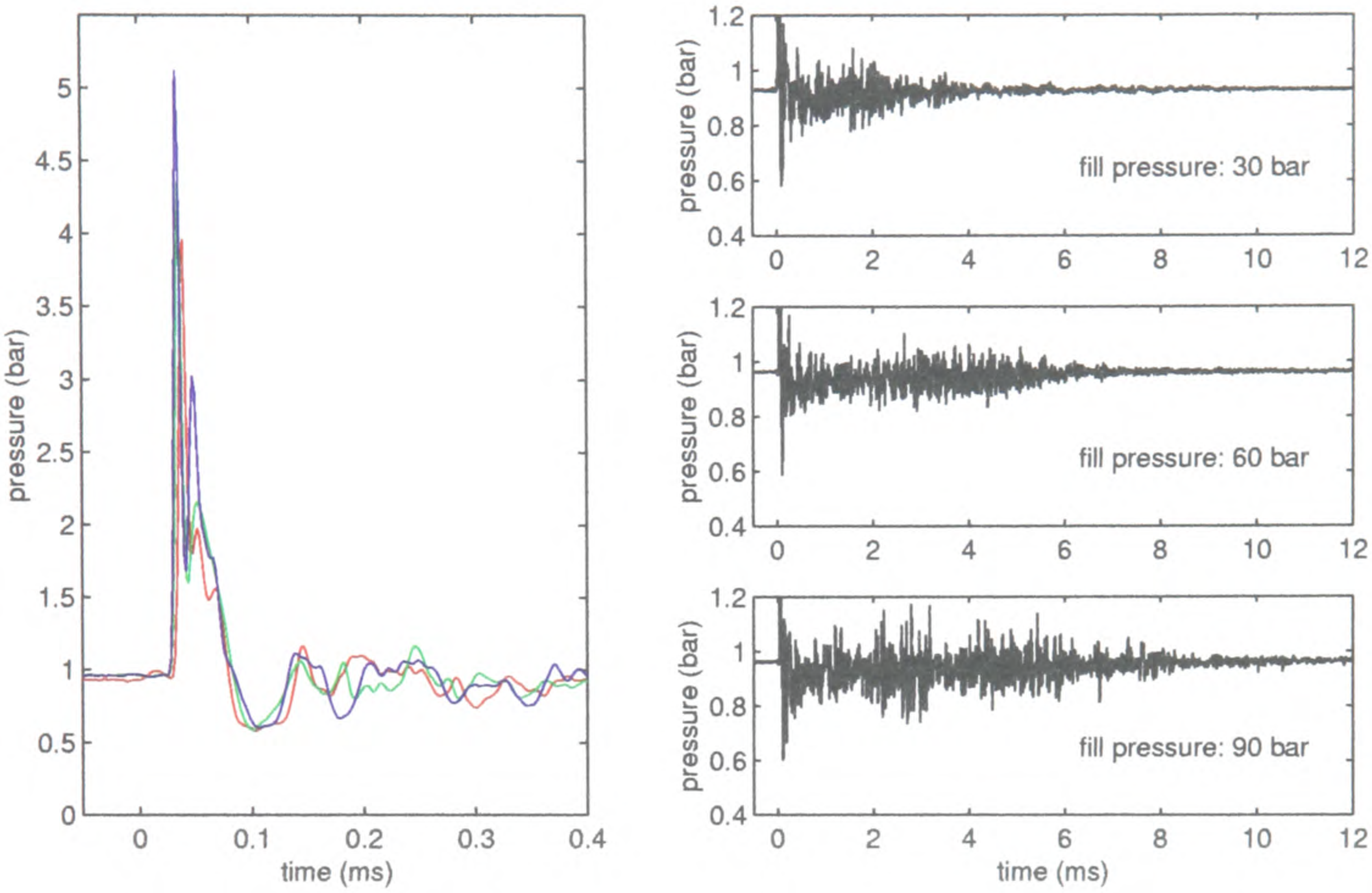


Figure 4.21 Pressure measured at the downstream nozzle transducer for fill pressure values of 30, 60 and 90 bar, with 20 μm polycarbonate membranes. Early segments of each of the three results are overlaid in the graph on the left; the graphs on the right show a portion of each run individually.

These results support the assessment (put forward in §4.2.1) that an extended shock system exists in the nozzle from the start of quasi-steady flow. The shock system drifts upstream as total pressure decays, as in classical quasi-one-dimensional nozzle flow. Across the shock system from upstream to downstream, pressure rises, while Mach number and total pressure fall. The upstream transducer is located within the region initially occupied by the shock system. As the system moves upstream in the course of a run, flow in the neighbourhood of the transducer moves further away from the supersonic state which exists somewhere upstream, and approaches the shocked subsonic exit condition. Local pressure rises as the downstream end of the shock system comes closer. The effect of an increase in fill pressure is to raise the starting total pressure and to shift the starting position of the shock system downstream. The transducer is then located closer to the upstream end of the shock system and the shock-free supersonic flow beyond it, in a region of lower pressure. This is why the pressure recorded at the upstream transducer drops away as fill pressure is increased.

These observations support the major finding of the programme of pressure measurements in the conical nozzle, which is that supersonic flow in the nozzle terminates upstream of the exit in an extended shock system. This is the case because pressure in the rupture chamber is not adequate to drive supersonic flow to the exit, even at fill pressures as high as 90 bar. It appears that a large fraction of the total pressure of the stored gas is lost in the flow from cylinder to rupture chamber.

4.2.5 Overview of Gas Flow in the Conical Nozzle

Flow in the conical nozzle is initiated by an unsteady shock which is attenuated significantly in its propagation through the nozzle. This is followed by a starting process which lasts for around 200 μs and introduces an oscillatory quasi-steady flow. The major qualitative difference between the flows in the conical nozzle and the Mach 3.5 contoured nozzle is that supersonic quasi-steady flow to the nozzle exit is never established in the conical nozzle at typical operating conditions. Instead, supersonic flow terminates in an extended shock system which begins upstream of the instrumentation used in this work — that is, at least 18 mm upstream of the nozzle exit. As total pressure decays with time, the shocks move upstream.

4.3 Optical Particle Detection

Optical particle detection (OPD) was used to record the timing of particle delivery from the conical nozzle. When used in conjunction with accompanying nozzle pressure measurements, OPD data describe the distribution of the particle payload within the gas flow. This is essential information, as it allows identification of the periods of gas flow which are important for particle acceleration. It also allows the influence of the particles on the gas flow to be assessed.

OPD experiments and nozzle pressure measurements were carried out simultaneously on a device with a conical nozzle, operating at a fill pressure of 60 bar, with 20 μm polycarbonate membranes. The test device was loaded with payloads of polystyrene microspheres of various particle diameter and total mass. The nozzle pressure measurements obtained in these experiments are reported in §4.4. This section is concerned primarily with the particle detection data. In §4.3.1, a representative OPD data set is described in detail to introduce the technique and to present and discuss the typical form of the particle delivery history. Results of the main programme of experiments, in which the make-up of the payload was varied, are summarised in §4.3.2.

4.3.1 A Representative Result

As explained in detail in §3.2, the OPD experimental technique functions by recording the time-varying intensity of light scattered by particles as they exit the device and pass through a probe volume which is centred 2 mm downstream of the nozzle exit. The output of an experiment consists of a time history called the *OPD signal*, which represents an uncalibrated measure of particle concentration.

The results of a typical optical particle detection experiment are presented in ***Figure 4.22***, which shows both the nozzle pressure measurements and the OPD signal over two time scales of interest. The particle detection signals have been corrected for background level, using the procedure explained in §3.2.2. In this particular experiment, a payload of 2.8 mg of 15.5 μm polymer spheres was delivered.

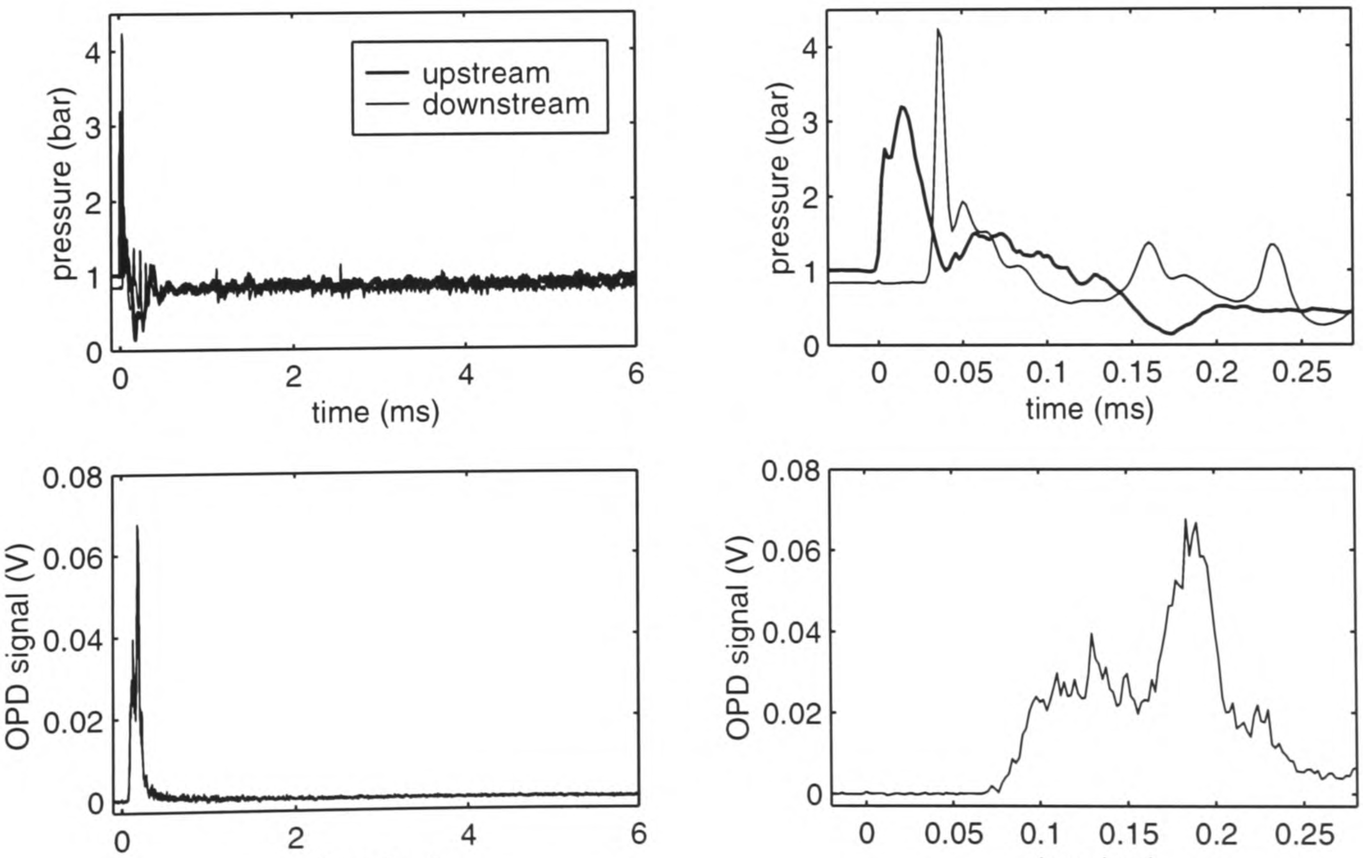


Figure 4.22 Results of a typical OPD experiment, with a 2.8 mg payload of 15.5 μm particles. The OPD signal (bottom) and accompanying pressure measurements (top) are shown on two time scales.

The first particles are detected approximately 70 μs after the starting shock passes over the upstream pressure transducer, and 40 μs after it passes the downstream transducer. The most intense concentration of particles occurs over the next 300 μs. The signal then drops sharply and decays gradually to zero over approximately 1 ms.

Further complete OPD results may be found in §4.4. Although timings vary, the pattern observed in delivery of this test payload is consistently repeated — the particle cloud concentration rises sharply, remains high (though variable) for some time, and decays slowly. The first particles to emerge from the nozzle do so near the end of the gas dynamic starting process, or as it gives way to quasi-steady nozzle flow, and particle delivery continues until some time after the establishment of quasi-steady nozzle flow at the exit. The significance of this finding is that gas in both the starting process and the quasi-steady nozzle flow are involved in the acceleration of drug particles.

4.3.2 Effects of Payload Mass and Particle Size

In this section, the timing of particle delivery is discussed on the basis of OPD signals recorded for various values of particle size and payload mass. OPD data are summarised here in the form of scalar parameters which are fully defined in §3.2.2. The parameters are the *time to first delivery*, the *20% rise time*, and the *90% delivery time.*

The passage of the starting shock over the downstream transducer is convenient as a precisely indentifiable reference time in the flow-field near the OPD probe location. In ***Figure 4.23***, the delay from detection of the shock at the downstream transducer to delivery of the first particles is shown as a function of particle size and payload mass. The mean value of this delay is 47 μs, and its standard deviation is 8 μs. The delay is not correlated with either payload mass or particle size.

Figure 4.24 is a plot of the interval from the detection of the starting shock at the downstream transducer to the time at which the OPD signal first exceeds 20% of its maximum value, as measured for all test payloads. This parameter indicates the time elapsed from the beginning of flow until the establishment of particle delivery at a substantial level. Its value ranges from 40 to 120 μs. This is less than, or comparable to, the typical duration of the flow starting process, noted in §4.2 as approximately 200 μs. The leading particles must therefore be entrained within the starting process.

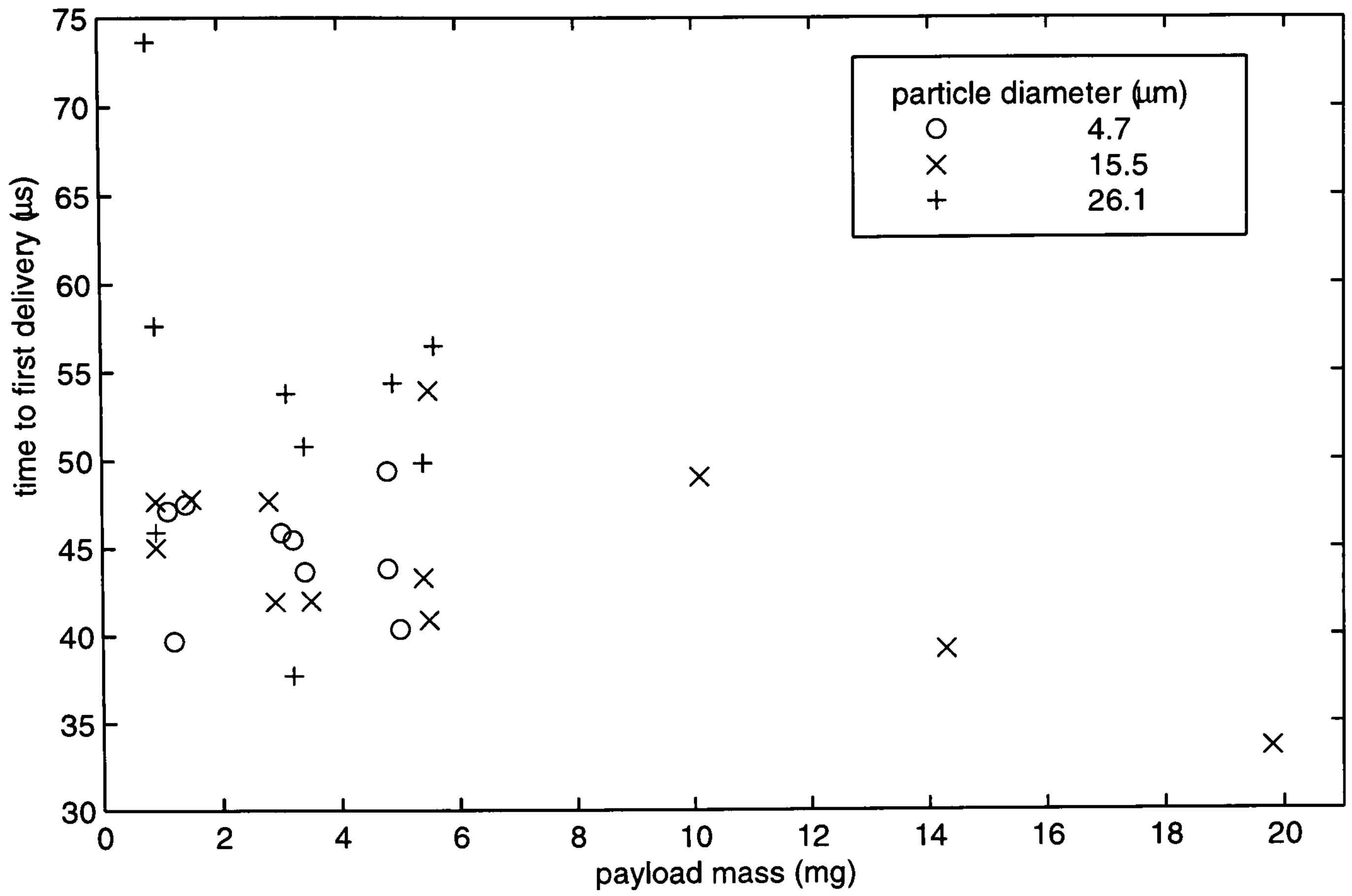


Figure 4.23 Time elapsed from detection of the starting shock at the downstream transducer to detection of the first delivered particles at the OPD location (2 mm downstream of the nozzle exit), plotted as a function of payload mass for three particle sizes.

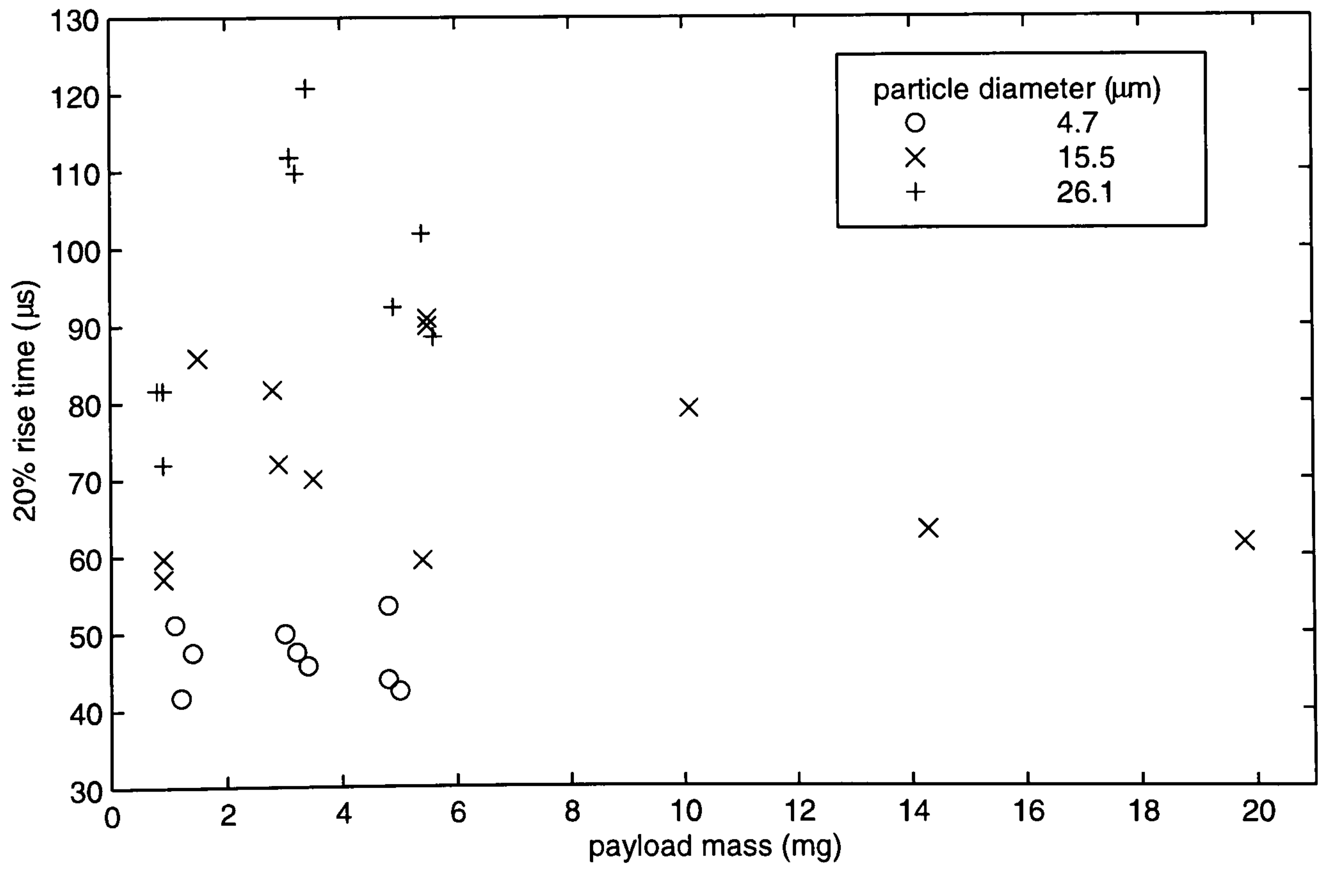


Figure 4.24 Delay from detection of the starting shock at the downstream transducer until the OPD signal reaches 20% of its maximum value, shown as a function of payload mass, for three polystyrene microsphere sizes.

The 20% rise time increases with particle size, from 47 ± 4 µs (mean ± standard deviation) for 4.7 µm particles, to 73 ± 13 µs for 15.5 µm particles, and 96 ± 16 µs for 26.1 µm particles. This result is a consequence of the fact that in a given flow-field, the smallest particles will accelerate most rapidly. However, the variation in the timing of first particle delivery (shown above in ***Figure 4.23***) did not display a similar trend. A possible explanation for this discrepancy is that the payloads of nominally large particles contained small quantities of particles of comparable size to the 4.7 μm spheres. This might come about through contamination by residue in the drug delivery device, or might be a characteristic of the size distribution of the particles as manufactured. The 20% rise time is independent of payload mass for the two smallest particle grades, though in the case of the 26.1 μm particles, it is significantly longer for the 3 mg payloads than for the 1 mg and 5 mg payloads. The key observation put forward in this analysis of the OPD data, however, is that particle delivery is always underway before quasi-steady flow is established.

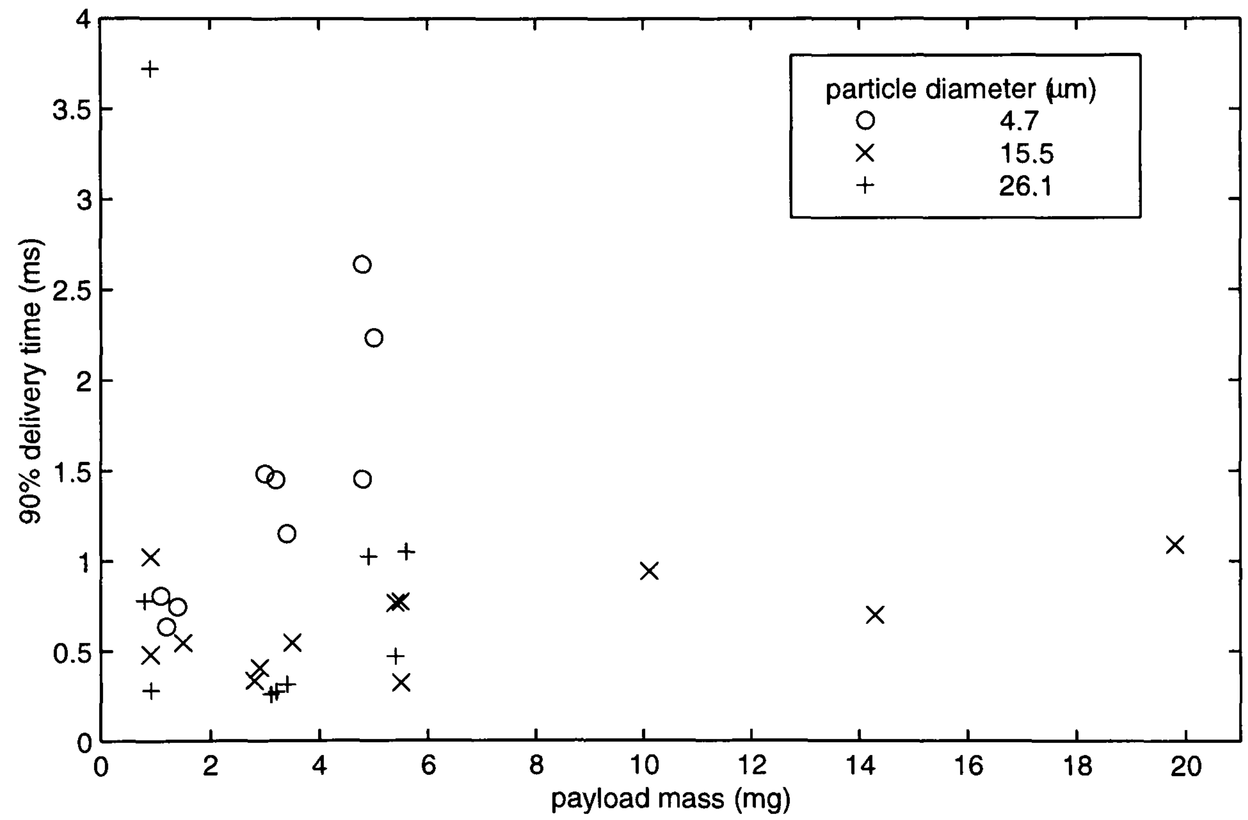


Figure 4.25 Time elapsed from the beginning of powder delivery until the integrated OPD signal accumulates to 90% of its final value, plotted as a function of payload mass, for three polymer microsphere sizes.

The *90% delivery time*, defined as the time elapsed from the beginning of delivery until accumulation of 90% of the total integrated OPD signal, is plotted in ***Figure 4.25***. This is a measure of the time taken for complete delivery of the payload. There is one apparently anomalous data point, representing a 90% delivery time of 3.8 ms for a 1 mg payload of 26.1 μm polystyrene spheres. This may be due to retention of drug particles in a sheltered joint in the structure of the drug cassette. The outlying datum is ignored for the purposes of this discussion.

The structure displayed in these results is not simple. Firstly, the delivery time is significantly longer for 4.7 μm particles than for either of the larger grades — the mean and standard deviation for each particle type are 1.40 ± 0.68 ms, 0.66 ± 0.26 ms and 0.56 ± 0.34 ms for 4.7 μm, 15.5 μm and 26.1 μm microspheres, respectively. Secondly, delivery time increases with payload mass for 4.7 μm particles, but is not significantly affected by changes in mass for the larger particles. In all cases, however, particle flow persists well beyond the 200 μs lifetime of the starting process and into the quasi-steady stage of gas flow. Taken in conjunction with the observed values of 20% rise time discussed above, these data show that both the starting process and the quasi-steady gas flow are involved in particle delivery for all payload configurations tested.

4.4 Pressure Measurements in the Presence of Payload

This section focuses on the nozzle pressure measurements which were made in conjunction with the optical particle detection experiments described above. The OPD experiments present an opportunity to examine the changes (if any) in the unsteady nozzle pressure brought about by the presence of payload.

A selection of pressure measurements are shown with their accompanying OPD signals as ***Figures 4.27*** to ***4.30***. Each figure includes pressure measurements from both transducers, along with OPD signals, for one or two values of payload mass. For comparison, a pressure measurement taken in a run with no payload is shown first in ***Figure 4.26***, with the same format and scale as the other figures.

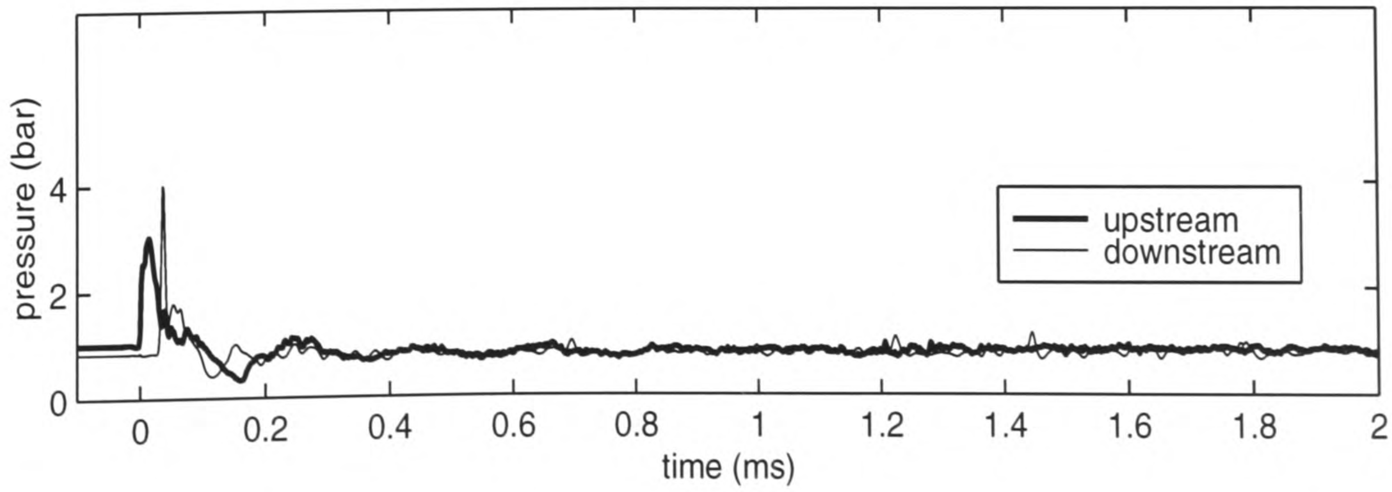


Figure 4.26 Pressure measurements taken in the conical nozzle, operating at 60 bar fill pressure, with 20 μm polycarbonate membranes and no payload.

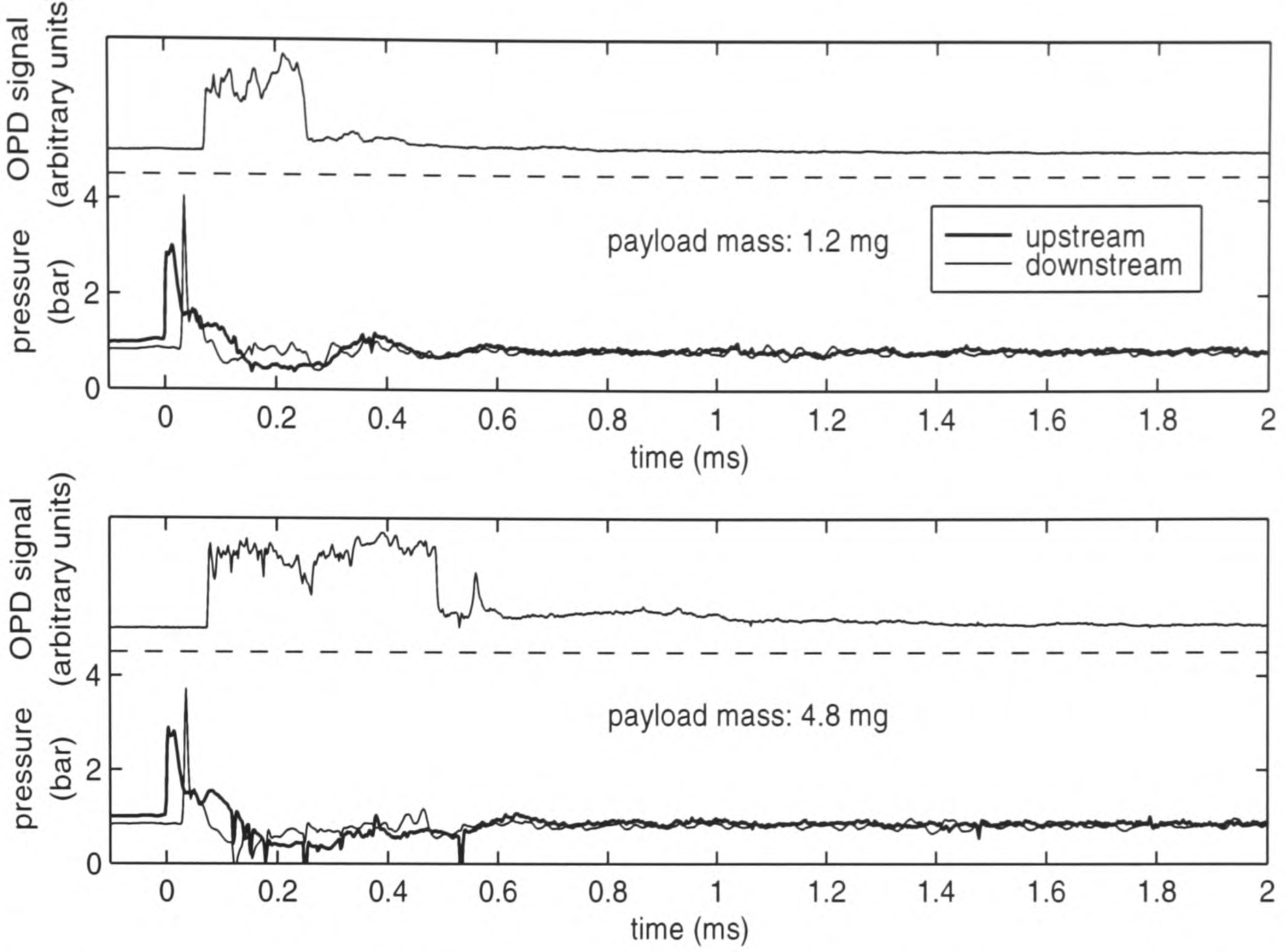


Figure 4.27 OPD signals and pressure measurements for 4.7 µm microsphere payloads of mass 1.2 mg (*top*) and 4.8 mg (*bottom*).

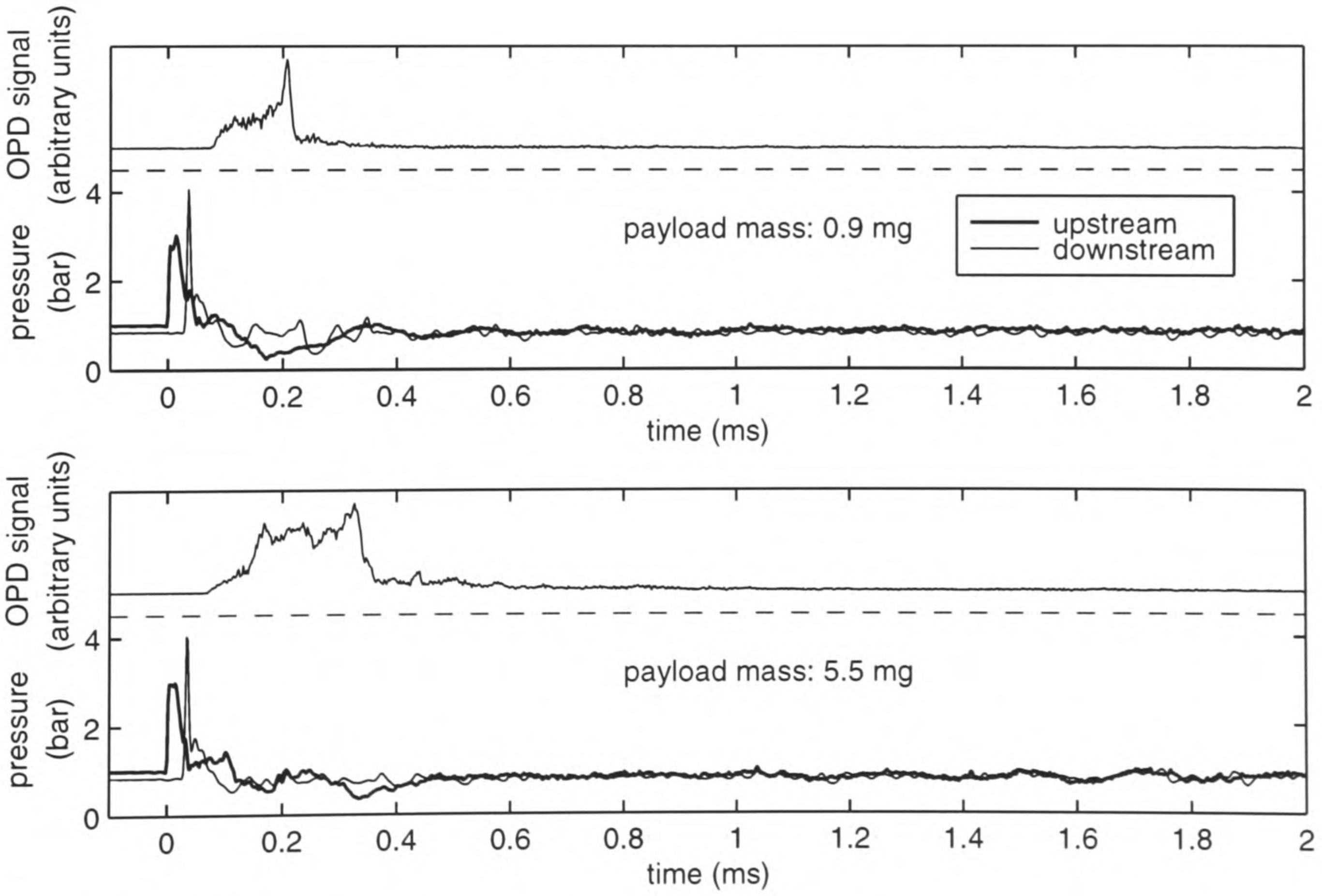


Figure 4.28 OPD signals and pressure measurements for 15.5 µm microsphere payloads of mass 0.9 mg (*top*) and 5.5 mg (*bottom*).

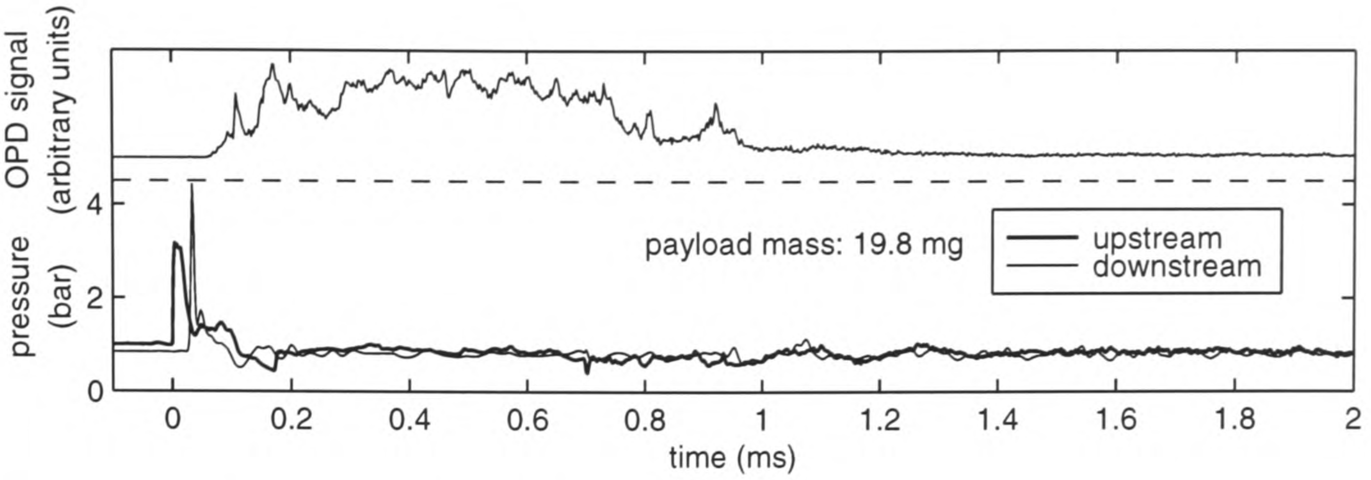


Figure 4.29 OPD signal and pressure measurements taken during delivery of 19.8 mg of 15.5 μm particles.

It is clear that the payload exerts a considerable influence on pressure in the nozzle. In runs without payload, pressure at the upstream transducer drops to a minimum approximately 180 μs after passage of the starting shock, rises again, and undergoes decaying oscillations. The addition of a payload modifies the pressure profile around this minimum. The period of lowest pressure is prolonged to a duration of the same order as the lifetime of high OPD signal. At the downstream transducer, the pressure minimum is often replaced by a pair of minima, separated in time by approximately 200 μs. This phenomenon is most distinct for payloads around 1 mg in mass.

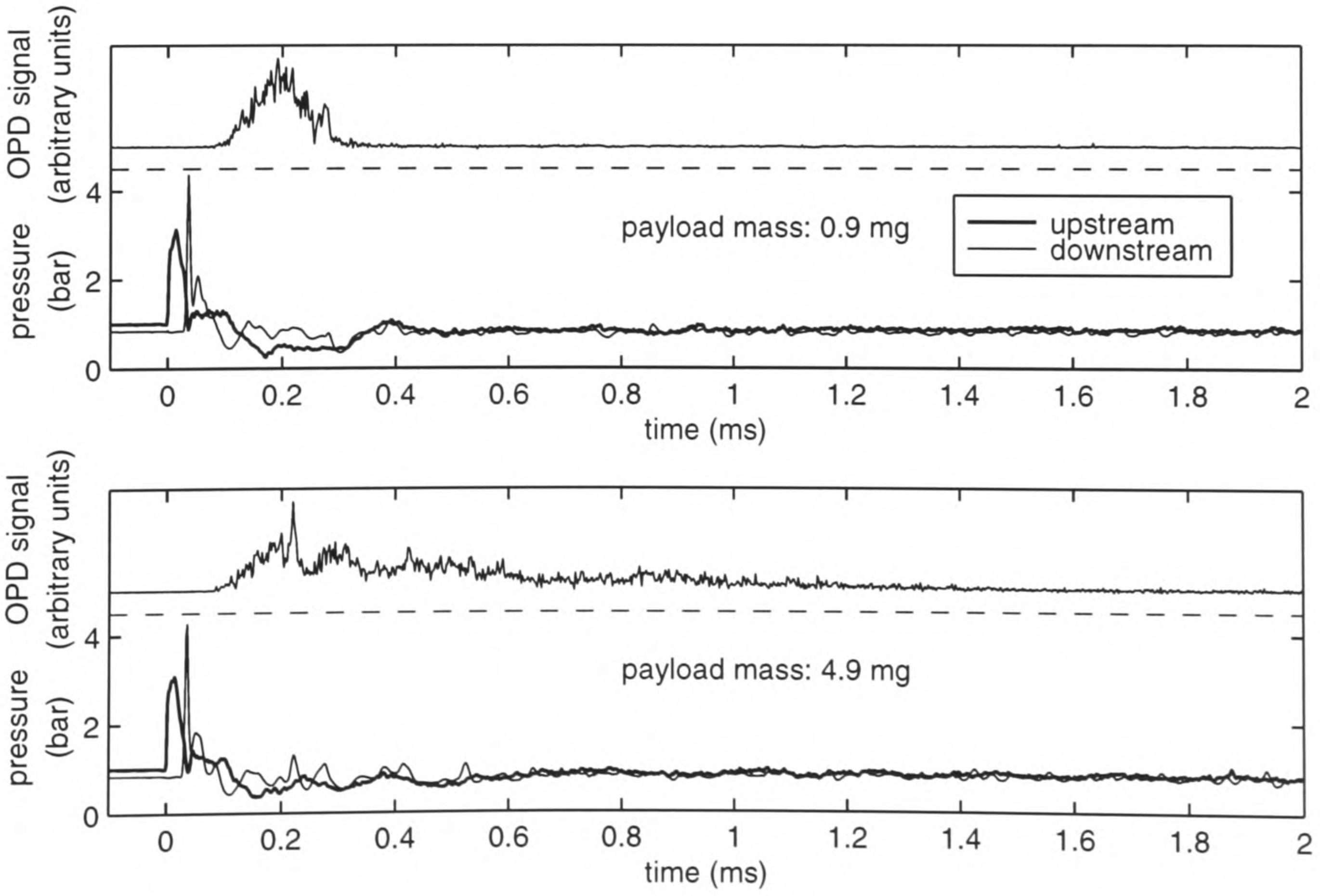


Figure 4.30 OPD signals and pressure measurements for 26.1 μm microsphere payloads of mass 0.9 and 4.9 mg.

The effect of payload on the nozzle flow, as observed at the upstream pressure transducer, is exemplified by the data shown in ***Figure 4.31***. A pressure measurement without payload is shown along with measurements in runs with 0.9 mg, 5.5 mg and 19.8 mg payloads of 15.5 μm particles. The scales in these graphs have been chosen to highlight the quasi-steady flow. When particles are added to the flow, there is at first a reduction in the rate of pressure increase after the minimum. At higher payloads, the value of the minimum is raised slightly, and the subsequent rise becomes as steep as it is in particle-free flow. This rise is interrupted by a second strong dip in pressure. For the largest payload, the initial dip is cut off by a sharp rise to an approximately steady level, which leads into a late and gradual pressure drop around 1 ms after the start of flow.

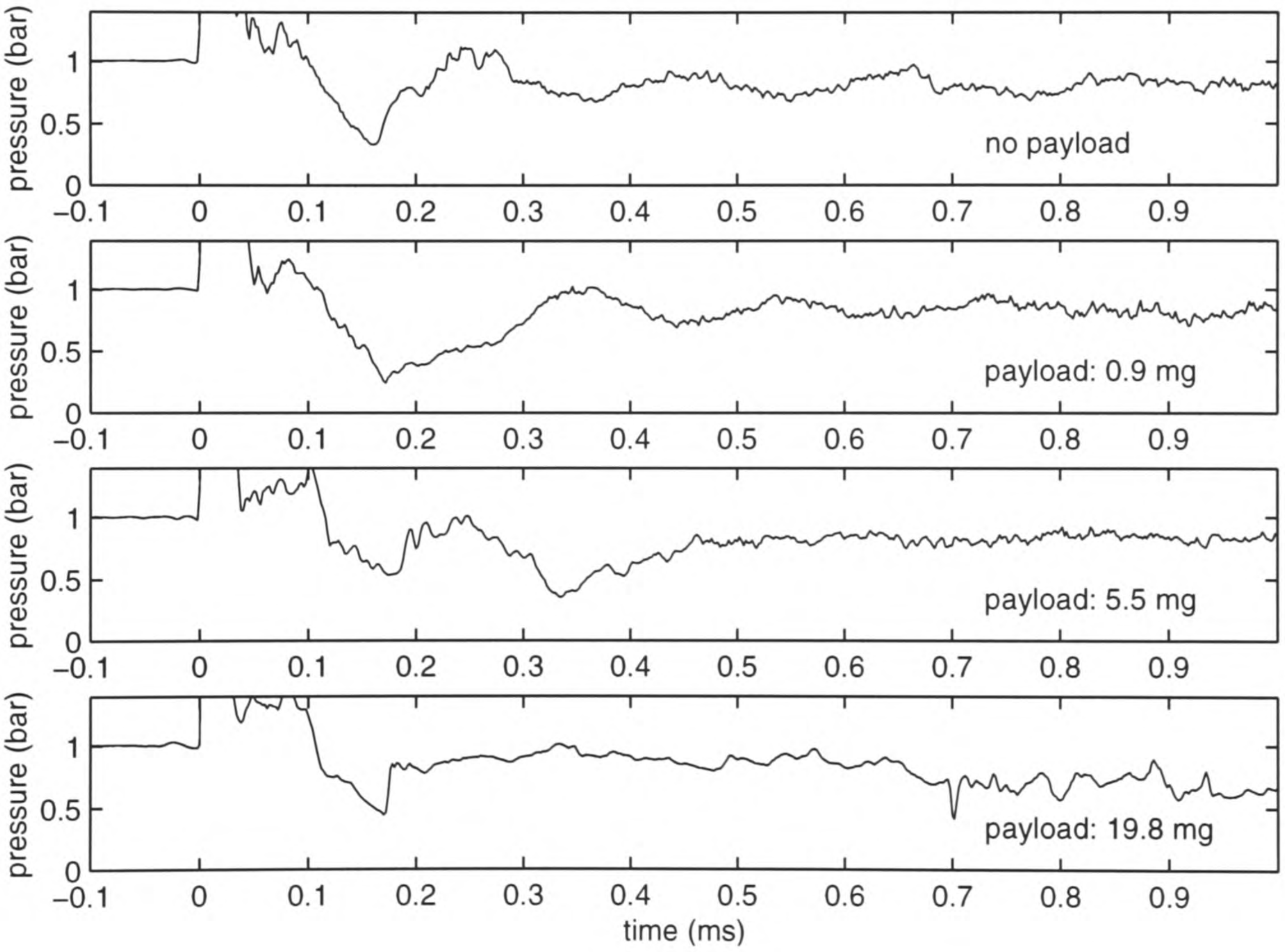


Figure 4.31 Pressure measurements at the upstream transducer, taken with no payload *(top)* and various payloads of 15.5 μm microspheres.

A possible interpretation of these data is that a pocket of low pressure gas is convected down the nozzle near the tail of the particle cloud. For small payloads, it appears in combination with the usual pressure dip, associated with the transition from shock tube to quasi-steady flow. With larger payloads, it appears as a distinct feature. By comparing the upstream and downstream pressure records for 1 mg payloads in ***Figures 4.27***, ***4.28*** and ***4.29***, it can be seen that a distinct second dip in pressure also occurs at the downstream transducer, slightly later than the prolonged interval of low pressure at the upstream transducer. This behaviour is discernible in some, but not all, of the downstream pressure measurements taken with larger payloads. In the extreme case of the 19.8 mg payload, a continuous pressure drop is apparent during particle delivery.

The simplest explanation for this hypothesised low pressure zone is that a pressure drop is set up across the particle cloud, from high pressure downstream to low pressure upstream. The drop is caused by flow in an upstream direction relative to the particles — in other words, particles are moving downstream faster than the gas in this part of the nozzle, and the pressure gradient balances particle drag forces. Alternatively, the particle cloud can be thought of as a porous plug through which gas flows with pressure losses. The theoretical possibility that particles can travel faster than the gas has been demonstrated by numerical predictions of gas and particle velocity, presented in Chapter 7. The changes in nozzle pressure due to the particles represent a large fraction of the absolute pressure, and are therefore likely to be associated with significant changes in gas velocity and density. Ultimately, this will impact on the performance of the drug delivery system.

The particle payload produces a second class of effects associated with the starting shock wave. It was pointed out in §4.2.3 that the initial pressure rise detected at the upstream transducer, in the absence of payload, occurs in two stages. In the pressure measurement shown in ***Figure 4.26***, for example, pressure immediately behind the starting shock wave is approximately 2.5 bar at the upstream transducer. After a period of microseconds during which pressure is constant, it jumps to approximately 3 bar. Comparison of the results for 1 mg payloads of various particle sizes in each of ***Figures 4.27***, ***4.28*** and ***4.30*** with the companion results for 5 mg, and with the results for no payload in ***Figure 4.26***, indicates that as payload mass is increased from zero, the strength of the starting shock increases while the strength of the second compression wave decreases. The overall pressure jump across the two-stage compression remains approximately constant.

This observation is confirmed by an analysis of the shock propagation speed, based on time of transit between the transducers. Shock speed is plotted in ***Figure 4.32*** as a function of payload mass. It increases by some 20 m/s, or 4%, as payload is increased from 0 to 5 mg, regardless of particle size. There is a relatively slight further increase as the mass of the 15.5 μm payloads is increased to 20 mg, but the scarcity of data points in this region makes conclusions questionable. ***Figure 4.32*** represents strong evidence that the particle cloud has a substantial effect on the dynamics of the gas. In fact, it shows that the influence of the cloud extends downstream to the theoretical limit, the starting shock. A mechanism for the increase in shock speed with payload mass is proposed in §6.4.2, on the basis of insights provided by time-resolved DGV.

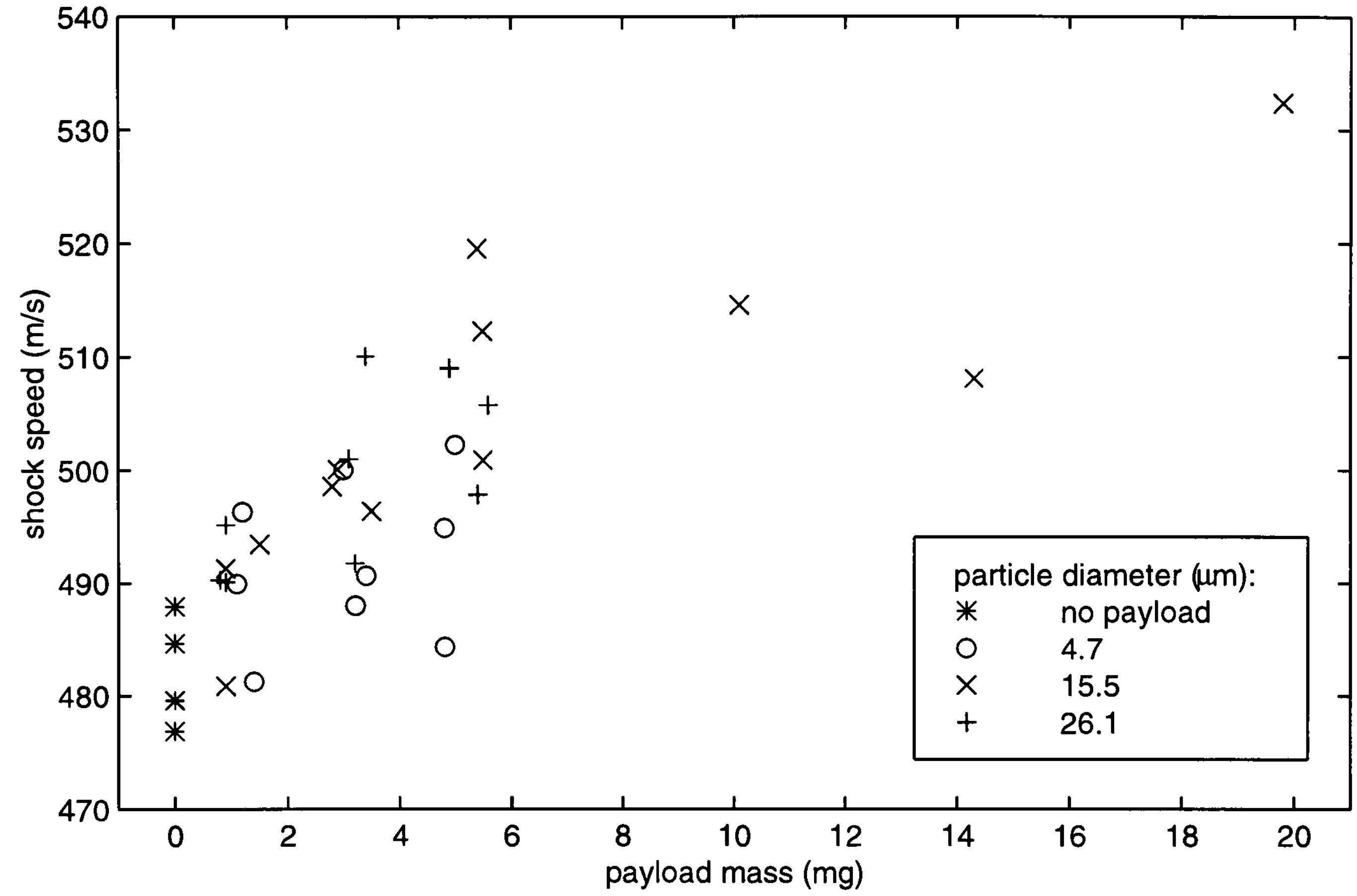


Figure 4.32 Variation of the speed of the starting shock in the conical nozzle for a range of particle payloads. The calculation of speed is based on the transit time measured between pressure transducers.

average payload (mg)	average shock speed (m/s)	average gas speed behind shock (m/s)
0	482	195
1.1	490	205
3.2	497	214
5.2	502	220

Table 4.1 Propagation speed of the starting shock, and gas speed immediately behind the shock, varying as a function of payload mass.

The velocity of gas immediately behind the shock in the instrumented region of the nozzle has been estimated from the measured shock speed between transducers, using the standard shock tube relations given by Anderson [10] and others. The results are listed in ***Table 4.1***, and will be referred to in Chapter 6 in the context of particle velocity measurements. The gas speed behind the shock increases from 195 to 220 m/s as payload mass is increased from 0 to 5 mg.

4.5 Conclusions

Measurements of unsteady pressure in the drug delivery devices under study, along with observations of particle delivery timing, have confirmed and refined the theoretical description of the drug delivery flow put forward in preceding chapters. Qualitatively similar gas flow histories have been observed in the Mach 3.5 contoured nozzle and the conical nozzle. Flow is initiated by a shock wave, which is followed by a starting process with a duration of the order of 200 μs. With 20 μm polycarbonate membranes, the gas velocity immediately behind the starting shock has been estimated as 586 m/s for the Mach 3.5 contoured nozzle without payload, and 195–220 m/s (increasing with payload mass) for the conical nozzle. When the starting process is completed and the associated unsteady waves are swept out of the nozzle, a quasi-steady nozzle flow is established. In the Mach 3.5 contoured nozzle, this stage is characterised by supersonic flow throughout the divergent part of the nozzle, sustained for (typically) 7 ms. In the conical nozzle, on the other hand, total pressure is never sufficient to drive quasi-steady flow to the nozzle exit, and an extended system of shocks is stationed in the downstream region of the nozzle from the beginning of flow. This inadequate total pressure is due to a large pressure drop in the flow from the cylinder to the rupture chamber, and a reduction in particle delivery velocity in the quasi-steady flow can be expected as a result.

For the conical nozzle, optical particle detection experiments have linked these observations of the gas dynamics to the central issue of particle acceleration. The flow of particles from the conical nozzle begins before the end of the starting process, and continues for a period of 0.3 ms to 2.5 ms (depending on particle size and payload mass), well into quasi-steady flow. Both the strongly unsteady starting flow and the quasi-steady nozzle flow, therefore, are involved in the acceleration of the drug particles.

Pressure measurements in the conical nozzle in the presence of payload, carried out in conjunction with particle detection, indicate a zone of low pressure near the tail of the particle cloud. The data suggest that the magnitude of the particles' influence on the gas flow may be strong enough to affect the flow velocity. However, the gas-particle interaction is a complex topic which requires further research if the nature and magnitude of the payload's effects on the gas flow are to be understood.

These findings constitute a description of the gas flow in the drug delivery devices under investigation, including basic information on the interaction between the gas and the drug particles. They will be combined with the DGV measurements of particle velocity, presented in Chapter 6, to move towards an integrated understanding of the drug delivery system's operation.

Chapter 5

Doppler Global Velocimetry: Methods

Doppler Global Velocimetry (DGV) is a new technique for the measurement of velocity fields in fluid flows. In a single non-intrusive experiment, it provides velocity data at high spatial resolution across an arbitrary plane slice through the flow-field of interest. The DGV systems under development in Oxford were utilised as the major experimental tool in this research into devices for transdermal delivery of drugs in dry powder form.

This chapter is an account of how DGV was applied to produce the velocity field measurements presented in Chapter 6. It begins with a brief general discussion of the nature of DGV and its place among velocimetry methods (§5.1), and proceeds with a technical description of the technique, and the Oxford implementation in particular (§5.2). A method for uncertainty analysis is presented in §5.3. The specific application of the technique to flow-fields of interest in this thesis is described in §5.4. Finally, §5.5 consists of analysis and discussions of the integrity and interpretation of DGV results obtained in this work.

Throughout this chapter, images of flow fields and measured velocity maps will be presented to illustrate various points about the measurement technique. For the purposes of this chapter, they are not treated as measurements or representations of physical flow fields, but as abstract examples of data sets captured by the measurement. All interpretation of DGV data relating to velocity measurements of drug delivery flow fields, in physical terms, is reserved for Chapter 6.

5.1 Background and Capabilities

Doppler Global Velocimetry was first documented in the literature by Komine [34] in 1991, and a number of research groups have reported developments and applications of the technique since then. Comprehensive reviews have been published by Meyers [8] in 1995, and, more recently, by Ainsworth *et al.* [9]. The DGV research programme in Oxford University was initiated, and continues today, with the goal of application in a short-duration gas turbine test facility, which is an arena of complex three-dimensional unsteady transonic flow. Details of DGV work in Oxford were first published by Ainsworth and Thorpe [35] in 1994, who demonstrated the technique with time-averaged measurements of the velocity field of a spinning disc, and discussed the feasibility of measurement in turbine applications. Subsequent developments of the technique, and application in the time-averaged measurement of the velocity field of a subsonic turbulent jet, were reported by Thorpe *et al.* [36,37].

The development of DGV has been motivated by the desire for non-intrusive whole-field measurements of complex flows. In an experiment consisting of the simultaneous acquisition of two images of a laser-illuminated flow-field, it delivers a measurement of the distribution of one, two or three velocity components across a plane slice through the field. Though at a relatively early stage of development, it offers potential advantages over comparable established technologies for fluid dynamics measurements. Laser Doppler Velocimetry (LDV) [38], first reported in 1964, is a powerful means of non-intrusively measuring up to three components of velocity at a single point. Particle Image Velocimetry (PIV) provides two-component measurement across a whole plane; Adrian [39] argues that it has been practised, in various forms, since 1932. However, it relies on the ability to resolve individual particles in the flow field by optical imaging, and although measurement of a third (out of plane) velocity component is possible, extensions to the basic technique are required. Both PIV and LDV are now widely accessible through commercially available systems.

The capability for whole-field measurement can lead to great reductions in testing time when compared with pointwise methods, and is therefore a highly desirable attribute in an experimental technique, particularly for the investigation of complex flows. Both DGV and PIV are whole-field techniques. In DGV, however, no velocity components have any special significance from a measurement point of view. DGV is intrinsically a single-component measurement method, but the measurement of multiple components is a simple matter of carrying out multiple measurements simultaneously (or sequentially, if the flow is steady or repeatable). It is not necessary to resolve individual particles in flow-field images for DGV, and so particle size and concentration, and optical resolution, are less critical for the success of a measurement than in PIV. While PIV requires at least two laser pulses (or, in some versions, two camera exposures) in rapid succession, DGV functions with continuous-wave illumination, or a single laser pulse. Finally, DGV can report a velocity value for every resolved point in a flow-field image, whereas in PIV, each velocity datum is derived from an analysis of a finite interrogation region in the raw image. DGV therefore has the potential to make more efficient use of available imaging resolution.

5.2 The Oxford Implementation of DGV

This section describes the DGV technology which has been developed in the Department of Engineering Science in Oxford University and applied in the research described in this thesis. The account given here is largely a summary of the work of others. Most of the DGV hardware was developed by Dr. Steven Thorpe. Some later elements, and the image processing software used in the current work, were added by Dr. Robert Manners in the Department of Engineering Science.

A comprehensive account of the development of DGV in Oxford is given by Thorpe [41], and is beyond the scope of this thesis. The description given here is intended to provide an understanding of the capabilities and limitations of the technique, and to provide the necessary background for later descriptions of the tasks involved in applying DGV to the drug delivery flow-fields which are the main topic of this thesis. Although the information given here refers specifically to the Oxford DGV systems, much of the description which follows is applicable to the DGV method in general. Major differences between the Oxford method and other versions are mentioned as they arise. This section begins with a general explanation of the principles of DGV and an outline of DGV apparatus in §5.2.1. The remaining subsections, §5.2.2 to §5.2.7, are concerned with technical descriptions of specific aspects of the system.

5.2.1 Overview of the DGV Method

DGV exploits the well-known Doppler effect — the shift in the frequency of wave radiation which is observed when the source and observer of a wave are in relative motion. In particular, when a moving particle scatters wave radiation, the particle's velocity is related to the frequency shift which is observed in the scattered wave:

$$\Delta\nu = \nu_0 \frac{\hat{o} - \hat{\imath}}{c} \cdot \boldsymbol{V} , \qquad (5.1)$$

where $\Delta\nu$ is the shift in frequency, ν_0 is the frequency of the wave incident on the scattering particle, c is the wave propagation speed, and $\boldsymbol{V}$ is the velocity of the particle. $\hat{o}$ and $\hat{\imath}$ are unit vectors in the direction of the observed and incident waves, respectively. The dot product in the equation implies that the frequency shift is sensitive only to the component of velocity in the direction of $\hat{o} - \hat{\imath}$.

The vectors in ***Eq. 5.1*** are shown in ***Figure 5.1***, a diagram which illustrates how the Doppler effect can form the basis of a generalised velocimetry technique. Wave radiation (in practice, laser light) of a known frequency is directed onto a particle whose velocity is to be measured. A frequency-sensitive detector gathers and analyses some of the radiation scattered by the particle. If the frequency and direction of both the incident and scattered radiation are known, one of the particle's velocity components can be inferred via ***Eq. 5.1***. This principle underlies both DGV and LDV.

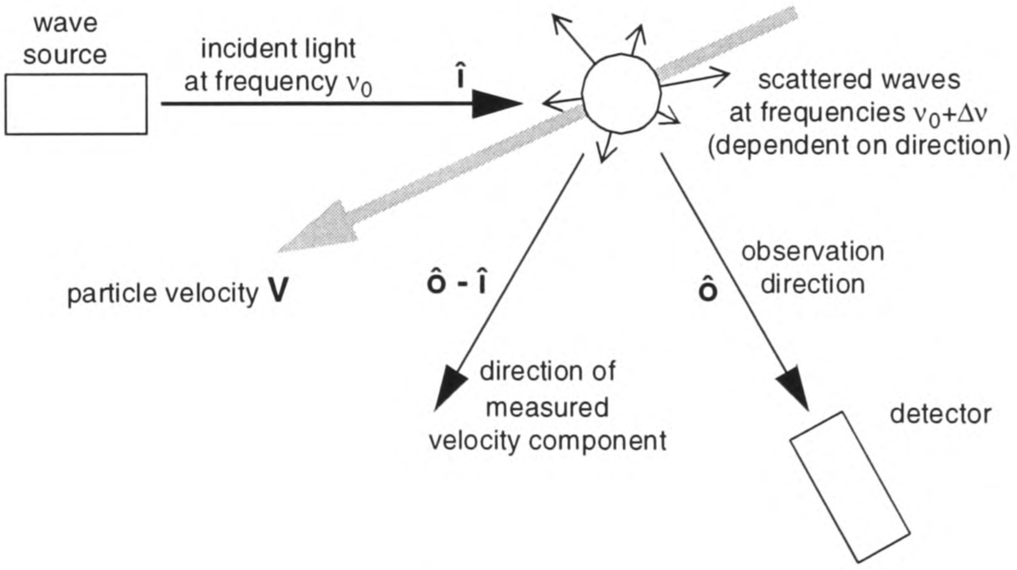


Figure 5.1 The principle of a Doppler velocimetry technique.

The essential innovation of DGV is in the use of iodine vapour, contained in a glass cell, as a frequency sensor. The iodine acts as an absorption filter, attenuating the scattered light by a fraction which depends on the light's frequency. This concept exploits strong absorption features in iodine, occurring in the visible spectrum at frequencies which coincide with emission frequencies available from several types of laser. The concept of the *iodine cell* and its practical implementation are described in more detail in §5.2.2. Given a calibration of an iodine cell's absorption characteristic as a function of frequency, it is possible to infer the frequency of light passed through the cell by measuring the fraction of the light which is absorbed (or transmitted) by the cell. The global (or whole-field) character of DGV stems from the fact that many light rays, of varying intensity and frequency, can be filtered simultaneously and independently using a single iodine cell. Consequently, the cell can be used to measure the frequency of light scattered from many locations in a flow-field at the same time. This is the basic concept of DGV.

An outline of the DGV system built in Oxford to realise this principle is shown in ***Figure 5.2***. The flow-field of interest is illuminated by a sheet of laser light which defines a measurement plane. Some of the light scattered (with Doppler shift) by particles in the illuminated plane is then gathered in an optical system, filtered by the iodine cell, and focused in a camera. Each location in the resulting image recorded by the camera corresponds directly to a location in the flow-field. An image of the flow-field formed in this way is called a *discriminated* image, because the frequency at which light is scattered from each point in the flow-field is encoded (by virtue of the iodine cell's frequency-sensitive absorption action) in the intensity or brightness of the corresponding point in the image.

However, the iodine cell is not the only influence on the intensity distribution in the discriminated image. Non-uniformities in illumination, in the concentration and size of light-scattering particles in the flow-field, and in optical equipment, all contribute to the image intensity distribution. Therefore, a second image of the same flow-field scene, called the *reference* image, must be formed with unfiltered light, so that the frequency-dependent effect of the iodine cell on the image can be isolated. The reference and discriminated images are formed in a DGV imaging system described in §5.2.3. The image processing methods used to extract frequency shift data (and ultimately, velocity data) from these images have been described in detail by Manners *et al.* [42], and are summarised in §5.2.6.

DGV has been implemented in both time-integrated and time-resolved formats. In time-integrated DGV, the measurement region is illuminated continuously, and images are formed over an arbitrary period of exposure of the recording camera or cameras. In time-resolved DGV, a pulsed laser is used to illuminate the flow-field for a very short time (typically of the order of 10 ns). This technique can be used to yield measurements of the essentially instantaneous velocity field in an unsteady flow. The relatively recent implementation of time-resolved DGV in Oxford is described in a forthcoming paper [43].

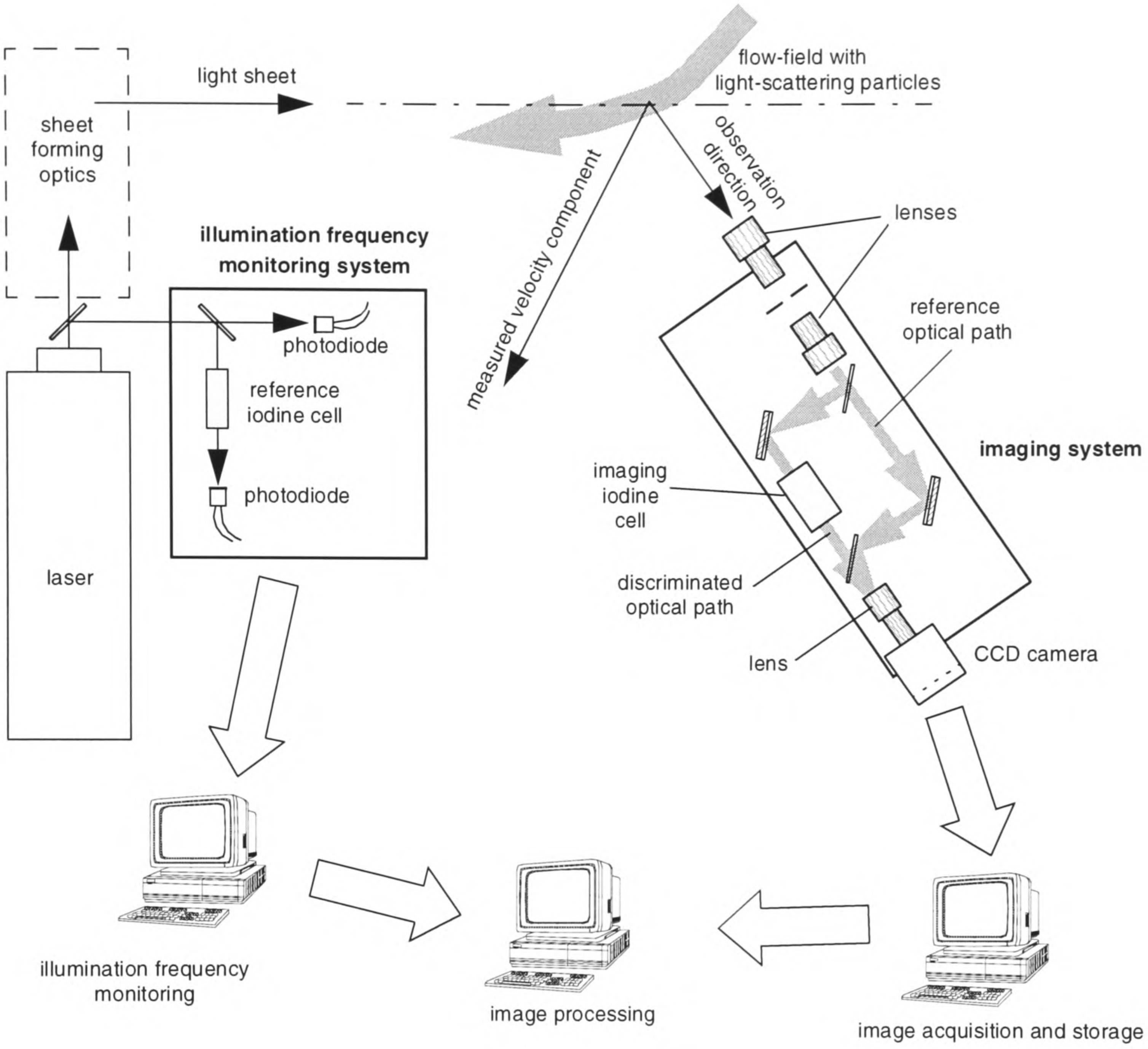


Figure 5.2 An outline of the Oxford DGV system hardware.

In the remainder of §5.2, the various components of the complete DGV system are described in detail. The iodine cell, the key frequency-sensing element of the system, is described in §5.2.2. The cell is built into an imaging system, described in §5.2.3, which arranges for the formation and recording of the discriminated image (which implicitly contains frequency shift information) and the reference image (which allows the influence of frequency shift to be decoupled from other influences on the image). The laser used to illuminate the flow-field is an important and specialised element of the system. In §5.2.4, the properties required of lasers for DGV are explained, and the specific lasers used in the Oxford systems are described. The apparatus used to measure the frequency of the laser light at source is described in §5.2.5. The image processing techniques by which a velocity map is derived from the flow-field images are outlined in §5.2.6. Finally, procedures for iodine cell calibration are explained in §5.2.7.

5.2.2 The Iodine Cell

As a consequence of the physics of electronic motion and electronic energy levels, radiation in certain relatively narrow regions of the electromagnetic spectrum is strongly absorbed by the iodine molecule. These neighbourhoods are the *absorption features* of iodine, and it is in these regions that iodine is useful as an absorption filter material. The iodine cell, which consists of a quantity of iodine vapour contained in a glass vessel, is the practical implementation of the frequency-sensitive light absorption filter which underpins the concept of DGV.

The performance of an iodine cell can be characterised in terms of the *transmission ratio*, which is defined as the fraction of incident light energy transmitted by the cell. The variation of transmission ratio with light frequency over a typical useful absorption feature, for one of the cells used in Oxford, is shown in ***Figure 5.3***. This absorption feature, occurring at a wavelength of approximately 514.5 nm (corresponding to green visible light) extends over a bandwidth of the order of 2 GHz (1.6×10^{-3} nm, in terms of wavelength), within which the normalised transmission ratio varies from unity at the edges of the feature to a minimum at the characteristic centre frequency of the feature. When the transmission profile (as shown in ***Figure 5.3***) of an iodine cell is known, the cell can be used as an instrument for the measurement of light frequency (within the range of the absorption feature, and relative to some reference point within the feature) through measurement of the transmission ratio.

In DGV work, the illuminating laser is tuned to emit at a frequency within the range of the absorption feature, such as the point marked "incident light" in ***Figure 5.3***. This light is scattered and Doppler-shifted by the particles whose velocity it is required to measure. The Doppler-shifted scattered light is directed through the cell, and filtered at a different transmission ratio, also marked on the graph. The magnitude of the Doppler shift is determined by measuring both of these transmission ratios (in practice, using different cells) and referring to the calibrated transmission profile of the iodine cell. Measurement of the transmission ratio is a matter of measuring light intensities, and the methods by which this is accomplished are described in the sections which follow.

As a mechanical structure, the iodine cell must contain the iodine vapour at low pressure (typically less than 1 torr) in such a way that light can pass through the cell without obstruction or aberration, and that the absorption properties of the cell, as a whole, are constant. Cells are constructed as a glass cylinder (of typical diameter 50 mm and length 60 mm) with optical grade flat glass plates as end walls, and are installed with the axis of the cylinder oriented approximately parallel to the direction of light propagation through the cell. They are built by placing iodine crystals in the cylinder, fusing the end walls to the tubular body, and then evacuating the cell through an orifice which is subsequently sealed.

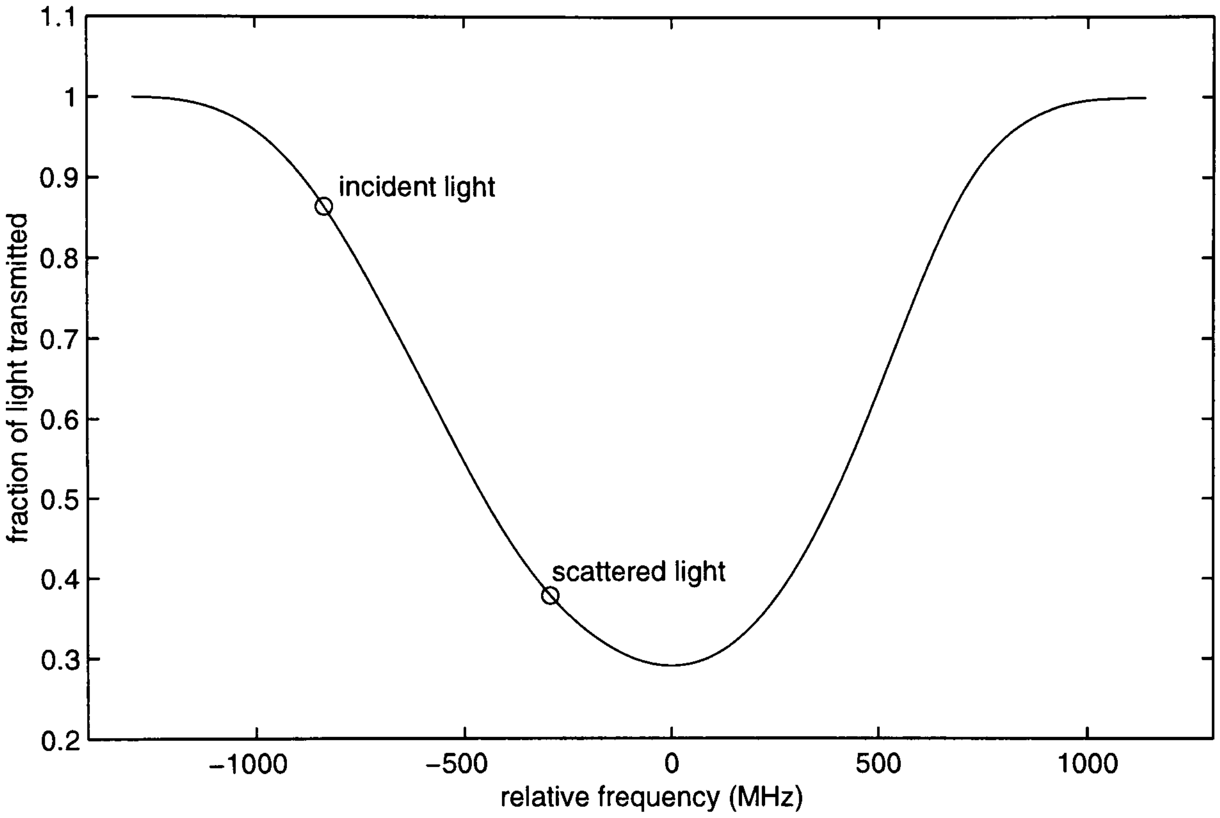


Figure 5.3 A typical transmission profile of an iodine cell, showing the shift in frequency and transmission ratio which light scattering by a moving particle might cause.

The need for constant absorption properties demands a constant density of iodine vapour in the cell. Ideally, constant vapour density would be achieved if the mass of iodine in the cell was sufficiently small to ensure that at operating temperatures, all the iodine would remain permanently in the gas state. Such a cell is known as a *starved* cell. In practice, however, it has been found difficult to control the quantity of iodine in the cell when it is sealed, since some is inevitably removed in the evacuation process. Of the four iodine cells used in the present work, one is a starved cell — the reference cell in the time-resolved DGV apparatus.

In the other cells (known as *saturated* cells), vapour and solid iodine exist in equilibrium. The temperature of the iodine crystals dictates their vapour pressure, and hence, the pressure of vapour in the cell. The crystals are accommodated in the *cold finger*, a hollow side arm. The temperature of the cold finger is monitored using a thermocouple and automatically controlled to within ±0.003 K, using a Peltier device to supply or remove heat as required. The temperature of the cold finger is typically held a few K below ambient temperature, so that crystals do not form elsewhere in the cell. In some DGV systems (such as that developed by Roehle and Schodl [44]), the need for such precise temperature control is removed by heating the cell to a temperature above the level at which all the iodine evaporates.

5.2.3 Imaging System

The optical imaging system developed in Oxford to realise the DGV concept is depicted in ***Figure 5.4***. Scattered light from the flow field is focused by the front lens to form a real image of the scene on the focal plane of the transfer lens, which acts as a collimator. The collimated light bundle is split into two paths, one of which passes through the iodine cell. Light on the two paths is then directed through a third lens and focused to form two side-by-side images in a digital charge-coupled device (CCD) camera. The whole image recorded by the CCD camera, containing both the reference and discriminated images, will always be referred to in this text as a *frame*, to distinguish it from the individual images of the flow field and avoid confusion.

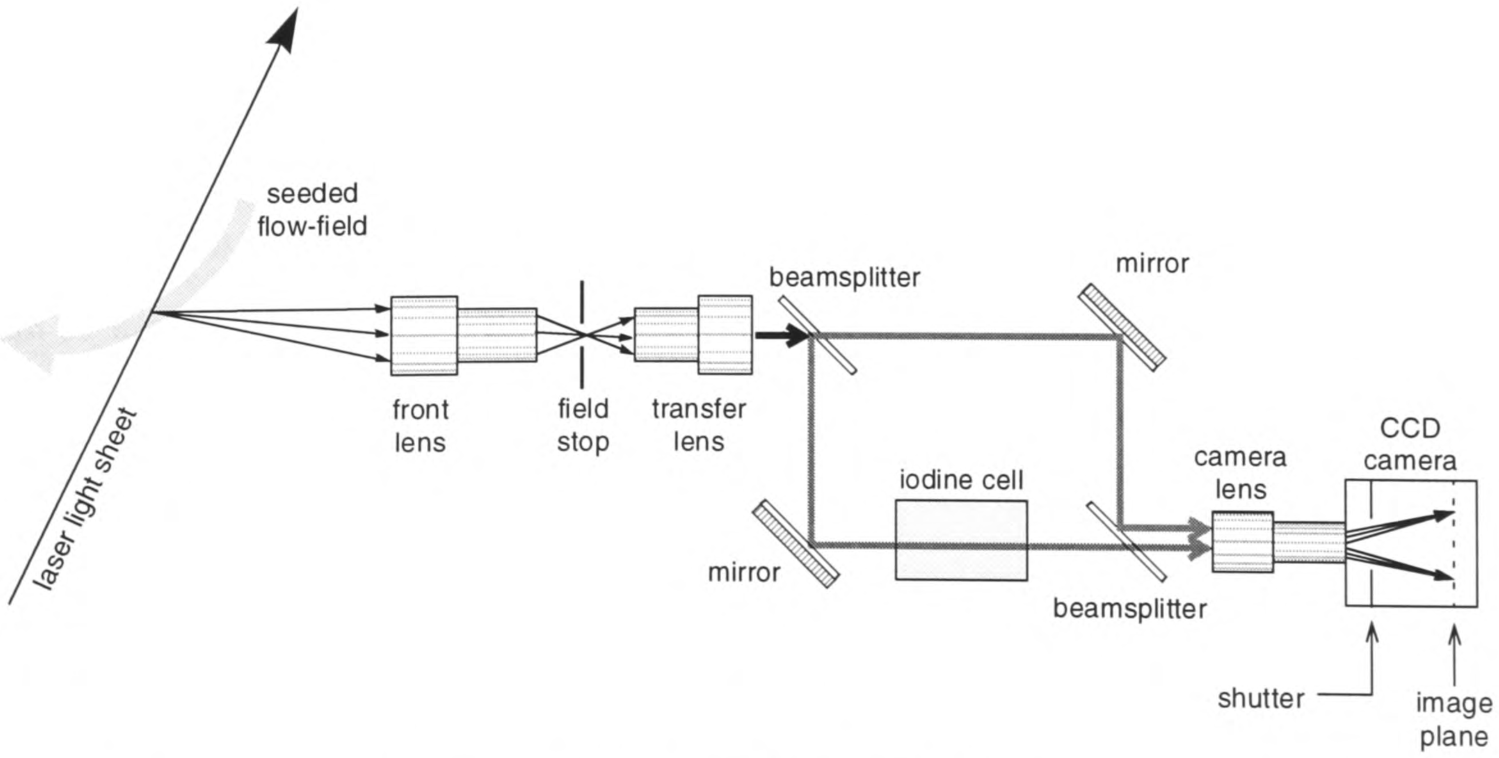


Figure 5.4 Schematic diagram of the DGV imaging system. The arrows show paths through the system are shown for light scattered from an arbitrary point in the flow-field.

In most DGV systems [34,40,44], each half of the split collimated bundle is focused in a separate camera. In the Oxford system, cost and complexity are reduced by acquiring both the reference and discriminated images on a single camera. This choice leads to a reduction in the size of the region over which measurement can be made at a given spatial resolution, since the number of pixels is halved with respect to a system in which two cameras are used. The mirrors and beamsplitters are adjusted so that each image occupies approximately one half of the final picture. A field stop, consisting of a plate with a rectangular slot to restrict the extent of the image, is located in the image plane between the front and transfer lenses. It ensures that the final reference and discriminated images have well-defined boundaries and do not overlap. To minimise optical aberrations and system size, all three lenses in the Oxford system are compound photographic units.

The camera, manufactured by Wright Instruments [45], is based on a Tektronix TK512-CF1 CCD chip. To reduce various types of image noise, the CCD chip is cooled to 200K by a Peltier device, and the readout time (the time taken to convert pixel intensity values to output digital signals) is long — that is, of the order of 1 s. The use of a slow-readout camera is feasible in the Oxford system because real-time measurements are not of interest. The CCD has 512x512 pixels, and pixel intensity values are digitised to 15 bits, equivalent to a precision of 1 in a maximum value of 32767. It is fitted with an electronically activated shutter, and interfaced to a personal computer which controls the camera and displays and stores the acquired frames.

Two complete imaging systems have been built in Oxford. In this research, one was used for time-integrated measurements, and the other was used for time-resolved work. The major difference between the two is in their lenses. In the system used for time-resolved work, lenses designed for use with 35 mm single lens reflex cameras were employed. The focal length of the front, transfer and camera lenses are 100, 50 and 135 mm, respectively. The front and transfer lenses used in time-integrated DGV, on the other hand, were intended for large-format cameras, and have larger absolute aperture diameters. Consequently, the diameter of the collimated bundle in the latter system is larger. The focal lengths of the lenses in this system are 60, 150 and 300 mm.

5.2.4 Illumination

This section deals with the processes by which light with the properties required for DGV is generated and delivered to the flow field to illuminate scattering particles. The most important properties required of DGV illumination are that the frequency of the light should fall within the span of an absorption feature of iodine, and that the frequency spectrum of the light should be narrow in comparison with the width of the absorption feature, which is of the order of 1 GHz. Radiation of such a narrow spectrum, at appropriate frequencies, can be provided by certain types of laser with special frequency-narrowing features. Illumination for DGV should also be accurately collimated, since the propagation directions of both the incident and scattered light must be known in order to convert measured frequency shifts into velocity data. Laser light meets this criterion.

In this section, the theory of operation of lasers is given in a bare outline, and the specific lasers used for DGV in Oxford are described. This information is drawn largely from a text by Beesley [46] and from the laser manufacturers' documentation [47-49].

The origin of radiation in a laser is the transition of electrons in atoms of a chosen material, the *lasing medium,* from high to low energy levels. On falling to a lower level, an electron emits a photon at a frequency ν which is related to the change in energy, ΔE, by:

$$\Delta E = h\nu \quad , \tag{5.2}$$

where h is Planck's constant. Many such transitions are possible for every type of atom, each one possessing a characteristic frequency and energy change. Transition and emission of this kind may occur spontaneously (at random), or may be stimulated by the absorption of a photon. In the latter case, known as *stimulated emission*, two photons are emitted, and are identical to the incident photon in frequency, phase and direction. Stimulated emission may be viewed as an amplification mechanism, since each incident photon is replaced by two copies of itself.

Atoms of the lasing medium are excited, by means of one of several *pumping* methods, to a higher energy level. Some atoms decay spontaneously, and the photons emitted by some of them are subsequently duplicated through stimulated emission. The physical arrangement of the laser selectively promotes this amplification. The lasing medium is enclosed in a cavity bounded by two mirrors, facing each other and mounted perpendicular to the laser axis. Photons which travel normal (or almost normal) to the mirrors are reflected back into the cavity, where they continue to stimulate the release of new photons with the same properties. Off-axis photons, on the other hand, are rejected. In this way, a parallel beam is established, growing in intensity as it traverses the cavity. Its frequency spectrum is dictated by the properties of the atomic transition (or transitions) involved in stimulated emission. One of the mirrors, called the *output coupler*, is partially transmissive, so that a fraction of this beam escapes to become the laser output. The other mirror is called the *high reflector*.

Amplification in the cavity depends not only on the direction of photons, but also on their frequency and phase. The mirrors form an optical resonator, and only light for which the cavity length (from mirror to mirror) equals a whole number of half-wavelengths can repeatedly traverse the cavity without destructively interfering. In terms of frequency, only the discrete frequencies or *longitudinal modes* which satisfy ***Eq. 5.3*** are amplified, where m is any integer, c is the speed of light and L is the mirror separation or cavity length.

$$\nu_m = \frac{mc}{2L} \tag{5.3}$$

Interference in the cavity also ensures that all amplified radiation is in phase.

For time-integrated DGV in Oxford, a Spectra Physics model 168 laser is used. This laser is described as *continuous wave*, meaning that its emission is nominally constant in time. The lasing medium is ionised Argon, pumped by a continuous electrical discharge (of up to 40 A) through the tube which contains the gas. In its simplest configuration, this laser outputs radiation in many spectra associated with different atomic transitions. Oscillation is restricted to the wavelength of interest for DGV, green 514.5 nm, by a prism, located in the cavity near the high reflector. Different wavelengths are refracted through the prism at different angles; by rotating the prism, the unwanted wavelengths are deflected off the cavity axis. The radiation is further restricted, to a single longitudinal mode within the spectrum of the 514.5 nm transition, by an *intracavity etalon* (Spectra Physics model 583). This device consists of a glass cube, mounted in the laser cavity, with two partially reflective faces aligned perpendicular to the laser axis. Radiation at any frequency not coinciding with one of the etalon's resonant modes (which are determined by the distance between the reflective faces) is suppressed. Fine control of the etalon length allows a particular mode of the host laser to be selected; this control is accomplished by a device which regulates the etalon's temperature to within ±0.01 K. As the etalon length changes, the laser frequency switches discontinuously to the mode which is closest to the frequency at which attenuation in the etalon, at its current length, is a minimum. This behaviour is called *mode hopping*. Single-mode output power of up to 400 mW has been obtained from the Argon ion laser in the present work.

For time-resolved experiments, a Spectra Physics GCR-130-10 laser [48] was used. This is a pulsed laser, delivering light in bursts of the order of 10 ns in duration. As this time is shorter than any time scales of practical interest in fluid dynamics, this laser can be used to acquire essentially instantaneous images, and hence instantaneous velocity maps, of unsteady flow fields. The active medium is a rod of Nd:YAG, which consists of yttrium aluminium garnet (YAG), a crystalline solid, doped with neodymium (Nd). This laser's operation is based on a transition of the Nd atom corresponding to a wavelength of 1064 nm, which is in the infrared.

The Nd atoms are pumped by flashes of white light of some 200 µs duration. The cavity contains a *Q-switch*, a device whose optical transmittivity can be electrically switched, very rapidly, between a low and a high value. The laser may be operated in *long pulse* mode, in which case the Q-switch remains open and laser action takes place over the entire duration of the flashlamp pulse. However, if the Q-switch is closed when the flash lamp pulse begins, oscillation of the laser cavity is prevented, and a large store of energy is built up while the Nd atoms are pumped. This energy is released in a brief, intense period of laser oscillation, which begins when the Q-switch is opened.

A Spectra Physics model 6530 *injection seeding* system [49] was installed to restrict the oscillation of the Nd:YAG laser in Q-switched operation to a single longitudinal mode. The seeding system includes a continuous wave diode-pumped Nd:YVO_4 laser which is intrinsically restricted to single-mode operation (within the bandwidth of the 1064 nm transition on which the Nd:YAG laser action is based), with an output power of the order of 1 mW. The seed laser's beam is injected into the Nd:YAG rod, where it stimulates emission in pumped Nd atoms. In the absence of seeding, oscillation of the Nd:YAG cavity is initiated by spontaneous emission events, releasing photons at frequencies randomly distributed across the spectrum of the 1064 nm atomic transition in Nd, which is approximately 30 GHz wide. The seed laser beam, however, has a narrow frequency spectrum, and is orders of magnitude more intense than any spontaneously emitted radiation. Consequently, oscillation at the seeded frequency builds up much more rapidly than the oscillation initiated by broadband spontaneous emission. The energy stored in the pumped Nd atoms is harvested in oscillation of the seeded mode before significant oscillation builds up in any other modes; the result is a single-mode pulsed laser beam.

Radiation at the seed laser frequency can only be amplified if that frequency matches a resonance of the Nd:YAG cavity. This is ensured through active automatic control of the Nd:YAG cavity length, which is accomplished by translating the high reflector on a piezoelectric stage. When the length of the cavity permits oscillation at exactly the frequency of the seeder, seeding is optimal, and the time taken from opening of the Q-switch for laser action to reach its maximum intensity is at a minimum. This time lapse, called the Q-switched build-up time, is used by the cavity length control system as a criterion to assess the quality of seeding. The frequency of the seed laser depends on its temperature, which is stabilised by another control system. The user can vary the seeder frequency (and hence the frequency of the seeded Nd:YAG output) by varying an input voltage (called the *seed laser control voltage*) which determines the set point of the temperature control system.

The 1064 nm infrared Nd:YAG output beam is passed through a *harmonic generator*, the main component of which is a crystal of potassium dideuterium phosphate (KD*P). This has the effect of doubling the frequency of a fraction (specified as 44%) of the radiation. The resulting green 532 nm light is separated from the remaining infrared by a dichroic (wavelength sensitive) beamsplitter. It is the green beam which is of value in DGV, as its (unseeded) frequency spectrum includes several iodine absorption features, and the frequency of single-mode seeded radiation can be tuned across a portion of this range. The efficiency of frequency conversion in the harmonic generator is strongly sensitive to the orientation of the KD*P crystal, and rotation of the crystal is a convenient means of varying the energy in the green pulse without altering the state of the laser. The typical output energy of the Nd:YAG laser in the single-mode 532 nm beam is 80 mJ per pulse.

5.2.5 Illumination Frequency Monitoring

The frequency of light scattered from the flow-field is measured through analysis of the flow field images. Knowledge of the frequency of the light at emission from the laser is also necessary to compute the Doppler shift, which is the difference between the two. In the Oxford DGV system, measurement of the laser emission frequency is enabled by a second iodine cell, the *reference cell.* The reference cell arrangement in the time-integrated system is shown in ***Figure 5.5***. A small fraction of the laser emission is sampled from the main beam, and then split into two beams of roughly equal intensity, one of which passes through an iodine cell. The intensities of both beams are monitored using photodiodes (type BPX65), whose signals are amplified and passed on to a computer via an analogue to digital converter. The ratio of the signals is proportional to the iodine cell's transmission ratio. Inevitable differences in photodiode sensitivity, amplifier gain and layout of the two optical paths cause the photodiode signal ratio to differ from the true physical transmission ratio. Knowledge of the absolute physical transmission ratio is unnecessary — any measured version of the transmission ratio which depends consistently on frequency is adequate.

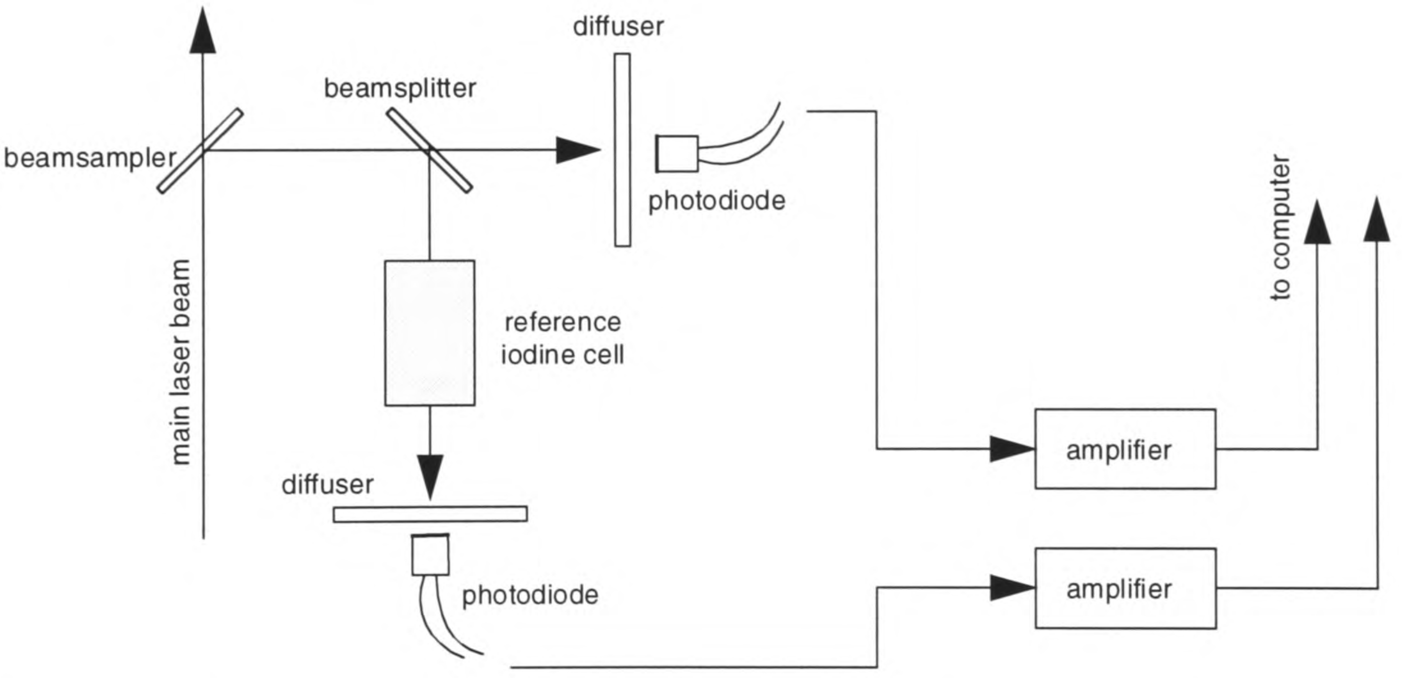


Figure 5.5 A schematic diagram of the reference cell system in the time-integrated DGV apparatus.

The reference cell arrangement in the time-resolved system is slightly different. The sampled beam illuminates a white target, which reflects light diffusely towards both photodiodes. Because of the short duration of the pulse, the photodiode signals cannot be monitored continuously. Instead, each signal is passed through an integrating amplifier, whose output rises, after each laser pulse, by an amount which is proportional to the light energy delivered to the photodiode during that pulse.

5.2.6 Image Processing

The raw data in a DGV experiment are captured in the form of a CCD frame comprising two similar images of the same flow field. The term *image processing*, in the context of DGV, refers to the sequence of operations by which this frame is translated into the velocity distribution which is the eventual output of the experiment. Various developers of DGV have attached differing levels of importance to image processing, consequently adopting quite different approaches to it. The Oxford group placed a strong emphasis on image processing from an early stage, and made a substantial effort in the development of software tools. In this section, the rationale for image processing is explained, and the Oxford method is outlined.

For any given point in the physical measurement plane it is required to generate a velocity value. There is a one-to-one mapping from this point to a point in the reference image formed on the CCD array, and another mapping to a point in the discriminated image. These conceptual mappings have a very real manifestation in the multitude of optical paths along which light propagates from particles in the flow field to the CCD array. The light which arrives at the point in the discriminated image is attenuated, through interaction with the iodine vapour, by a fraction which depends on its frequency.

The purpose of the image processing stage of a DGV experiment is to calculate the ratio of the image intensity at each point in the discriminated image to the intensity at the corresponding point in the reference image. The major challenge is to identify corresponding pairs of points correctly. This task, called *alignment*, is non-trivial because the reference and discriminated images are offset on the CCD array by an initially unknown spatial shift; because they are deformed through perspective and optical effects; and most significantly, because these deformations are not, in general, identical for both images.

The approach taken in the Oxford DGV system, among others, is to calibrate the transformations from the geometry of the measurement plane to the geometry of both images, by analysing the positions in which known fiducial marks on a test card are imaged. A numerical specification of these transformations is stored, and is later recalled for each individual velocity measurement and used to map pixels in both the reference and discriminated images of the flow field into common, comparable positions. All analysis and manipulation of images is implemented in software. The current Oxford image processing technique is described in some detail by Manners *et al.* [42].

The calibration stage, which will be referred to as *alignment calibration*, must be carried out whenever the imaging system or the light sheet is moved, or when the imaging system is adjusted internally. A test card, consisting of a regular grid of white circular dots on a black background, is located in the measurement plane, and illuminated with diffuse laser light. A CCD frame containing both reference and discriminated images of the card is captured. Images of a white "normalisation" card and a black card are also captured, under the same illumination. Each frame is passed to a PC and saved immediately after it is acquired.

Sample images are shown in ***Figure 5.6***, in a schematic diagram of the alignment calibration image processing procedure. The first frame shown is the raw CCD frame of the dots card. The reference and discriminated images are side by side, delimited by a dark vertical band at the centre of the frame. This band is, in fact, an image of the field stop. These particular images were chosen as an extreme example of the effects which make alignment necessary. Note that the grid of dots is visibly deformed by perspective effects due to a non-normal viewing of the measurement plane, and that the image intensity (brightness) of the dots varies across the frame because of illumination and imaging non-uniformities. Curvature of the columns of dots is also present, though barely detectable — it indicates optical distortion, due to a severely non-normal viewing angle and short viewing distance.

The black frame is subtracted from the dots and normalisation frames, to cancel the effect of zero offset in the CCD camera electronics, and of scattering from the black background of the test card. The black-corrected dots frame is then divided, pixel by pixel, by the black-corrected normalisation frame. This removes intensity variations due to non-uniformity in illumination and in the optical hardware, and provides a frame in which image intensity ranges from zero, in background regions, to a uniform value near unity, inside images of the dots. Precise coordinates are then assigned to each dot by locating its centroid. The normalisation ensures that centroid locations are not biased by gradients in illumination intensity.

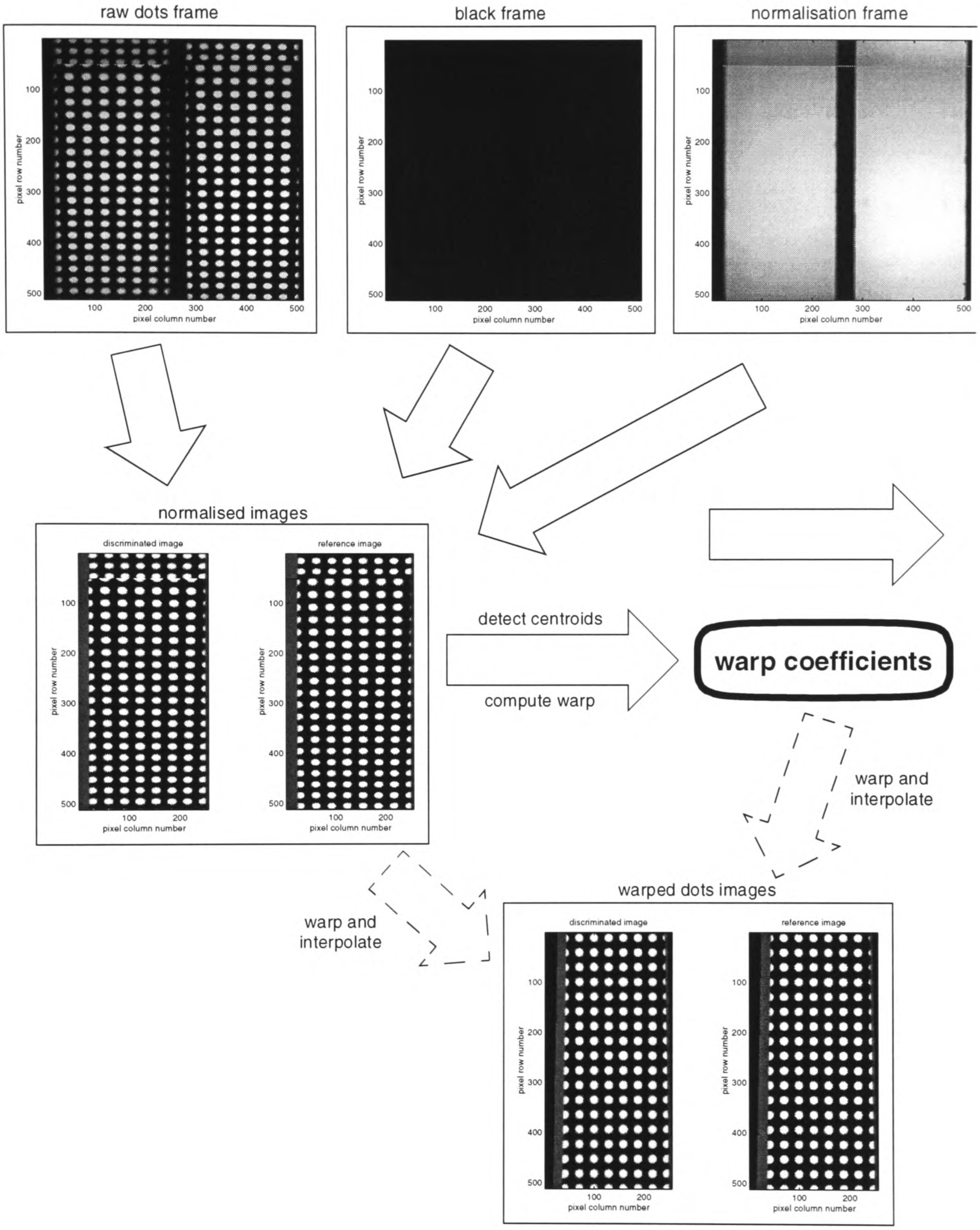


Figure 5.6 Schematic diagram of the process leading to computation of warp coefficients. Computation of the warped dots images is an optional stage which demonstrates and checks the warp.

A geometrical transformation is now computed which maps the centroid locations in each image onto a uniform grid of unit spacing — a distribution which is geometrically similar to the true distribution of the dots on the physical test card. The transformation, known as a *warp*, is a set of polynomial functions. It contains all the information that is required to mutually align the reference and discriminated images of a real flow field so that their ratio can be computed.

Application of the warp to every pixel in the reference and discriminated images transforms these images onto two different irregular arrays of pixels. In this form, the images are still not quite prepared for direct comparison. The final step is an interpolation which re-evaluates the warped image intensity distributions on identical regularly spaced pixel grids.

The function of the warp and interpolation operations is best illustrated by a pair of warped dots images, shown at the bottom of ***Figure 5.6***. The dots in the warped images are laid out on a regular grid (as they are on the test card which was originally imaged), without distortion or perspective effects. More importantly, dots appear in the same positions in the discriminated and reference images. This fact is demonstrated more precisely in an image intensity profile taken through the centre of a row of dots in both images (***Figure 5.7***). The dot cross-sections are aligned to within a fraction of a pixel width. Further intensity profiles of warped dots images, in this format, will be presented at various stages in this thesis to support discussion of alignment accuracy and optical performance.

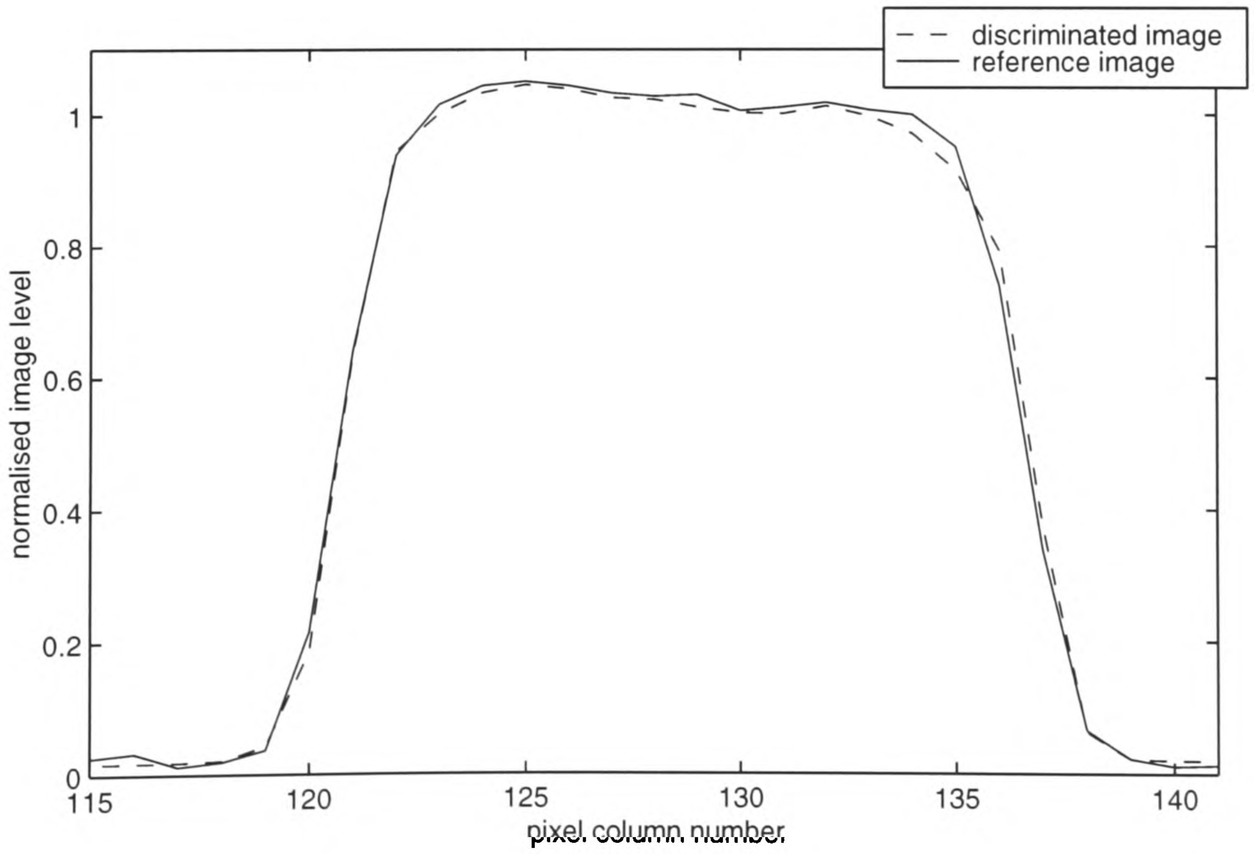


Figure 5.7 Intensity profiles in warped reference and discriminated images of alignment dots.

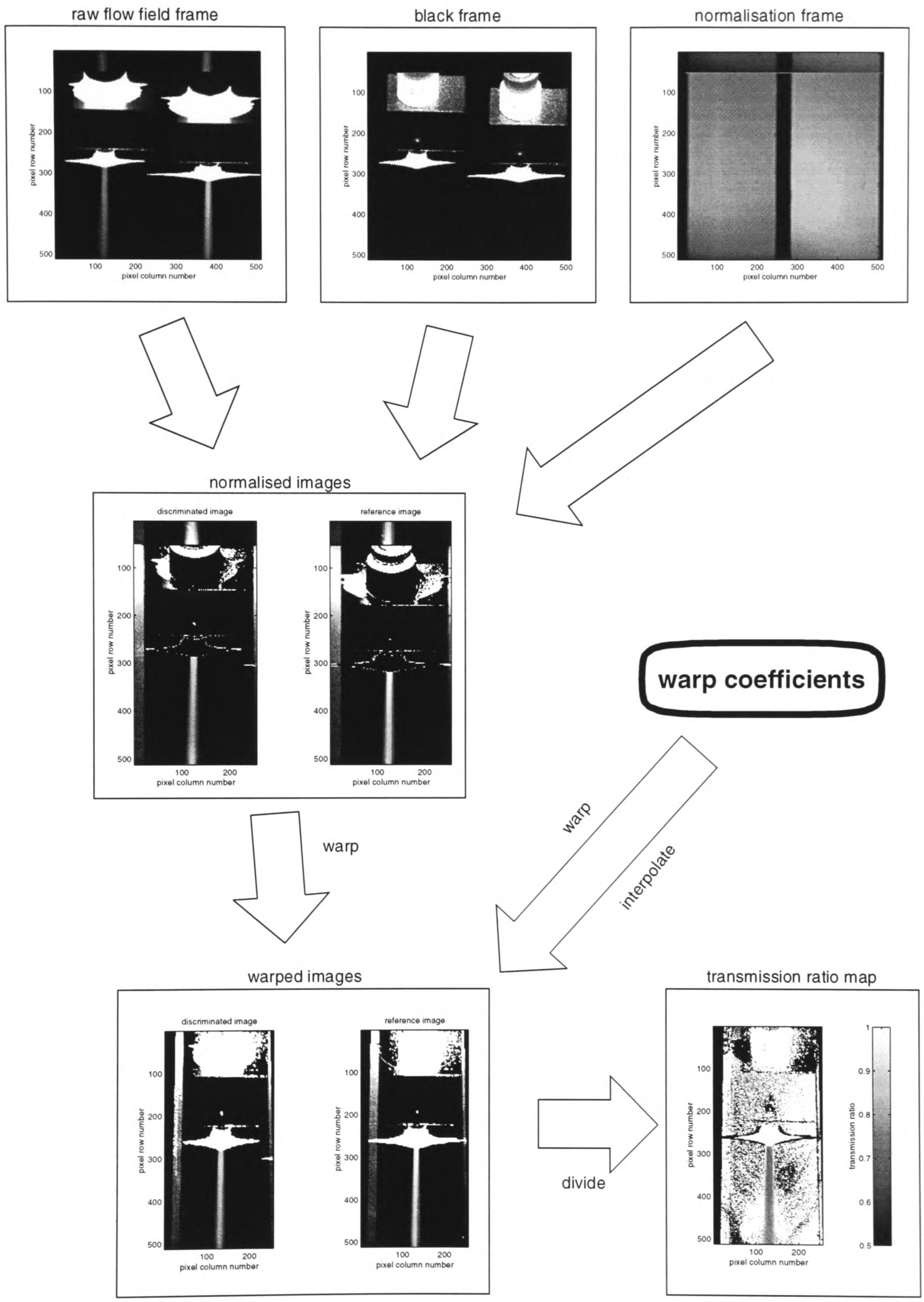


Figure 5.8 Schematic diagram of the processing of a flow field CCD frame

Once a set of warp coefficients have been computed and stored, they can be used with all flow field images acquired as long the imaging configuration remains unchanged. The various stages of image processing of a typical flow field frame, leading to a computed distribution of transmission ratio, are depicted in ***Figure 5.8***. The flow field itself appears in the flow field images as a light vertical band. As in the alignment calibration process, black and normalisation frames are required. The black images are images of the flow field scene, under the illumination conditions used for the measurement, in the absence of flow and light-scattering particles. The black frame is subtracted from the flow field frame to suppress the offset in the CCD camera output and light scattered from background objects. The normalisation images, again, are images of a stationary white target. In this case, their function is to correct for any light intensity differentials between the two sides of the imaging system which are not due to the iodine cell. Variations between individual mirrors and beamsplitters, and between CCD pixels, may cause such differentials.

The warp coefficients, computed and stored in an earlier alignment calibration procedure, are recalled and used to warp the normalised images. Each pixel in the discriminated warped image, and the pixel at the same position in the reference image, now correspond to images of the same location in the measurement plane. Differences between these images are due solely to light absorption by iodine vapour. It is straightforward to compute the ratio of discriminated image intensity to reference image intensity, pixel by pixel, to obtain a map of transmission ratio. This ratio is multiplied by the transmission ratio for the frequency of the light with which the normalisation card was illuminated. The resulting ratio values correspond directly to transmission ratios measured in calibration of the imaging iodine cell. They are looked up in that calibration to give frequency (relative to the arbitrary zero selected in the calibration). The difference between this frequency distribution and the relative frequency of illuminating light is the Doppler shift frequency, which translates, via ***Eq. 5.1***, into velocity.

This manipulation of pixel intensity values in the warped and interpolated images can be represented more clearly in mathematical terms. Let I be the total light energy, scattered from a particular flow field location, which is gathered by the front lens of the imaging system and delivered to the first beamsplitter, within the collimated bundle. Let ν_s be the shifted frequency of this light. Then, light with energy $k_d T(\nu_s) I$ and $k_r I$ arrive at pixels in the discriminated and reference images, respectively, where k is an attenuation constant associated with optical elements and CCD pixels, the subscripts d

and r refer to the discriminated and reference images, and T is the imaging cell transmission ratio (a function of frequency). If N is the light energy gathered in acquiring the normalisation images, then the intensities of the normalised flow field images are given by

$$I_d = \frac{k_d T(\nu_s) I}{k_d T(\nu_n) N} = \frac{T(\nu_s) I}{T(\nu_n) N} \tag{5.4}$$

in the discriminated image, and

$$I_r = \frac{k_r I}{k_r N} = \frac{I}{N} \tag{5.5}$$

in the reference image. If the warp is determined with perfect accuracy in the alignment calibration, then the measured ratio of discriminated to reference image intensity for this pixel will be:

$$R = \frac{I_d}{I_r} = \frac{T(\nu_s)}{T(\nu_n)} \ . \tag{5.6}$$

On multiplication by the normalisation transmission ratio value $T(\nu_n)$, the contribution of the normalisation images to the overall image ratio is removed, and the above expression reduces to $T(\nu_s)$, the transmission ratio associated with the frequency of scattered light. That frequency, ν_s, is then calculated by interpolation on the calibrated profile of the imaging iodine cell, an example of which was shown in ***Figure 5.3***. Similarly, the laser emission frequency ν_0 is computed from the measured reference cell transmission ratio and its calibration. The Doppler shift, $\Delta\nu = \nu_s - \nu_0$, is then substituted into the Doppler equation, repeated here:

$$\Delta\nu\lambda_0 = \boldsymbol{V} \cdot (\hat{\boldsymbol{o}} - \hat{\boldsymbol{\imath}}) \ . \tag{5.7}$$

λ_0 is known to sufficient accuracy from the laser specifications as 514.5 nm for the Argon ion laser [47] or 532 nm for the Nd:YAG laser [48]. The magnitude of the vector $\hat{\boldsymbol{o}} - \hat{\boldsymbol{\imath}}$ is calculated from an analysis of the measurement geometry. The value of the velocity component in the direction of $\hat{\boldsymbol{o}} - \hat{\boldsymbol{\imath}}$ can then finally be calculated.

5.2.7 Iodine Cell Calibration

The imaging apparatus and the image processing methodology are designed to deliver a measurement, for every pixel in a warped interpolated image of the flow-field, of the fraction of light, destined for that pixel, which is transmitted by the iodine cell. Conversion of this information into Doppler shift information requires a calibration of the dependence of transmission ratio on frequency.

All the cells used in the present work were calibrated in their installed configurations. This removes any possibility that the distance traversed by light passing through the cell, or its interaction with the cell windows and walls, may be different under calibration and operational conditions. In general, calibration is approached by monitoring the transmission ratio of the cell in question while sweeping the laser emission frequency over the range of interest, and measuring that frequency by a standard which is independent of the cell.

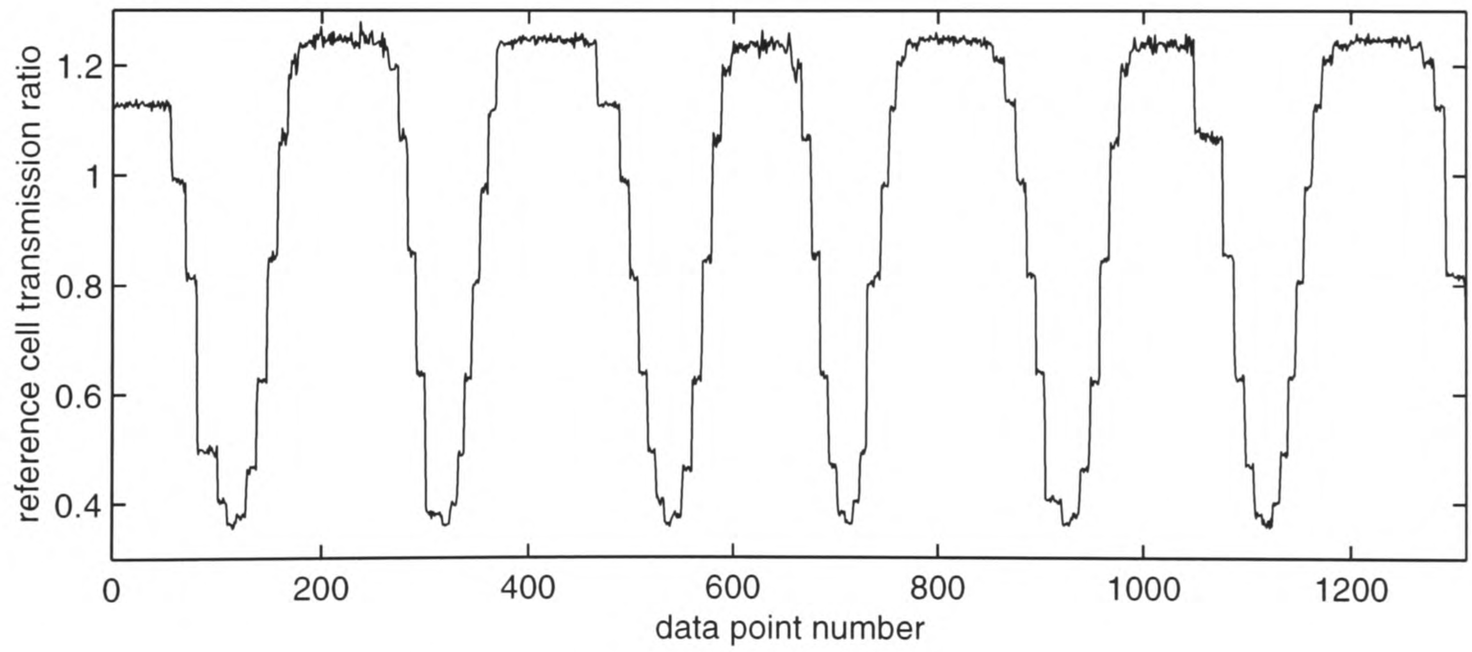


Figure 5.9 Variation of measured transmission ratio during calibration of the reference iodine cell for use in time-integrated DGV, with 6 successive scans of the iodine absorption feature.

In calibrating an imaging cell, a reference cell which has already been calibrated is used to measure frequency. The frequency standard against which a reference cell is calibrated depends on the laser with which it is to be used. For cells used in the time-integrated DGV apparatus, the 143 MHz spacing between discrete longitudinal modes of the Argon ion laser (explained in §5.2.4) provides a convenient yardstick of frequency. To perform a calibration of the reference cell, mode hops are induced by varying the laser's etalon temperature while monitoring the transmission ratio of the reference cell. The laser frequency is swept back and forth across the useful bandwidth of the iodine absorption feature several times in discrete mode hops. Continuous drift of laser frequency between mode hops (due to variation of the cavity length with temperature) is minimised by allowing the laser several hours from start-up to reach thermal equilibrium, and by carrying out the calibration as quickly as possible. The variation of transmission ratio during such a calibration procedure is shown in ***Figure 5.9***. In subsequent data processing, a transmission ratio value is identified for every period of nominally constant frequency between mode hops. If frequency drift between hops is negligible, the laser frequency should revisit each mode, or frequency value, several times as it repeatedly scans across the iodine cell's bandwidth. In this way, multiple measurements of transmission ratio are taken for each mode, and their average is calculated. The final calibration curve is a spline through these average transmission ratios. This last stage of the procedure is illustrated in ***Figure 5.10***.

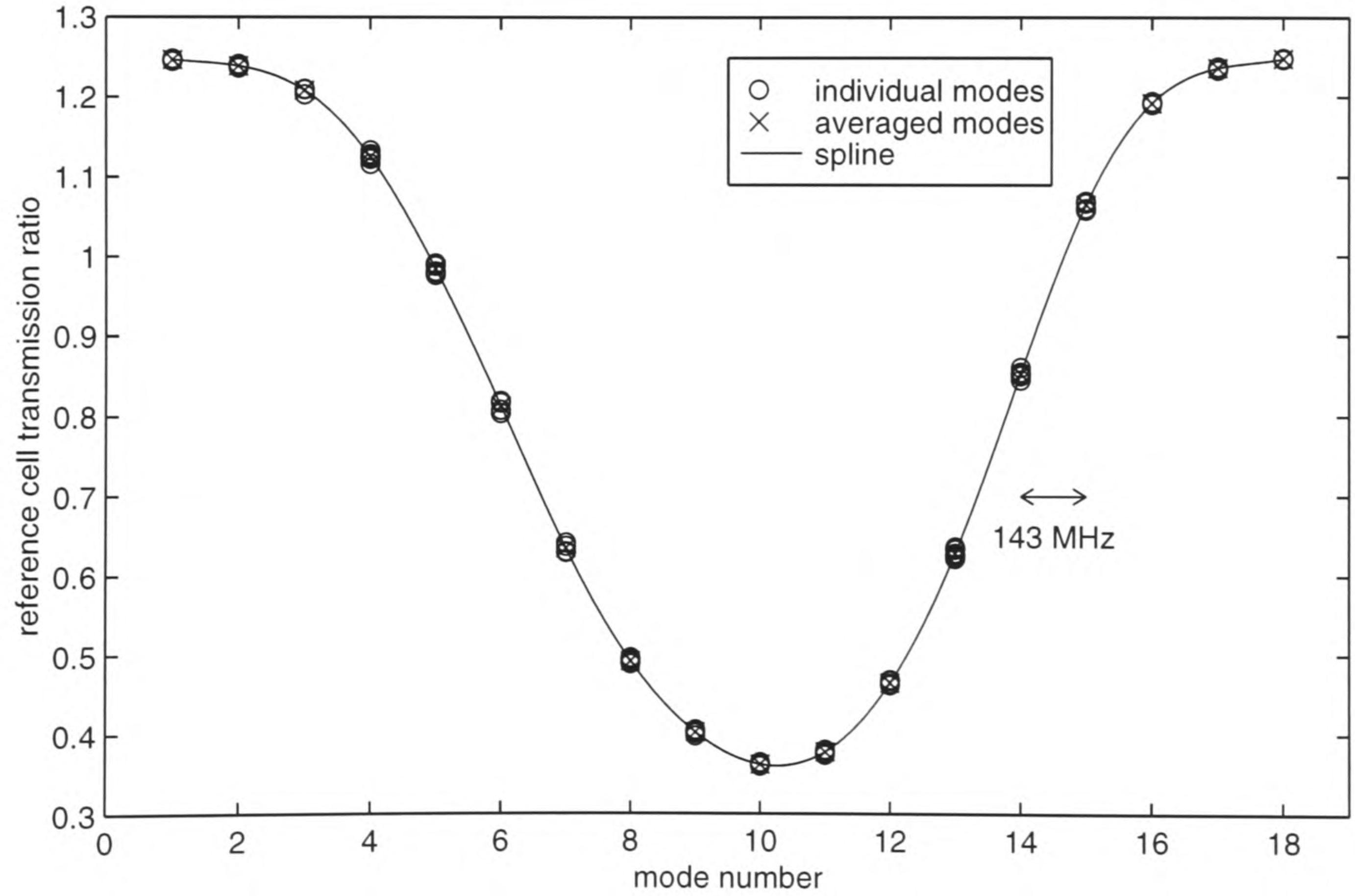


Figure 5.10 Processed data from a reference cell calibration, showing the transmission ratio values measured on multiple passes of each laser mode, their averages for each mode, and the final spline through these points, which is used as the calibrated cell profile.

A different technique is used for calibration of the reference cell in the apparatus for time-resolved DGV, since the Nd:YAG laser does not undergo mode hops. A *Fabry-Perot inteferometer* is used to measure the laser emission frequency. This device, which is analysed in detail by Hercher [50], consists of a pair of partially transmissive convex spherical mirrors facing each other across a cavity which is equal in length to the mirrors' radius of curvature. When monochromatic light is passed through the device, an interference pattern consisting of concentric rings can be observed. The distribution of the rings, given by ***Eq. 5.8*** below, depends on the frequency of the light.

$$\rho_m = \left[\left(m - \frac{4r}{c}\nu\right)r^3\lambda\right]^{1/4} \tag{5.8}$$

ρ_m is the radius of the *m*th ring, *r* is the radius of curvature of the mirrors in the interferometer, *c* is the speed of light, λ is the absolute wavelength, and ν is the frequency relative to an arbitrary zero (in effect, determined by the arbitrary designation of ring $m = 1$). λ is specified as 532 nm for the frequency-doubled Nd:YAG output, and the quantity $c/(4r)$, the *free spectral range* of the Fabry-Perot device, is specified as 2 GHz. This represents the range within which frequencies can be uniquely determined; if values of frequency are permitted for which $(4r/c)\nu$ exceeds 1, ambiguity may occur.

The emission frequency of the seeded Nd:YAG laser was stepped by adjusting the frequency control input voltage on the seed laser, allowing a few seconds after each adjustment for the Nd:YAG cavity length to adjust to the new seed laser frequency. Meanwhile, the Fabry-Perot inteferometer was used to measure the laser emission frequency, while the reference cell transmission was monitored using the apparatus described in §5.2.5.

The green output beam of the Nd:YAG laser was directed onto a white paper screen in front of the Fabry-Perot interferometer. Images of the resulting interference patterns were acquired on a CCD camera which, for convenience, was used *in situ* in the DGV imaging system. The camera was exposed for periods of 25 – 60 s, so that measured frequencies represent averages over about 250 to 600 laser pulses. Ring patterns were analysed only in the image formed on the reference side of the CCD frame, though the discriminated side of the frame would have been an equally valid choice.

A sample image of a ring pattern is shown in ***Figure 5.11***. The very bright outermost ring is actually an image of the structure of the Fabry-Perot interferometer. The ring radii ρ_m were determined in pixels as half the distance (along a horizontal diameter) between the brightest pixels in opposite sides of each ring. Because the mirror radius *r* is not known in terms of pixels, it is effectively unknown. Given the radii of two rings, *i* and *j*, *r* can be eliminated from ***Eq. 5.8*** to give:

$$\xi = \frac{j\rho_i^4 - i\rho_j^4}{\rho_i^4 - \rho_j^4}, \qquad (5.9)$$

where ξ is defined as $(4r/c)\nu$. Inspection of this expression shows that errors in ρ_i and ρ_j will translate into large errors in frequency if the difference between ρ_i and ρ_j is small. However, since the spacing between neighbouring rings decreases as radius increases, it is advantageous to use rings of low radius. Therefore, in images where more than two rings are available, the innermost ring is taken along with each of the other rings in turn to determine several values of ξ (and r), which are then averaged to yield a final value. This approach systematically avoids ring pairings with the smallest spacings. The plot in the right hand side of ***Figure 5.11*** shows the six ring radii measured in the accompanying fringe image. As a check on the average values of ξ and r computed from the five pairs of rings provided by this image, they were substituted into ***Eq. 5.8*** to generate a prediction of the variation of ring radius. The prediction is also shown in the figure, and gives a satisfactory fit to the raw points.

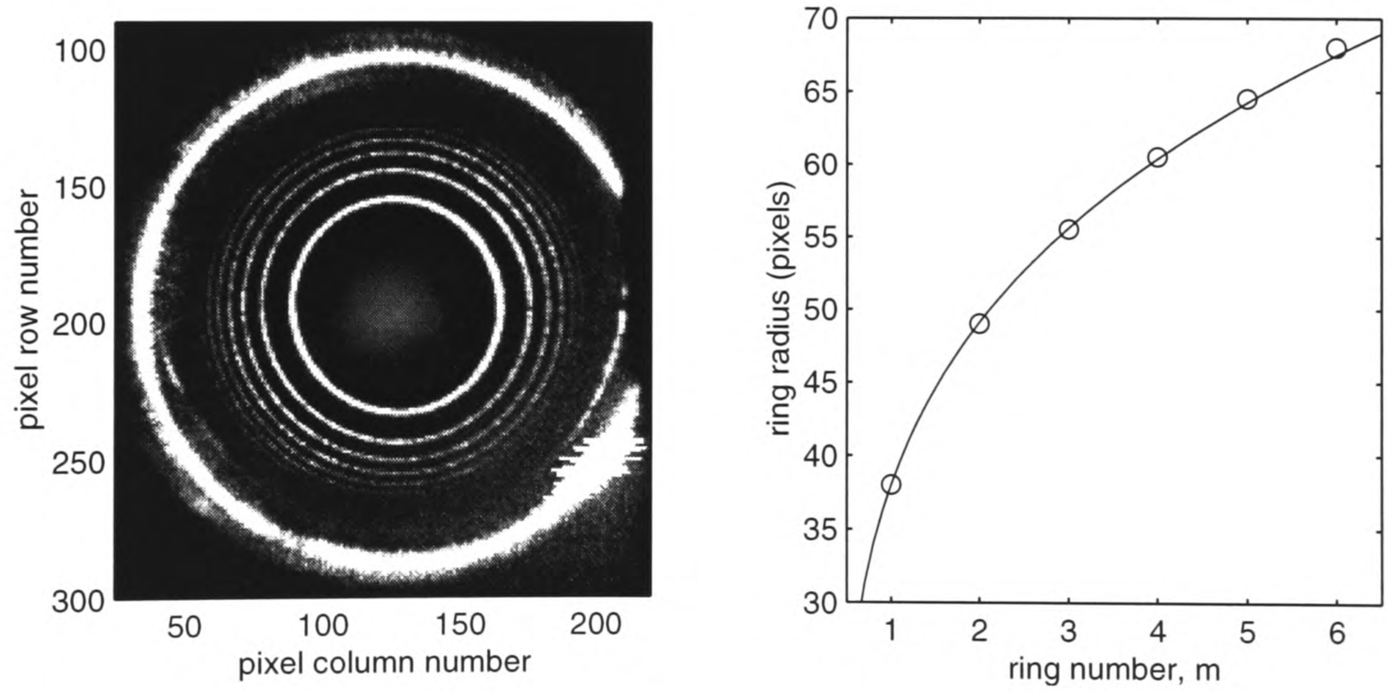


Figure 5.11 *Left:* an image of a typical fringe pattern observed with a Fabry-Perot interferometer. *Right:* ring radii in that image, along with the theoretical prediction (solid curve) based on average frequency and mirror radius computed from the measured ring radii.

Values of ξ, and hence of frequency, were calculated in this way from ring patterns imaged at each of 24 settings of the seed laser control voltage. The results are plotted in ***Figure 5.12***, and suggest a linear relationship between frequency and control voltage. As shown in §5.3.2 of this thesis, the deviations from a fitted linear correlation are consistent with the finite precision (of one pixel) to which the ring radii are measured. This limit on accuracy has been incorporated into an uncertainty analysis of time-resolved measurements, which is presented in §5.3.2.

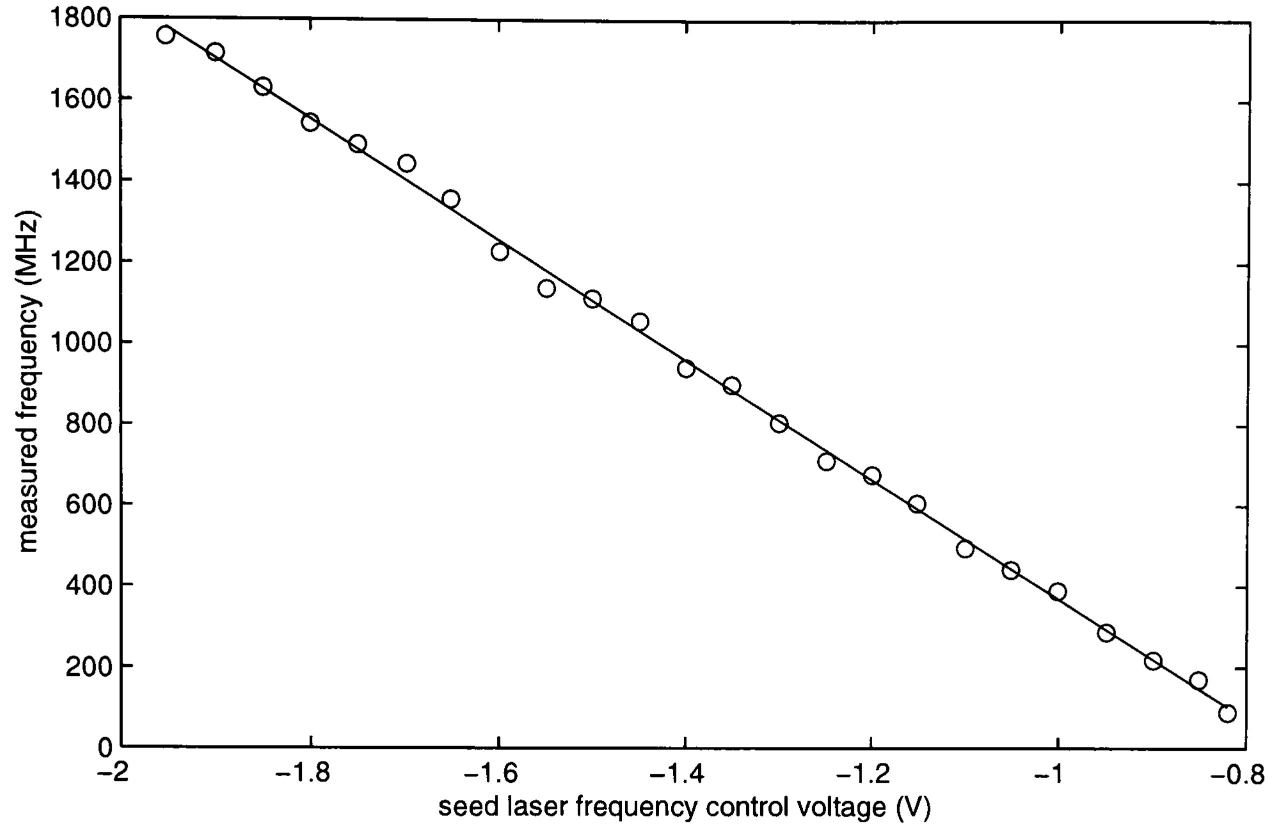


Figure 5.12 Measured frequency of Nd:YAG laser emission, plotted as a function of the seed laser control voltage, with a least-squares linear fit.

The apparent linear dependence of frequency on seed laser control voltage is taken as the correct description of the operation of the Nd:YAG laser. Hence, the working calibration is based on the variation of reference cell transmission ratio (recorded as an average over the exposure period of each Fabry-Perot interference pattern image) with frequency (based on a linear fit to the measured frequency values as a function of seed laser control voltage). The final calibrated transmission profile is a spline curve fit to this variation. This profile is shown in ***Figure 5.13*** along with the raw data on which it is based.

The procedure for calibration of the imaging iodine cell was the same for both time-resolved and time-integrated DGV systems. A white test card was illuminated with diffuse laser light, the frequency of which was varied in steps. At each frequency setting, a CCD frame of the card (consisting, as usual, of reference and discriminated images) was captured and stored, while the reference cell transmission ratio was monitored. A black frame was also taken, by exposing the CCD in the absence of illumination. In postprocessing, the ratio of discriminated and reference image intensities was calculated and averaged over the largest possible region in every frame. For some early imaging cell calibrations, an alignment calibration was carried out and transmission ratios were calculated rigorously on the basis of the warped images. However, it was found that these transmission ratios differed negligibly from the ratios computed when the images were aligned crudely by estimating the offset between them in the frame. The fact that image intensity gradients were low, and the use of spatial averaging over thousands of pixels, made precise alignment unnecessary.

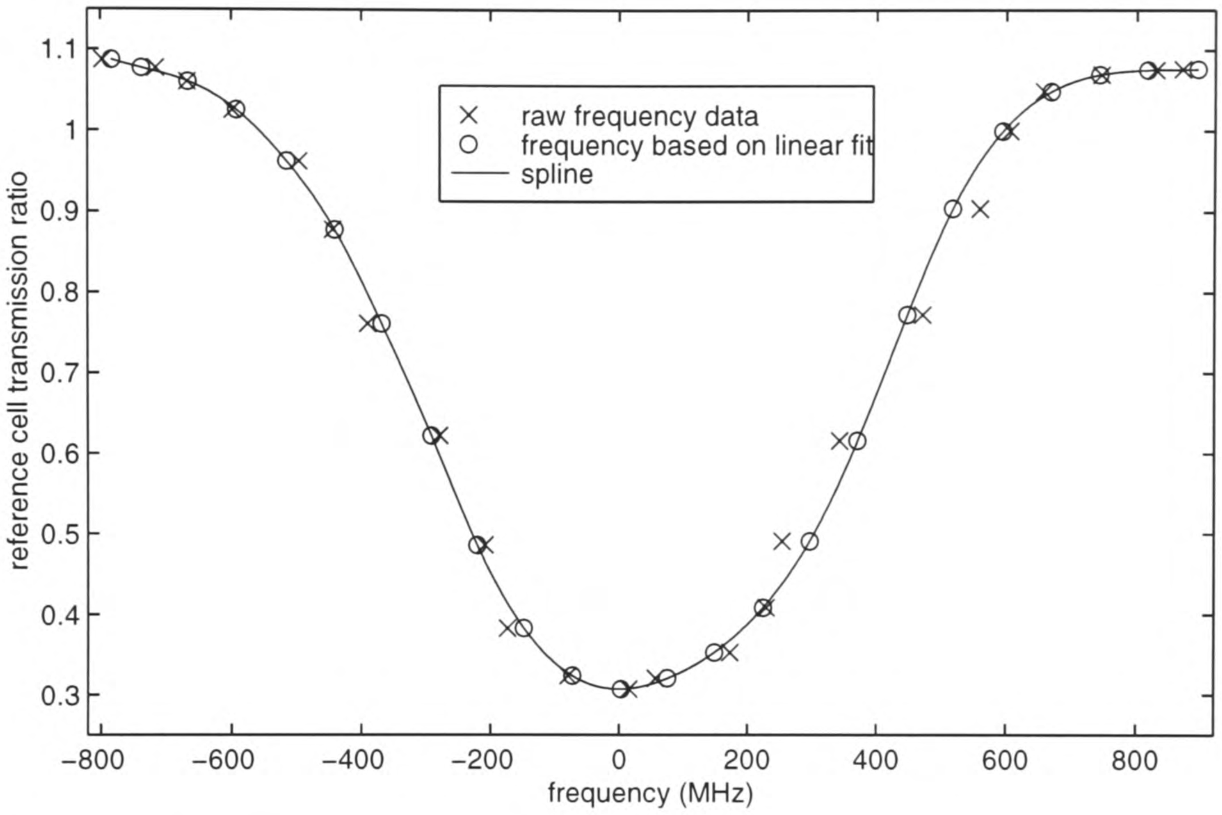


Figure 5.13 Calibrated transmission profile of the reference cell used in time-resolved DGV apparatus. Also shown are the raw data points based on measurements of frequency using the Fabry-Perot interferometer, and the points on which the calibration is directly based, which were derived by assuming a linear dependence of frequency on seed laser control voltage.

For each imaging cell, a profile of transmission ratio as a function of frequency was then generated by finding a correlation between the transmission ratio of the imaging cell and that of the reference cell. A typical correlation for a calibration in the time-integrated facility is shown in ***Figure 5.14***. It is closely modelled by a quadratic curve fit, as shown; correlations for the cells in the time-resolved facility, on the other hand, are almost exactly linear. Calibrated transmission profiles of the imaging cell were finally produced by applying the quadratic fit to all the points of the reference cell profile.

For completeness, calibrations typical of those used in this work are shown in ***Figure 5.15***. Profiles of the reference cell and imaging cells used in both DGV systems are shown. All calibrations were checked before important experiments, and renewed whenever the reference cell was rearranged, and on a few occasions when appreciable drift had occurred.

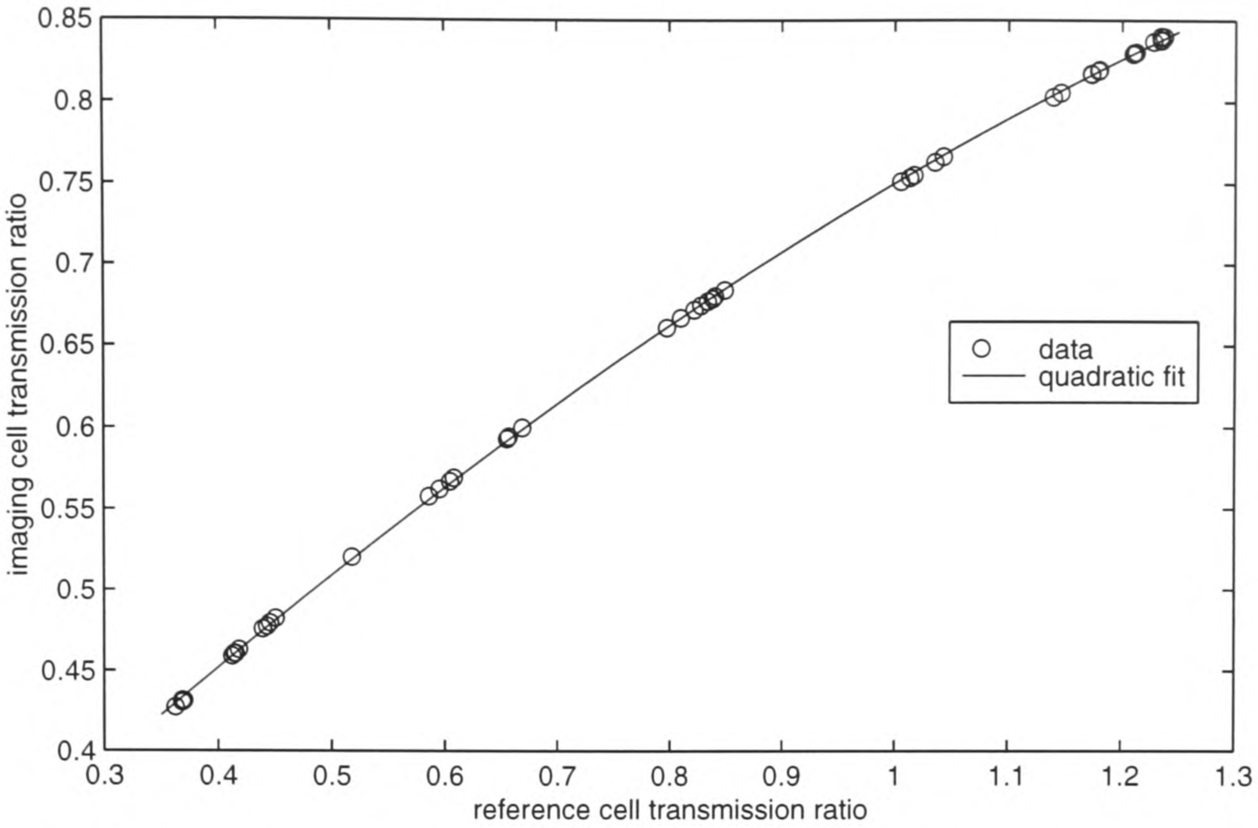


Figure 5.14 Correlation between reference and imaging cell transmission ratios measured during a calibration of the imaging cell in the time-integrated DGV system.

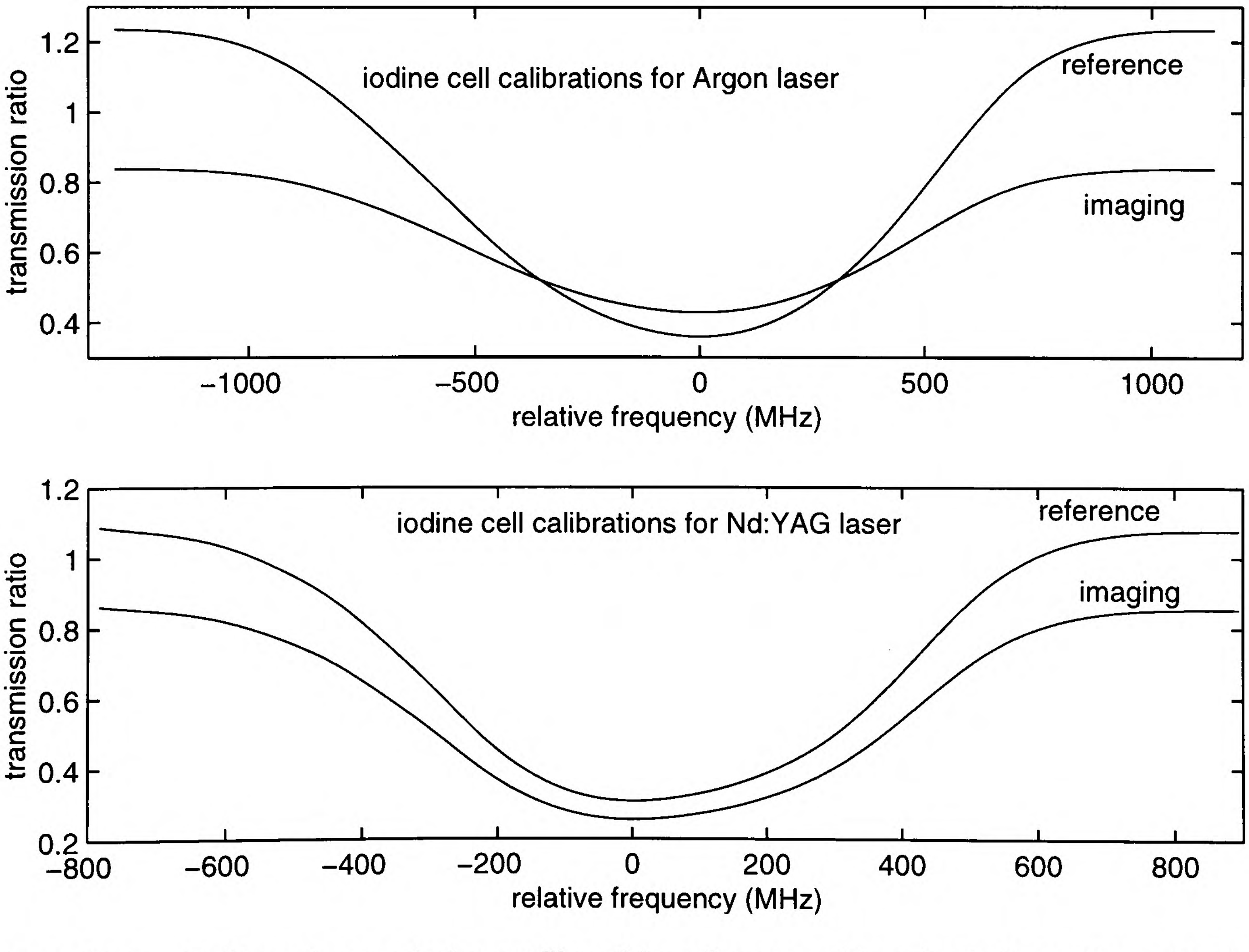


Figure 5.15 Calibrated transmission profiles of the reference and imaging iodine cells used with the time-integrated (*top*) and time-resolved (*bottom*) DGV systems. Frequencies are specified relative to the frequency of minimum transmission.

5.3 Uncertainty Analysis

Through consideration of the various hardware and software influences on the final measured velocity, an estimate of the uncertainty associated with every pixel in a DGV measurement of a velocity field can be made. The method of computing uncertainty is explained in this section. For three representative velocity measurement results, calculated estimates of the velocity uncertainty field are presented in Chapter 6.

The uncertainty analysis method described below is essentially the method established by Thorpe *et al.* [37] and Thorpe [41]. Its predictions have been shown to be realistic in measurements of two known velocity fields, a spinning disc and a turbulent free jet. Two minor changes to the method have been made in the course of present research — a model for the error occurring in calibration of the iodine cell for time-resolved measurements has been added, and a slight modification has been made to the estimation of misalignment error.

The conventional mathematical basis of general uncertainty analysis, described, for example, by Bevington [51], is followed. Quantitative statements of uncertainty in a measurement represent the root mean square (r.m.s.), over many repetitions of an experiment, of the deviation of the measurement result from the true value of the measured parameter. Uncertainty will be denoted by ε, subscripted with the name of the quantity to which it pertains. Additional subscripts will be used when a more exact specification is necessary. Uncertainties in derived quantities which are functions of measured parameters follow the rules listed below, in ***Eqs. 5.10*** to ***5.12***.

$$C = A \pm B \Rightarrow \varepsilon_C^2 = \varepsilon_A^2 + \varepsilon_B^2 \tag{5.10}$$

$$C = \frac{A}{B} \Rightarrow \left(\frac{\varepsilon_C}{C}\right)^2 = \left(\frac{\varepsilon_A}{A}\right)^2 + \left(\frac{\varepsilon_B}{B}\right)^2 \tag{5.11}$$

$$C = AB \Rightarrow \left(\frac{\varepsilon_C}{C}\right)^2 = \left(\frac{\varepsilon_A}{A}\right)^2 + \left(\frac{\varepsilon_B}{B}\right)^2 \tag{5.12}$$

Thorpe [41] identified five sources of error in DGV measurements:

- iodine cell calibration;
- determination of the laser emission frequency;
- instability of the iodine cell profile;

- noise in the CCD measurement of light intensity;
- image misalignment.

The first three contributions may conveniently be grouped together as an overall error in frequency shift, associated with the iodine cells. The fourth originates as an error in image intensity; the fifth, it will be shown, can also be described in terms of an error in intensity. In the sections which follow, the analysis of each source will be outlined, and the method for combining their contributions into an overall velocity uncertainty will be given.

5.3.1 Uncertainty Due to the Iodine Cells in Time-Integrated DGV

As described in §5.2.7, the reference iodine cell for time-integrated measurements is calibrated by measuring cell transmission as the emission frequency of the Argon ion laser is swept back and forth across the iodine absorption bandwidth several times. The calibrated value of transmission ratio *T*, at each of the emission frequencies visited, is the mean of all measurements of transmission ratio at that mode. The uncertainty in transmission ratio, ε_T, is taken as the standard deviation of all those measurements. This is converted to uncertainty in frequency by ***Eq. 5.13***.

$$\varepsilon_\nu \frac{dT}{d\nu} = \varepsilon_T \tag{5.13}$$

$\frac{dT}{d\nu}$ is calculated as the local slope of the calibrated transmission profile. In reality, of course, the laser frequency drifts continuously between mode hops to some extent, because it depends on cavity length, which varies with temperature. Consequently, successive measurements of transmission ratio which are assumed to represent the same mode do not, in fact, take place at precisely the same frequency. This is the major cause of uncertainty in the calibration. The uncertainty ε_ν is calculated for each of the calibration points, and for practical purposes, a curve is fitted so that the frequency uncertainty at any point on the cell profile, between modes, is available. This is the first of the five error sources listed at the start of §5.3. The frequency uncertainty distribution for the imaging cell calibration is precisely the same, since it refers to the reference cell as a standard. This is the second of the five error sources.

In previous studies, with calibrations based on two sweeps of laser frequency across the iodine cell bandwidth, the continuous drift of the laser frequency has been measured and incorporated in the uncertainty analysis. In the current work, however, drift which occurs between mode hops is accounted for implicitly, as it contributes to the variations between measurements of transmission ratio at nominally identical frequencies on successive sweeps.

Variations in iodine cell transmission, the third of the error sources, have previously been found to be due to ambient temperature variations [41]. They contribute an r.m.s. uncertainty of 1.5 MHz to frequency measurements.

5.3.2 Uncertainty Due to the Iodine Cells for Time-Resolved DGV

The iodine cells used for time-resolved DGV are calibrated by an entirely different method, described in §5.2.7, using a Fabry-Perot interferometer for measurement of light frequency. The main influence on the accuracy of this calibration technique is the accuracy with which ring radii are determined in digital images of the fringe patterns formed in the Fabry-Perot device. Accuracy is limited by pixel size in these images. If the maximum intensity on a ring is identified at pixel i, then the true intensity peak may be located, with uniform probability, anywhere between $i - 1/2$ and $i + 1/2$. The standard deviation of this distribution, which is interpreted as the r.m.s. error in radius of the ring, is 0.29 pixels. This value is used with ***Eq. 5.9***, which defines the relationship between ring radius and frequency, to determine uncertainties in each of the frequency values derived from each Fabry-Perot ring pattern. Averaging of the frequencies computed from several pairs of rings in each pattern reduces the uncertainty. The resulting frequency uncertainty is plotted as a function of frequency in ***Figure 5.16***. Also shown is the deviation of measured frequency from the frequencies predicted by a linear least-squares fit for frequency as a function of control voltage. The estimate of uncertainty is in agreement with the deviation over most of the frequency range, though the uncertainty based on finite image resolution becomes much greater as it grows rapidly in the high end of the frequency range. This growth occurs as the radius of the innermost ring in the interference pattern approaches zero.

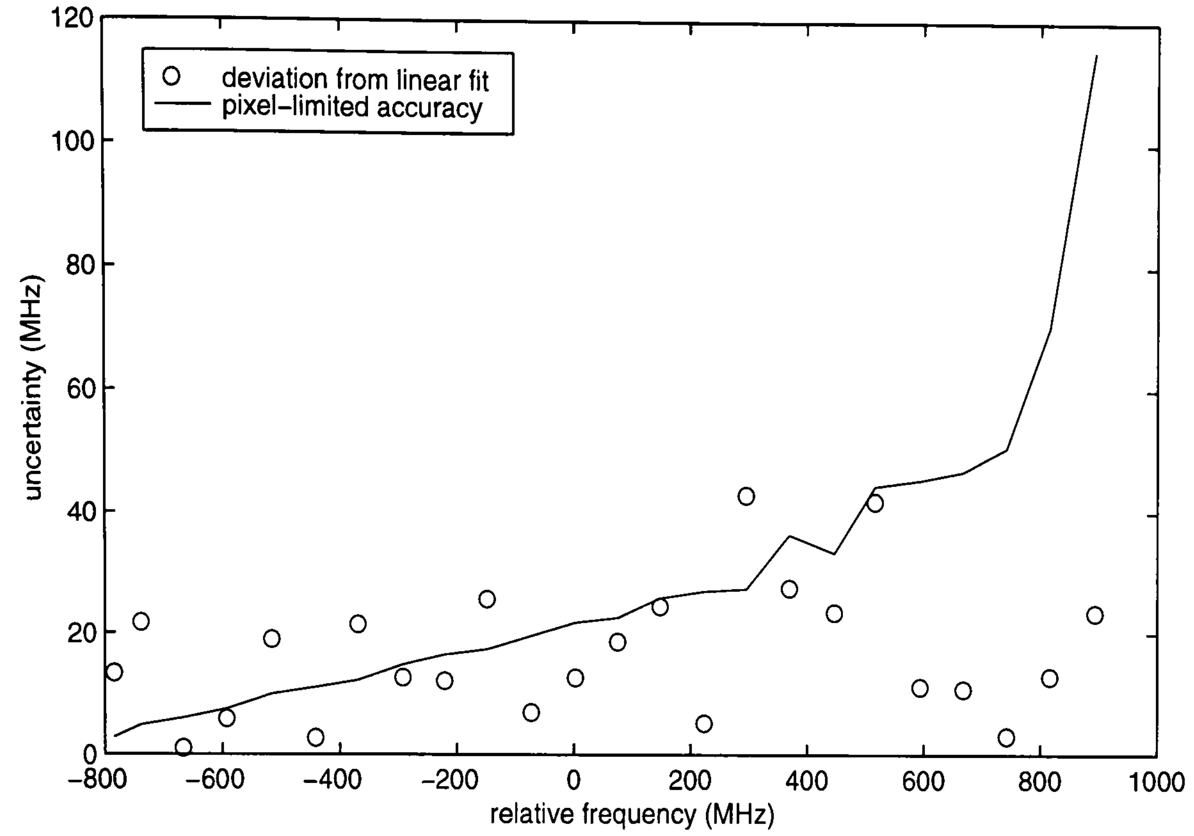


Figure 5.16 Theoretically estimated frequency uncertainty in an iodine cell calibration based on Fabry-Perot ring patterns, compared with the deviation of frequency from an assumed linear dependence on seed laser control voltage. The relative frequency zero is at minimum transmission.

The second major contribution to uncertainty in the iodine cell data for time-resolved measurements is long-term instability of the cell transmission profile. As mentioned previously, the reference cell is believed to be a starved cell, in which, at typical operating ambient temperature, all the iodine is in the vapour state. The vapour density, and consequently the cell transmission characteristics, should remain constant without any active control of vapour density. However, repeated calibrations showed some day-to-day variation of the cell's transmission profile, possibly because some iodine vapour tended to crystallise out in small quantities when room temperature was low. In order to incorporate this effect in an uncertainty analysis, the difference in transmission ratio between two sample calibrated profiles, taken at different times and known to be representative of extremes in cell behaviour, was taken as a conservative estimate of the standard deviation of cell transmission.

The contribution of iodine cell variation to the uncertainty in frequency shift measurement is presented in ***Figure 5.17***, along with the uncertainty associated with the finite resolution of the Fabry-Perot interferometer images, and the combined uncertainty due to these sources. On the low frequency side of the iodine cell transmission profile, the variability of the cell transmission dominates the overall error. Very near the transmission ratio minimum at the centre of the profile, uncertainty (in terms of frequency) tends to infinity. On the high-frequency side, which is utilised for the measurements described in this thesis, the contribution of iodine cell instability is minor and the uncertainty associated with analysis of the Fabry-Perot interference patterns dominates over a broad region where overall error ranges from 25 to 40 MHz. At higher frequencies, the effect of cell instability overrides.

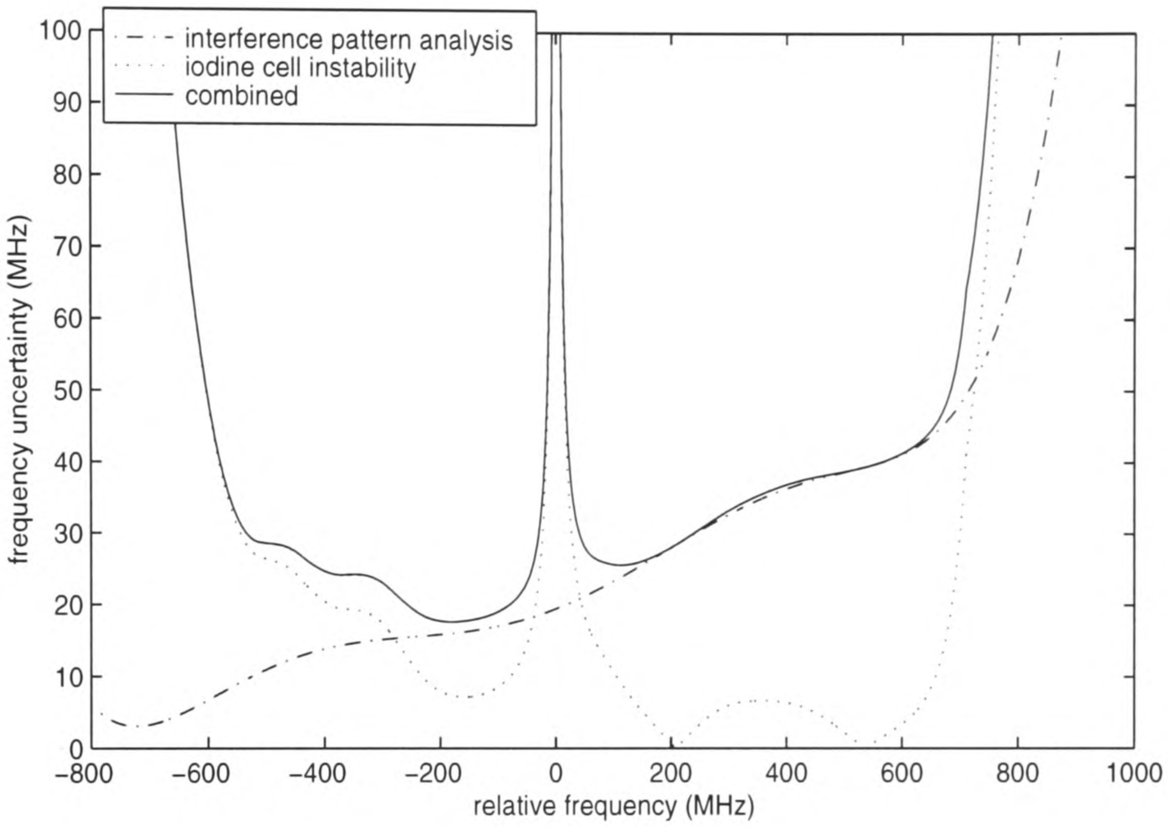


Figure 5.17 A plot of the overall frequency uncertainty in calibration of the reference iodine cell for use in time-resolved measurements, and its two components, due to finite resolution in images of the Fabry-Perot ring pattern and instability of the transmission characteristic.

5.3.3 Error due to CCD Noise

For each pixel in the CCD array, the camera electronics deliver a digitised value of intensity which is susceptible to three sources of error. Of these, *dark current*, which is strongly temperature-dependent, is negligible in the cameras used in the Oxford system because the temperature of the CCD chip is maintained at 200 K. *Read-out noise*, which is introduced in the process of converting pixel charge to voltage signals, is also insignificant, by virtue of the slow read-out rate of these cameras. The major error source is *shot noise*, which is a fundamental limit on the accuracy to which the number of electrons in a pixel can be counted. For a pixel with a capacity of N_{fw} electrons, the uncertainty in the digital intensity reading I is given by:

$$\varepsilon_I = \sqrt{\frac{I I_{max}}{N_{fw}}} \tag{5.14}$$

where I_{max} is the maximum intensity reading, corresponding to a pixel electron count of N_{fw}. Given values of 640,000 and 32,767 for N_{fw} and I_{max} respectively, for the cameras used in the Oxford DGV system, it is straightforward to calculate shot noise for every pixel in a given image.

5.3.4 Error due to Misalignment

The processes of alignment calibration and alignment were outlined in §5.2.6. If these processes function with perfect accuracy, pixels at the same position within the warped reference and discriminated images represent the same spatial region in the measurement plane. However, the accuracy of alignment is finite, and effectively, each pixel is displaced from its correct aligned position. The result of this displacement is that an incorrect intensity value will be measured for each pixel, if there are local intensity variations. If the displacement of a particular pixel due to misalignment is characterised in terms of distance in x and y as ε_x and ε_y respectively, then the resulting error in intensity is given by ***Eq. 5.15***:

$$\varepsilon_I = \varepsilon_x \frac{\partial I}{\partial x} + \varepsilon_y \frac{\partial I}{\partial y} \; . \qquad (5.15)$$

In practice, the partial derivatives in the above equation can be approximated by central difference expressions. The method for evaluation of ε_x and ε_y is less obvious. Thorpe [41] assumed that misalignment everywhere in the image was equal to 0.05 pixels, a figure derived from empirical tests of the imaging and image processing techniques at a fundamental level. Manners *et al.* [42] took misalignment error as the r.m.s. deviation between the locations of dot centroids in the warped images and the vertices of a regular orthogonal grid at which the dot centroids would appear under a perfectly accurate warp. However, the latter method does not measure the *relative* misalignment of the reference and discriminated images. In the present work, misalignment is assessed by a direct empirical method. Profiles of image intensity along lines through dot centres, as represented in ***Figure 5.7***, are analysed in both discriminated and reference warped images of the dots alignment cards. The dot intensity profiles appear as a roughly uniform plateau of approximately unit normalised intensity, separated from the low background intensity by an edge smeared out over some five pixels. (The fact that the edge is not resolved within one pixel is due to optical aberrations). The offset between the two images of each edge of each dot is measured where the normalised image intensity is 0.5, approximately half of its maximum value. For every dot within the region of interest in the image, this yields two data points for misalignment in the x direction, and two for the y direction. The errors ε_x and ε_y in ***Eq. 5.15*** are then supplied by a global root mean square of the offsets in each direction.

5.3.5 Overall Uncertainty

The preceding sections described how three contributions to the measurement uncertainty are evaluated. Of these, one, the error associated with the iodine cells, is expressed in terms of frequency. The other two, shot noise and misalignment, are expressed in terms of intensity in the warped discriminated and reference images. It is required to combine these into an estimate of overall uncertainty in velocity.

First, the uncertainties in intensity must be converted to uncertainties in frequency. Using the usual rules for combining uncertainties, the uncertainty in measured transmission ratio T can be shown to be given by:

$$\varepsilon_T = \sqrt{\left(\frac{\varepsilon_{I,d}}{I_r}\right)^2 + \left(\frac{I_d \varepsilon_{I,r}}{I_r^2}\right)^2} \quad , \qquad (5.16)$$

where the subscripts r and d refer to the reference and discriminated images, respectively. This is translated into an error in frequency according to:

$$\varepsilon_\nu = \frac{\varepsilon_T}{\left(\frac{dT}{d\nu}\right)} \quad . \qquad (5.17)$$

Again, $\left(\frac{dT}{d\nu}\right)$ is computed from the calibrated transmission profile of the cell in question. It remains to combine the errors in frequency and convert them into errors in the measured velocity component. This operation is summarised in ***Eq. 5.18***.

$$\varepsilon_V = \frac{\lambda_0}{|\hat{o} - \hat{i}|} \sqrt{\varepsilon^2_{\nu,\,iodine} + \varepsilon^2_{\nu,\,misalignment} + \varepsilon^2_{\nu,\,CCD}} \qquad (5.18)$$

Given the complete data set associated with a DGV experiment — raw CCD frames, warped images, iodine cell calibrations and velocity field — the methods outlined above may be used to estimate the uncertainty in measured velocity at each pixel. Uncertainty field estimates calculated in this way for three representative velocity measurements are shown alongside the main presentation of DGV results in Chapter 6.

5.4 Application to Transdermal Powder Delivery Flow-fields

In the preceding sections of this chapter, the DGV method has been discussed in general, without reference to any particular flow field. Some aspects of the experimental method are, of course, specific to the flow being studied. The primary purpose of this section and the next is to complete the description of the DGV method, in the context of the current work, with an account of the detailed configuration for application to the drug delivery flow fields of interest. Areas which are regarded as routine or well-understood elements of the experimental method are dealt with in this section, §5.4. They include the geometric configuration of the illumination, the flow-field and the imaging system, and topics specific to the application of time-integrated and time-resolved measurements. This section begins with a note on the significance of the particles from a laser anemometry point of view. Some more complex issues, relating to the integrity and interpretation of the measurements, are dealt with in §5.5.

5.4.1 Seeding Particles or Model Drug Particles?

This is an appropriate point at which to mention a simple feature in which this work differs from the majority of applications of DGV and other laser anemometry methods. Usually, the aim is to measure the velocity field of a fluid. Because the fluid itself is not detectable in many optical techniques, foreign *seeding* or *tracer* particles are added to the flow. They must scatter radiation adequately, but should be small enough to follow the fluid velocity field with negligible lag.

In this study of flow fields in drug delivery, however, the velocity field of primary interest is that of the solid drug particles. The particles loaded into drug delivery devices for velocity measurements, in the majority of experiments documented here, should not be regarded as seeding particles. They are either drug particles, or particles chosen for their similarity to the drug, and are an integral part of the system being investigated. Measurements of the velocity of silica powder are an exception to this statement. It was hoped that the silica particles would follow the gas flow closely enough to give an indication of gas velocity.

Consequently, the tasks of selecting, manufacturing and delivering appropriate seeding particles do not arise in this work. A drawback associated with the role of the particles in these measurements is that the light-scattering properties of the flow-field cannot be manipulated by altering the seeding concentration or particle type at will. In any attempt to alter the payload for this purpose, the possible impact on the overall two-phase flow-field of interest must be considered.

5.4.2 Experiment Geometry

The layout of the imaging system and the light sheet in relation to the flow field is an important feature of the experiment design. It determines the measurement plane and the visible region of the measurement plane, as well as the direction of the measured velocity component. It has a bearing on the choice of laser emission frequency, which is also discussed in this section.

The subject of the DGV experiments in this work is the axial component of particle delivery velocity. This component has the greatest importance to device function, since it will determine the drug particles' penetration of the target surface. The chosen arrangement of the measurement plane, illumination vector and observation vector is depicted in ***Figure 5.18***. The measurement plane (defined by the plane of the light sheet) is vertical, containing the nozzle centreline and bisecting the nozzle. The illumination vector is directed upstream, parallel to the centreline, and the observation vector is in a horizontal plane. The measured velocity component lies in the horizontal plane defined by those two vectors. Radial velocity in the central plane is vertical, and therefore makes no contribution to horizontal velocity components. Swirl velocity in this measurement plane would be entirely horizontal, but it is expected to be negligible. Therefore, in this measurement configuration, only axial velocity contributes to the measured Doppler shift.

Another attraction of this setup is that it enables characterisation of the entire three-dimensional flow field. Assuming that the flow is axisymmetric, then only axial and radial variations exist, and they are fully described by a measurement on any plane which contains the centreline.

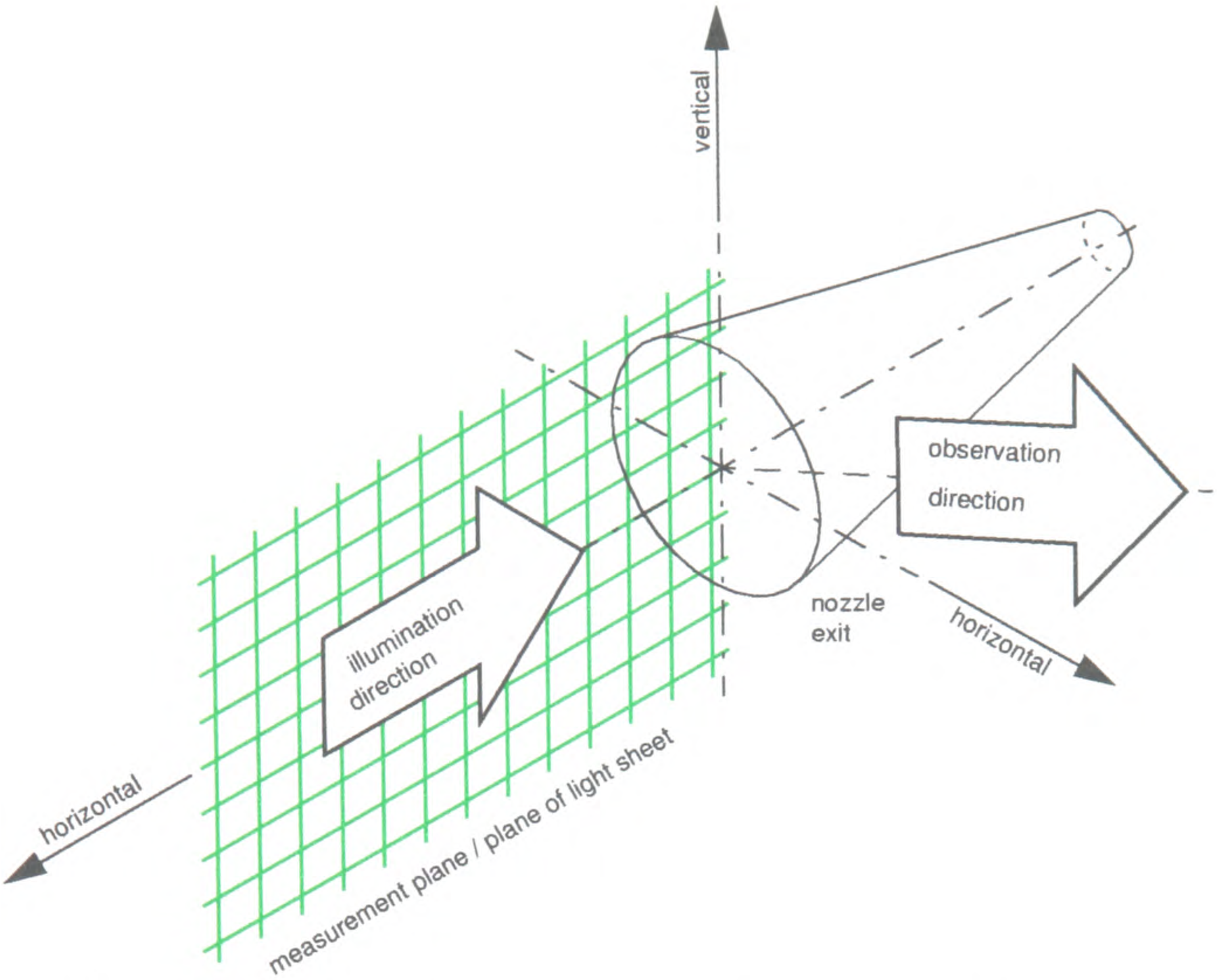


Figure 5.18 The arrangement of the light sheet and drug delivery device in DGV experiments.

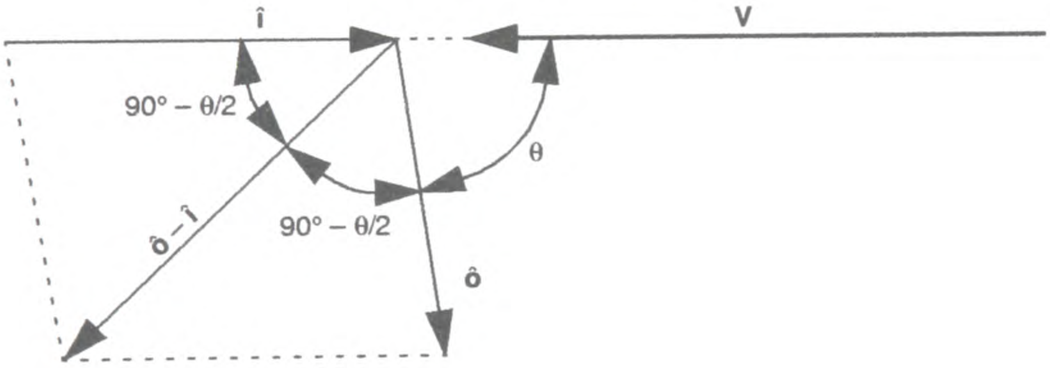


Figure 5.19 A diagram of the vectors involved in the Doppler shift equation, as they occur in a generalised geometry for measurement of axial velocity in the present drug delivery flow fields. All the vectors shown are in a horizontal plane.

The sensitivity of the frequency shift to axial velocity is determined by the arrangement of the illumination, observation and velocity vectors, which appear in the Doppler shift equation. A diagram in ***Figure 5.19*** shows a plan view of the vectors, as they occur in a generalised form of this measurement geometry. The Doppler shift equation is repeated here:

$$\Delta\nu = \nu_0 \frac{\hat{o} - \hat{i}}{c} \cdot V \quad . \qquad (5.19)$$

For this particular geometry, with the assumption of zero swirl velocity, the dot product is evaluated as follows:

$$\Delta\nu = \frac{\nu_0}{c}(\hat{o} \cdot V - \hat{i} \cdot V) = \frac{\nu_0}{c}(-V_x\cos\theta + V_x) \quad , \tag{5.20}$$

where the subscript x denotes the axial velocity component. Rearrangement gives the form of the Doppler shift equation which is normally used in the present work:

$$V_x = \frac{\Delta\nu\lambda_0}{1 - \cos\theta} \quad . \tag{5.21}$$

Eq. 5.21 demonstrates how the angle θ controls the sensitivity of frequency shift to velocity. Since the measurable frequency shift is limited by the width of the iodine cell profile, θ effectively determines the maximum measurable velocity. The choice of θ has other implications. According to the Mie theory [38] (which describes wave scattering by spheres with a diameter greater than the wavelength of radiation), the intensity of scattered light should tend to increase (although not monotonically) as θ decreases, and this has been confirmed experimentally in the present work. At values of θ very far from 90°, the depth of field in the image becomes large, and it may become difficult to generate an image which is sharp across the field. It has also been observed experimentally that significant optical distortion occurs when θ is low and the imaging system is placed close to the measurement plane. In summary, the maximum measurable velocity and the intensity of available light increase as θ decreases, but image quality may deteriorate. In practice, particle velocities in the flow fields created by drug delivery devices have been found to be high enough to make the maximum measurable frequency shift the critical criterion in the selection of θ.

The choice of laser emission frequency ν_0 is a related issue. The chosen ν_0 must allow enough width on the iodine cell transmission profile to accommodate the largest Doppler shifts which occur, while transmission ratios in the flat regions of the profile, near the centre and the edges, must be avoided. In the flat regions, sensitivity of transmission ratio to frequency is low or zero, and moderate errors in the measurement of transmission ratio are magnified into large frequency errors. The large frequency shifts encountered in the current work require a choice of ν_0 which will maximise the available width of the transmission profile, so that the largest possible viewing angle, θ, can be chosen.

It is apparent from ***Figure 5.15*** that the low-frequency side of the iodine cell profiles used in time-integrated work offers a slightly wider useful frequency range than the high frequency side. On a rough subjective test, a shift of perhaps 850 MHz can be measured on this absorption profile without encountering very low sensitivity. The highest velocities measured with the Mach 2.5 nozzle approached 1200 m/s; a viewing angle of approximately 46° was used in the final experiments, translating the 850 MHz range into about 1400 m/s. Since positive Doppler shifts were most convenient to measure in the layout of the laboratory, emission frequencies in the low frequency end of the profile, at transmission ratios between 1.15 and 1.2, were chosen. For time-integrated measurements of the flow from the conical nozzle, with maximum velocities of under 600 m/s, a viewing angle of 71° was used, implying a clear measurable velocity of 650 m/s.

Transmission profiles of cells used for time-resolved work are also shown in ***Figure 5.15***. In terms of available frequency range, there is no clear advantage to using either side of the profile. Again, the arrangement of the laboratory favours positive Doppler shift. The selection of an emission frequency on the high frequency side of the profile, at a low transmission ratio, ensures that the direction of frequency shift tends to raise the intensity of the discriminated image, helping to reduce shot noise contributions to velocity error. Reasonable sensitivity is available over a range of 500 MHz. Velocities approaching 800 m/s were measured in this configuration; a viewing angle of 45° was chosen, permitting measurement of up to 910 m/s.

The light sheets formed from the beams of the Argon ion and Nd:YAG lasers were approximately 1 mm and 0.5 mm thick, respectively. For time-integrated measurements of particle velocity from the contoured nozzle (exit diameter 2.5 mm), the laser light was formed into a sheet approximately 16 mm wide (in the vertical direction). With the conical nozzle, as discussed in the next section, it was necessary to minimise the light sheet width in order to boost illumination intensity. The width chosen was typically 25 mm, or approximately two exit diameters. Nd:YAG light sheets used for time-resolved measurements were typically 50 mm wide. Light sheet thickness was always set to the lowest possible value, which was limited by the divergence of the laser beam.

5.4.3 Illumination Intensity

The intensity of the raw flow-field images used for DGV can be controlled by altering the laser output power (or pulse energy) at source, or by varying the aperture of the imaging system. This control of image intensity is essential. If the image is too bright, local saturation of the CCD occurs, and velocity cannot be computed reliably. If the image is too dim, of course, shot noise obliterates the meaningful data. Ideally, the maximum intensity in the region of interest in the raw flow-field images should be just below CCD saturation level.

For the time-integrated measurements, maximum image intensity was limited by the power available from the laser. The aperture of the imaging system was set at the largest practicable value, which was restricted by imaging issues discussed in §5.5.2. The laser was always run at the maximum power available in single-mode operation, and even so, images with maximum intensity of half the CCD saturation level were rarely acquired. By adjusting the illumination optics, the width (or extent in the plane of the sheet, normal to the propagation direction) of the light sheet was kept as low as practically possible while covering the entire jet diameter. In this way, the available power was concentrated in the region of interest.

The larger exit area of the conical nozzle compounded the difficulties associated with low light levels in time-integrated velocimetry. Because any given payload is distributed over a larger area at the exit of a larger nozzle, the particle concentration is lower, and less light is scattered (and gathered for imaging) per unit area of the measurement plane. A light sheet of approximately two nozzle exit diameters in width was used. In an attempt to boost scattered light intensity from the dim regions of low particle concentration on the periphery of the jet, the light sheet was shifted upwards, in its own plane, so that the centreline (and line of maximum intensity) of the light sheet was about 3 mm above the nozzle centreline. This approach leads to loss of data in the lower half of the jet in cases of particularly low scattering levels, giving a false impression of asymmetry in some measured velocity fields at a first glance. However, it ensures that at least one half of the jet diameter is adequately illuminated.

In time-resolved work, pulse energy from the Nd:YAG laser was always more than enough to saturate the CCD camera. The aperture of the imaging system was held at an intermediate value, f/8, and laser output in the green 532 nm beam of interest was attenuated as required by detuning the harmonic generator.

5.4.4 Application of Time-Integrated and Time-Resolved DGV

Both the time-integrated and time-resolved DGV techniques have been applied in this research. In time-integrated DGV, the measurement plane was illuminated continuously in time by the Argon ion laser. The CCD array was exposed for 1 s (typically), a period many times longer than the lifetime of the drug delivery flow-field. During this period, the drug delivery device was actuated manually, and the light scattered from the flow-field was integrated on the CCD array over the duration of particle flow (several milliseconds). Each image recorded in this way, and the resulting velocity map, constituted a representation of the whole history of the flow-field.

In time-resolved DGV experiments, the pulsed Nd:YAG laser was used to illuminate the flow-field for approximately 10 ns, while the camera was again exposed for a period of the order of 1 s. Data yielded by these measurements represented the (practically) instantaneous state of the flow-field at the moment of illumination. The technical arrangement which enabled delivery of the laser pulse at a chosen instant in the flow history is described in the next section.

5.4.5 Illumination Timing for Time-Resolved DGV

The Nd:YAG laser used for time-resolved measurements delivers light pulses of approximately 10 ns duration at a rate of one every 100 ms. The lifetime of particle delivery in the short-duration flow field of the drug delivery device is typically less than 5 ms. Clearly, automatic coordination of the laser and the actuation of the drug delivery device is required. For a successful measurement, a pulse must be delivered within the particle delivery period of less than 5 ms. Furthermore, in order to map temporal variations in the flow field with successive measurements, it is desirable to control the timing of the pulse within the flow field history to an accuracy of tens of microseconds or less.

One approach is to trigger the laser in response to signals from an instrument, such as a pressure transducer, monitoring the flow field. Although a pulse can be demanded from the laser at any time by an external signal, large deviations from the 10 Hz nominal pulse repetition rate may cause a partial or total loss of pulse energy, or failure to obtain single-mode radiation. Alternatively, the drug delivery device can be fired automatically, using a solenoid and some triggering electronics. Attempts to initiate the flow field electromechanically, on a trigger signal from the fixed-rate laser pulsing, failed because of temporal variability in the actuation system. Mean delays of 38 ms or more, from a trigger signal to the beginning of flow in the device, could be obtained, with a standard deviation of the order of 3 ms. Even if this level of repeatability allowed images to be obtained at some stage in the particle delivery lifetime, on a reasonable fraction of attempts, it would certainly not give any reliability in the choice of measurement timing within the flow history.

A hybrid of these two strategies has been employed successfully. The gross timing of the drug delivery actuation with respect to the laser's 10 Hz cycle is set by triggering a solenoid in response to a signal issued by the laser with the most recent laser pulse; fine tuning of the instant of measurement is then provided by requesting the next laser pulse a little early, on a trigger signal from a pressure transducer. Much of the electronic hardware involved was designed and assembled by Dr. Steven Thorpe; the author contributed the solenoid trigger circuit and some modifications to the timing of the command signals. ***Figure 5.20*** contains a diagram showing the flow of signals through the apparatus. Examples of some of the signals themselves are shown in ***Figure 5.21***. In normal operation, a custom-built fixed-rate repetitive pulse generator causes the laser to fire every 100 ms by issuing a signal to its lamp trigger input. When the user closes a "fire" switch, one input to an AND gate circuit is activated. When the next laser pulse raises the other input to the AND gate less than 100 ms later, the gate's output goes high and triggers a transistor switch circuit which supplies power to the solenoid. The delay from the rise of this signal to the commencement of flow can be controlled by selecting the voltage supplied to the solenoid. A mean delay of 85-95 ms is set, ensuring that flow begins well before the next regular laser pulse is due. A piezoelectric pressure transducer, manufactured by PCB, located in the rupture chamber, acts as a monitor of the unsteady flow. Its output is passed through a dedicated charge amplifier and an AD524 amplifier to a commercial pulse generator (Hewlett Packard 8013B). The AD524 acts as a buffer, and allows addition of an offset to the pressure transducer signal so that the fixed input voltage level required to trigger the Hewlett Packard (HP) pulse generator can be provided at any desired level of the pressure transducer signal. The amplifier was used to offset the pressure signal by -1.0 V, so that the HP pulse generator triggered when the pressure signal was approximately 1.9 V. This occurs late in the filling of the rupture chamber, typically 70 μs before bursting of the downstream membrane.

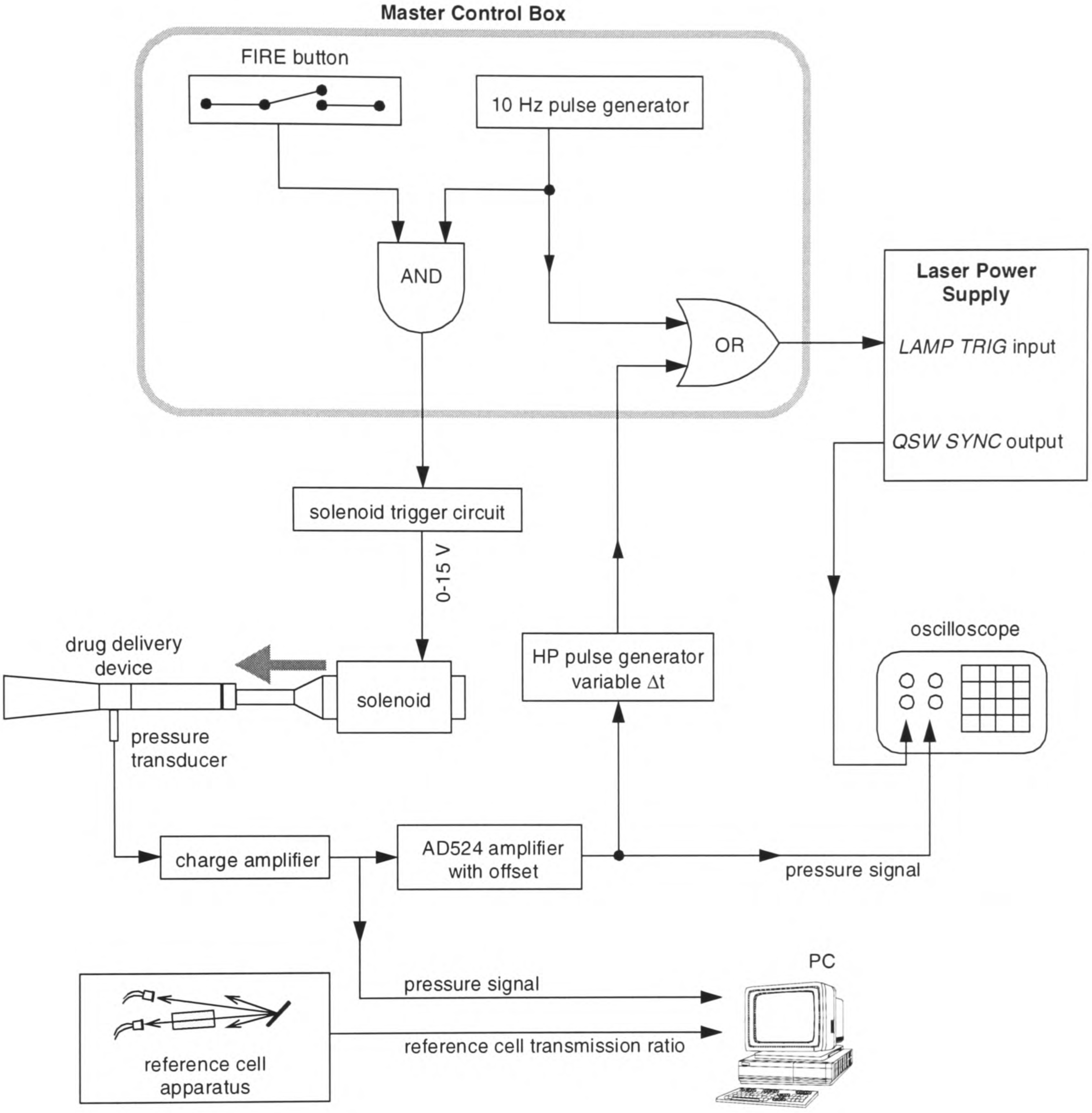


Figure 5.20 Schematic diagram of the flow of information in the system which governs synchronisation of the drug delivery flow with the pulsed laser.

On recognising a trigger from the amplified pressure transducer signal, the HP pulse generator delivers a well-defined rectangular pulse which is acceptable to the laser. Importantly, it introduces an adjustable delay ranging from 35 ns to a second. Its output is routed through an OR gate to the laser's lamp trigger input. The chosen width of this command pulse is relatively large, at 20 ms, to ensure that it persists long enough to override the next periodic 10 Hz pulse (which is routed to the laser through the other input of the OR gate). The laser pulse is then delivered after an adjustable delay of (typically) 140 to 180 μs, known as the Q-switch delay, during which neodymium atoms in the laser are pumped. Finally, the pressure transducer signal, and a laser output signal which is synchronous with the light pulse (precisely, it is synchronous with Q-switch opening), are stored on an oscilloscope. This record is used to measure the delay from bursting of the downstream membrane (assumed to occur when pressure reaches a second maximum) to illumination of the flow-field.

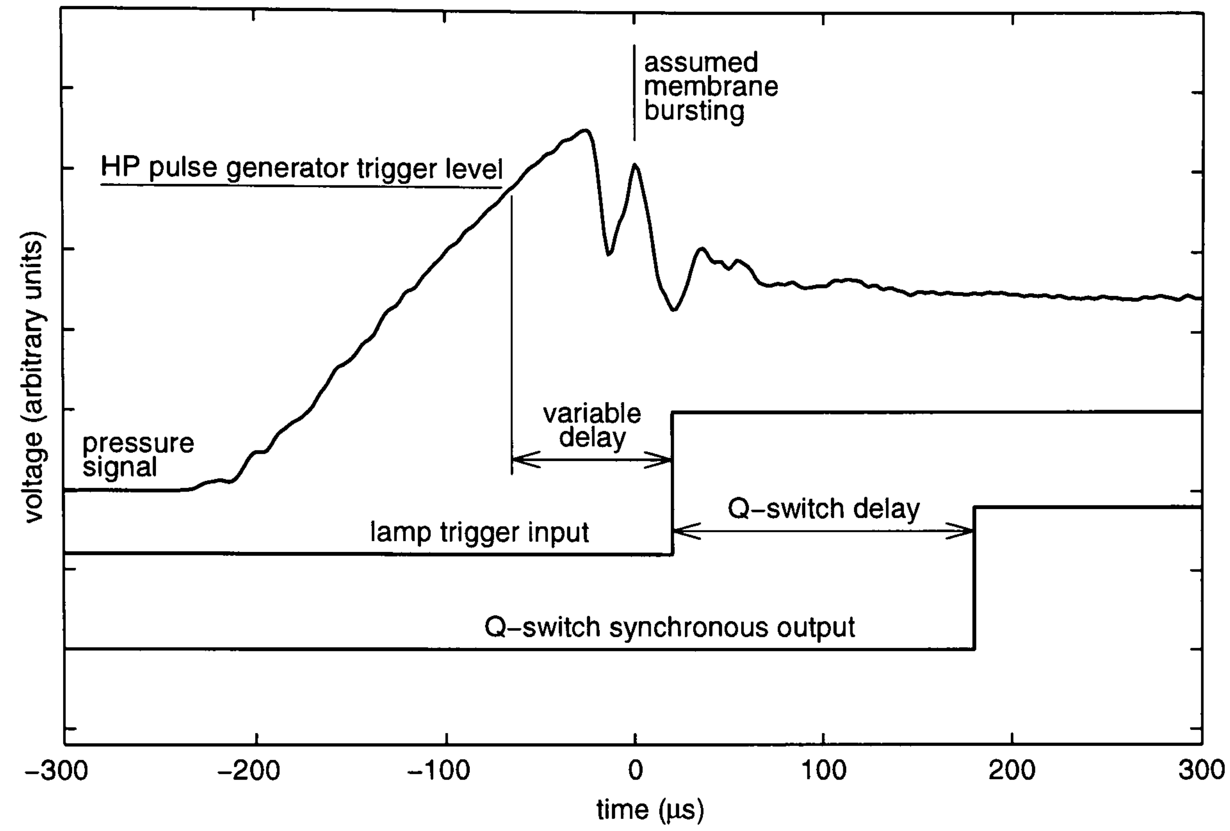


Figure 5.21 Examples of signals in the triggering system for time-resolved DGV.

The pressure transducer signal is also passed to the PC which records data from the reference iodine cell monitoring Nd:YAG emission frequency. This provides a means to identify the reference cell transmission ratio which was measured on the particular laser pulse delivered while the flow-field was active.

The net effect of this system is to bring one laser pulse forward in time by up to 15 ms in response to the flow-field. It ensures that a light pulse is always delivered in the lifetime of the flow and that its timing with respect to the evolution of gas flow field can be consistently chosen to within microseconds. The limit on the consistency of the timing is set by the variability of gas dynamic processes between the period just before the membranes burst, when the system triggers, and the initiation of flow in the nozzle. The variation in the delay from triggering of the HP pulse generator to detection of the starting shock has been measured as ±8 μs. This is added to by variations in the timing of the delivery of the first particles; optical particle detection experiments described in Chapter 4 suggest that this variation is of the order of ±3 μs for a 1 mg payload of 4.7 μm spheres.

5.5 Studies of Measurement Integrity

This section addresses various areas in which there were grounds for concern as to the validity of DGV measurements on flow fields from drug delivery devices. The first two subsections deal with spurious effects on the measurement of time-integrated velocities. Experiments were designed to demonstrate and quantify these effects, and solutions were devised. In §5.5.3, the feasibility of velocity measurements within the nozzle is discussed, with supporting experimental evidence. Finally, the physical meaning of the results of time-integrated measurements, in which an unsteady history is represented by a single velocity field, is discussed in §5.5.4.

5.5.1 Spurious Detected Light

In general, in DGV experiments, one of the purposes of a black image is to negate the effects of any light reflected by any objects, other than scattering particles in the flow-field, which appear in the image. It is assumed that the black-corrected image corresponds to the image which would be formed by the light scattered only by the particles whose velocity is of interest. Light from any other source is spurious, as makes a non-physical contribution to the velocity field computed from the raw image. It became apparent in the course of time-integrated DGV experiments with drug delivery devices that light was being scattered from the flow field by some material other than the drug particles. This section contains an account of experiments carried out to assess the extent of the problem, some theory for the source of the problem, and a description of methods devised to minimise its impact.

The problem was first detected when time-integrated frames of the flow from a device with an empty drug cassette were captured in the usual way. Two typical raw reference images, for the Mach 2.5 and conical nozzles, are shown in ***Figure 5.22***. These figures show that in the absence of powder, sufficient light is scattered from the measurement plane to form a strong image. If the mechanism by which light is apparently scattered from these particle-free jets also occurs when drug particles are present, the measurement of drug particle velocity will be based on a combination of the Doppler shift due to scattering by drug particles and the unknown Doppler shift associated with this unexpected scattering.

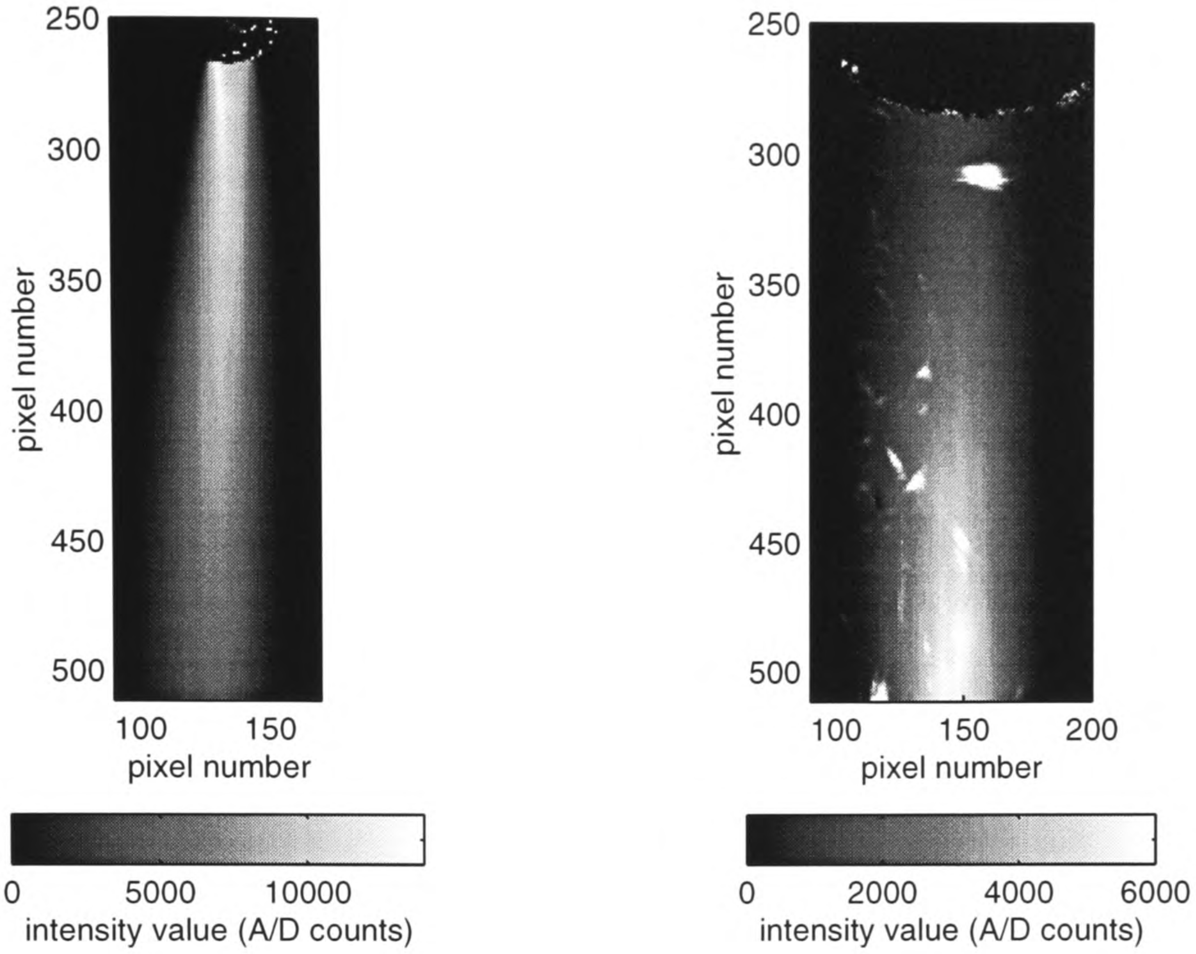


Figure 5.22 Black-corrected reference DGV images of the jets from devices fitted with Mach 2.5 (left) and conical (right) nozzles, both fired with empty drug cassettes. Image intensities are expressed as CCD analogue-to-digital (A/D) counts. Flow is from top to bottom.

Some conceivable mechanisms for scattering of light from the jet are: Rayleigh scattering (scattering by molecules); scattering by contaminants from within the device (dirt, lubricant or O-ring fragments, for example); scattering by membrane fragments; scattering by condensates in the gas, formed when constituents of the gas (helium and its impurities, with some air) undergo thermodynamic state changes in the flow field; and finally, scattering by atmospheric contaminants which become entrained in the jet.

Whenever feasible, these possibilities have been investigated experimentally. The contribution of debris in the cylinder was ruled out by thoroughly cleaning a cylinder and fitting new, unlubricated O-rings; no reduction was achieved in the intensity of the drug-free jet image. The contribution of membrane fragments was tested by firing a device which was assembled without membranes; again, no improvement was observed.

Evidence was found for the possibility that light scattering in the absence of drug is due to constituents of the gas in the jet. The make-up of gas delivered by the device varies throughout the duration of device actuation, as the following considerations show. Immediately after a shot, the cylinder contains gas, consisting mainly of helium, at atmospheric pressure. It is reasonable to assume that while the device is disassembled for refilling and reloading, the nozzle and rupture chamber are thoroughly purged of helium and filled with air. The residual gas in the cylinder may be infiltrated, or replaced entirely, by atmospheric air, depending on the time for which the valve is left open before the next fill. On filling, this residual air/helium mixture is compressed and diluted in helium. On firing the device, the valve opens and air in the rupture chamber is mixed with the helium-air mixture flowing from the cylinder. The membranes burst, and the first gas species to flow from the nozzle exit is air, behind the starting shock wave; this is followed by the mix of gases which filled the rupture chamber before bursting; finally, the helium-rich mixture stored in the cylinder is delivered in the quasi-steady nozzle flow.

Impurities in the helium (BOC grade A) are specified by its manufacturer as less than 0.004% by volume. They are therefore negligible in comparison with the quantity of air added to the helium.

To test the hypothesis that this air was responsible for the unwanted scattering, a cylinder filling procedure was devised which was expected to reduce the proportion of air in the cylinder. First, a trickle of helium through the filling system was maintained while coupling it to the cylinder, with the aim of excluding atmospheric air. The cylinder valve was then opened, and helium was allowed to flow through the filling line and into the cylinder as usual. When the pressure in the cylinder reached the regulated supply pressure (the desired fill pressure), the cylinder valve was closed and the helium supply cut off. The cylinder was then immediately discharged into the filling system. The filling system was allowed to vent through a bleed valve until pressure fell to just above atmospheric pressure, at which point the valve was closed. The cylinder was then refilled and discharged in this way several times, without disconnecting it from the filling system. Each fill adds helium to the air/helium mixture in the cylinder, and each subsequent discharge leaves a small mass of a new mixture with a reduced air content. Repetition of this procedure, called *flushing*, should successively reduce the mass of air in the cylinder.

Images of the jet produced by devices without drug, filled using the procedure described above, are shown in ***Figure 5.23***. In each case, the fill procedure involved five flushes before the final fill. The images should be compared with the images in ***Figure 5.22***, which were captured under similar conditions without flushing. For the Mach 2.5 nozzle, the flushing procedure is apparently successful, since the maximum digitised intensity of the image of the drug-free jet is reduced from a value of approximately 13,000 to 1,000. This result is preserved in repeats of the experiment. For the conical nozzle, on the other hand, little or no improvement is achieved.

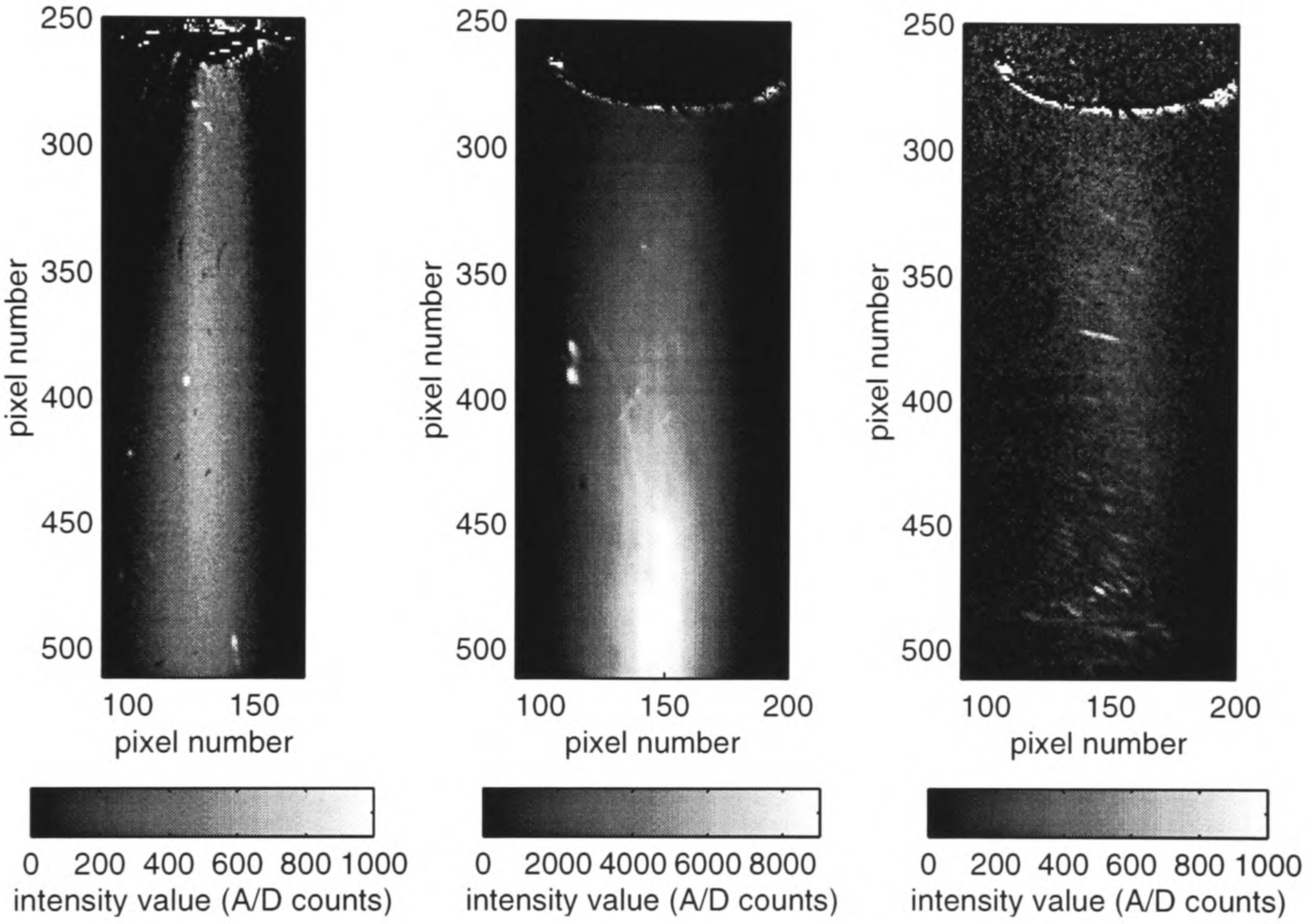


Figure 5.23 Black-corrected reference DGV images of jets from devices without any drug payload. *Left:* Mach 2.5 nozzle, flushed 5 times. *Middle:* conical nozzle, flushed 5 times. *Right:* conical nozzle, flushed 5 times and fired with air jet across exit. Note differing grey scales. Image intensities are expressed as CCD analogue-to-digital (A/D) counts. Flow is from top to bottom.

The success of cylinder flushing as an approach to reducing the scattering signature of the drug-free jets from the Mach 2.5 nozzle indicates that air in the cylinder is a major source of scattering particles. This theory is supported by the following gas dynamic analysis, which shows that water vapour in the air initially stored in the cylinder is likely to condense in the nozzle. If the cylinder is initially filled with air at atmospheric pressure and some known relative humidity, the mass fraction of any air constituent in the mixture which exists in the cylinder after filling can be calculated. For water vapour, this calculation requires knowledge of the relative humidity, and a hygrometric chart [52]. The variations of pressure and temperature in the cylinder and rupture chamber as the pressurised mixture blows down through the nozzle are calculated using the numerical model described in Chapter 7. (Time histories of the predicted pressure and temperature can be seen in Chapter 7). The static pressure and temperature in the rupture chamber are than taken as the total pressure and temperature in flow through the nozzle, and the gas state at all points in the nozzle is determined by assuming isentropic quasi-one-dimensional flow, with Mach number depending only on area. Assuming that the mass fraction of each constituent of the mixture remains constant, it is then possible to calculate the partial pressure of each mixture component at every point in the nozzle.

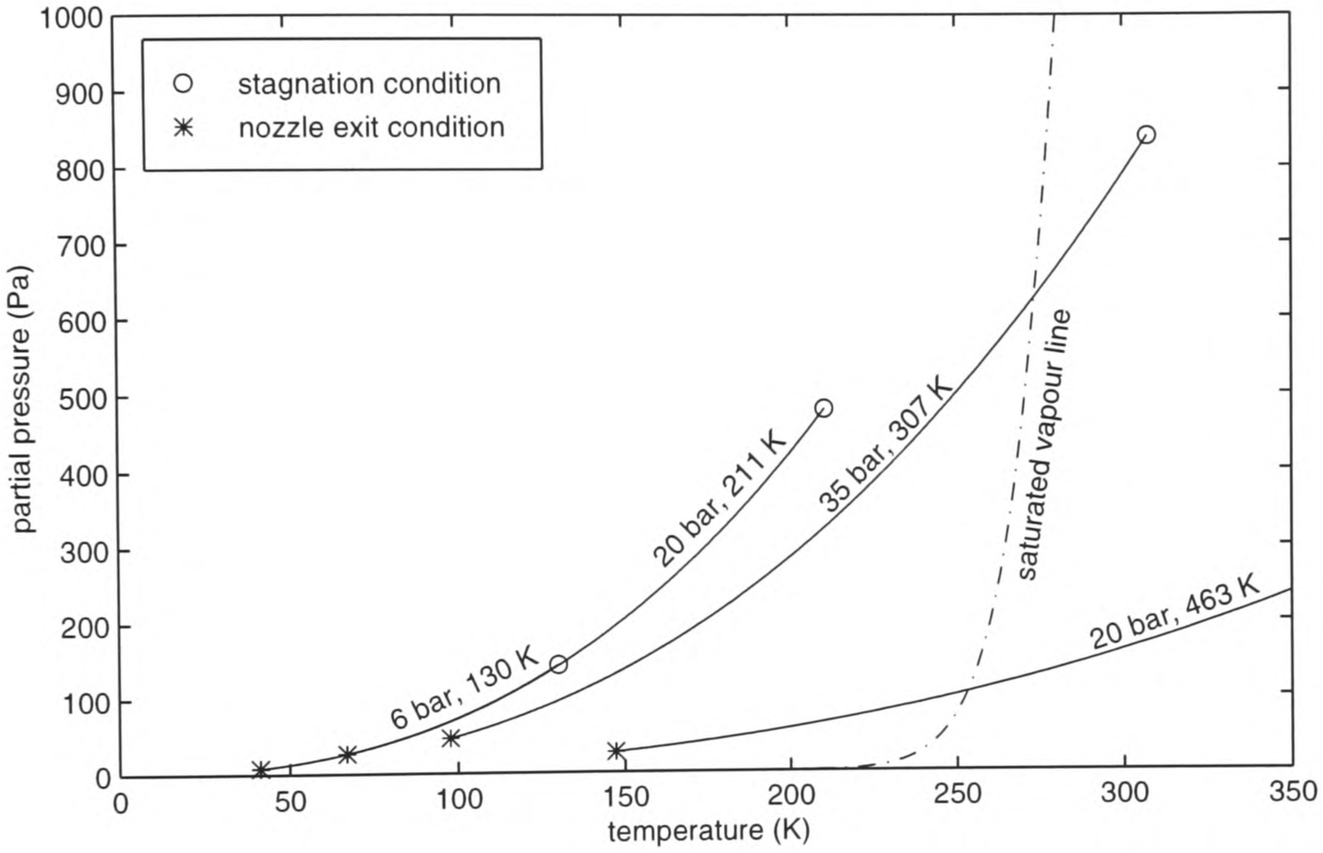


Figure 5.24 Calculated variation of temperature and partial pressure of water vapour in quasi-steady flow in the Mach 2.5 nozzle, for various instantaneous reservoir conditions, with the water vapour saturation line. Annotations show instantaneous total temperature and pressure.

Figure 5.24 shows the results of this analysis, applied for water vapour, assuming that the cylinder was full of air at 50% relative humidity before filling to 60 bar with helium. A selection of axial profiles of water partial pressure and temperature in the nozzle are shown, on a pressure-temperature graph. The vapour saturation line is also shown. For high-pressure, high-temperature states, which occur early in the flow, the axial nozzle profiles cross the saturation line, suggesting that water vapour does condense in the nozzle.

A similar analysis has been applied for nitrogen and oxygen, using saturation line data from Gray [53], and suggests that they could condense in the nozzle, but only in the most extreme conditions conceivable. The result for oxygen is illustrated in ***Figure 5.25***. The possibility that water vapour might condense behind the starting shock wave has also been examined, with the conclusion that the temperature rise associated with the shock overrides the pressure rise, pushing the state of the water vapour in the direction of superheating. Both experiments and analysis, therefore, support the theory that water vapour originating in the cylinder, condensing as it cools in flow through the nozzle, is responsible for light scattering from the drug-free jet.

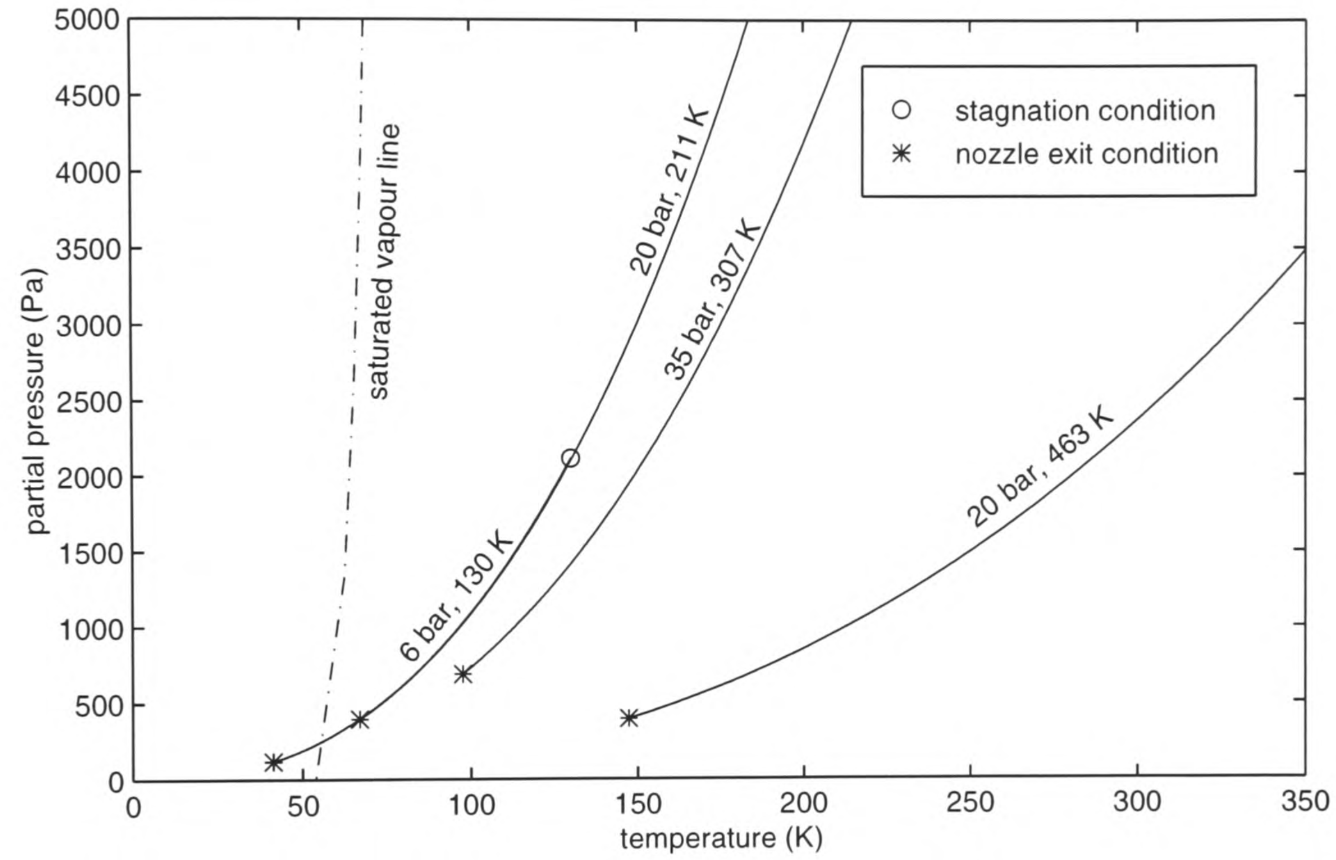


Figure 5.25 Calculated variation of temperature and partial pressure of oxygen in quasi-steady flow in the Mach 2.5 nozzle, for various instantaneous reservoir conditions, with the saturation line for oxygen. Annotations show instantaneous total temperature and pressure.

As suggested in ***Figure 5.23***, attempts to eliminate air from the cylinder did not succeed in reducing the optical signature of the drug-free jet from the conical nozzle. Another potential source of non-drug light-scattering material was investigated. Material contained in the air surrounding the device could become involved in the jet either through its presence in the nozzle before actuation (the volume of the conical nozzle is much greater than that of the Mach 2.5 nozzle), or by entrainment in the established jet. The first remedial action which attempted to undermine these hypothetical mechanisms was to control the atmosphere in which the device operated. This idea was implemented by firing the device in a box measuring 160 × 160 × 400 mm, oriented with its long edges parallel to the nozzle axis. The drug delivery device protruded by about 70 mm into the box, through the centre of an end wall. For 30 – 60 s prior to firing, the box was purged with a stream of compressed air from the laboratory main, introduced through an opening in the upstream end wall and allowed to escape through an opening in the downstream wall. Just before device firing, the flow of purge air was stopped and the downstream endwall of the box was removed.

This method was successful in reducing light scattering from a drug-free jet, suggesting that some constituent of the atmosphere in the laboratory was responsible for unwanted light scattering. However, when a device with a drug payload was fired inside the box, a cloud of low-velocity, almost stationary drug particles was formed which lingered inside the box for several seconds. The influence of this cloud on the time-integrated velocity measurement yielded a velocity field with strong local variations. In some cases, the tracks of individual particles were apparent as streaks across the image. It is presumed that this low-velocity particle cloud was due to an internal flow, driven or initiated by the jet from the device, which caused the particles to recirculate inside the box.

In an alternative approach to control of the local atmosphere, the conical nozzle device was set up in free air, as usual, and fired while a steady low-velocity jet from the laboratory compressed air supply was directed vertically downwards across its exit. A black-corrected reference image obtained in this way is shown as the third image in ***Figure 5.23***, which illustrates that this approach was successful in reducing the signature of the drug-free jet. Unfortunately, it is inevitable that this external crossflow will have some effect on the flow field which is the subject of measurement. However, the crossflow velocity is expected to be small in comparison with the measured axial velocity of particles in the jet from the drug delivery device. Although there must be a complex three-dimensional interaction between the drug delivery jet and the applied crossflow, the greatest effect of the crossflow on the particles must be to alter the vertical component of their velocity — a component to which the DGV apparatus, as configured throughout the experiments described in this thesis, is not sensitive. This expectation that the particles' axial velocity is not affected is supported by the results of the definitive DGV measurements taken with crossflow, presented in Chapter 6, which show no detectable asymmetry in the velocity field.

The attempt to explain the nature of non-drug light-scattering material in the jet from the conical nozzle has been unsuccessful. The ineffectiveness of flushing the cylinder with helium indicates that condensation of water vapour stored in the cylinder is not a major contributor to non-drug light scattering. The partial success of firing into a box flooded with air from a compressed supply, as a means of reducing the optical signature of the drug-free jet, suggests that scattering is due to material from the atmosphere which is somehow entrained and concentrated in the jet.

The optical particle detection technique, described in §3.2, was used to gather a little more information about the phenomenon of light-scattering from the drug-free jet from the conical nozzle. A device fitted with nozzle pressure transducers and with an empty drug cassette was fired across a narrow laser beam, and the variation of light scattering intensity was recorded using a photodiode and a data acquisition system. A result showing particularly strong light scattering is shown in ***Figure 5.26***. It suggests that the non-drug scattering takes place very late in the run, some 20 ms after the initiation of flow, when detectable activity on the pressure transducer has come to an end. Though this result does not shed light on the mechanism of light scattering, it is an important finding for time-resolved DGV. Because the unwanted light scattering occurs much later than any drug delivery observed in the optical particle detection experiments of Chapter 4, instantaneous drug particle velocity fields can be measured without the worry of interference from non-drug material.

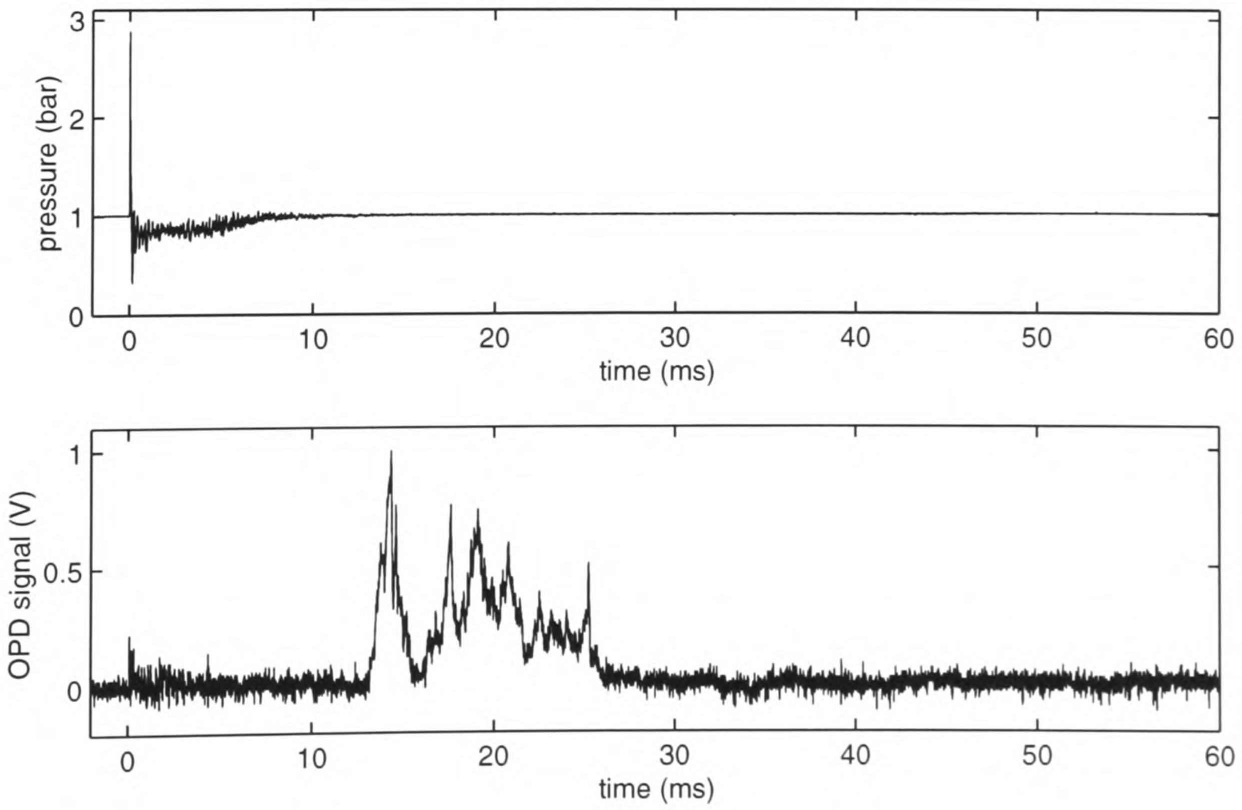


Figure 5.26 Pressure at the upstream nozzle transducer (*top*) and optical particle detection signal (*bottom*), recorded for a conical nozzle device without drug.

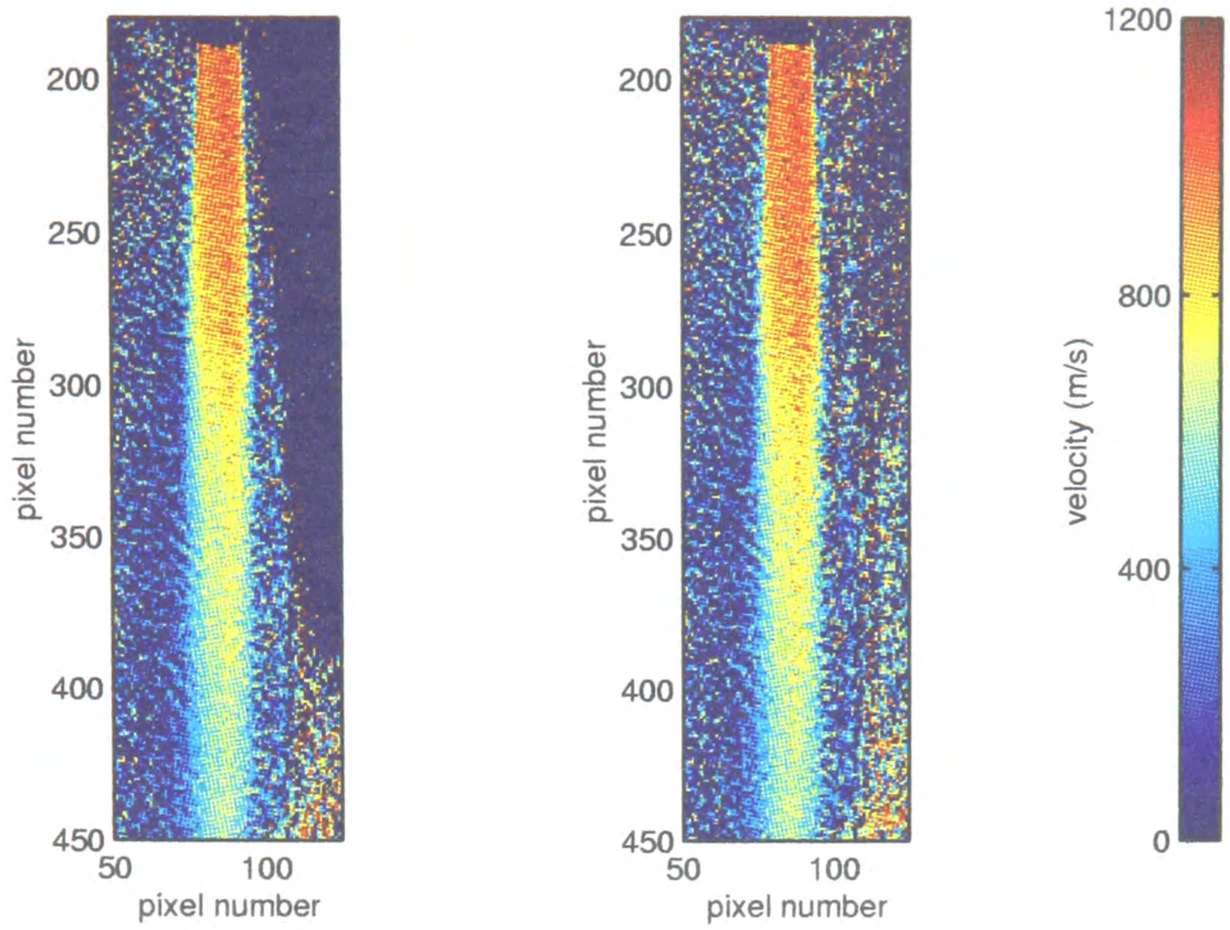


Figure 5.27 Time-integrated velocity field measured for 0.9 mg of 4.7 μm polymer microspheres delivered from the Mach 2.5 nozzle, with (*left*) normal image processing and (*right*) an image of a drug-free jet used in place of the black image.

Despite the precautions described above — flushing of the cylinder with helium, and, in the case of the conical nozzle, delivery of a slow air flow across the nozzle exit — some spurious light scattering still occurs. It remains to carry out a quantitative assessment of the extent to which this residual spurious signal distorts the measured velocity field of the drug particles. For both nozzle types, a typical CCD frame of a jet without drug particles, in place of the usual black frame acquired without a device firing, was subtracted from a frame of a particle-bearing jet. The resulting corrected frame was passed through the usual image processing algorithm to yield a velocity map which may be regarded as an approximation to the measurement which would be made if drug particles were the only light scatterers in the flow field. The velocity maps obtained in this way for flows from the Mach 2.5 nozzle and the conical nozzle are shown in ***Figure 5.27*** and ***Figure 5.28*** respectively, along with velocity maps generated in the usual way, without any compensation for spurious light scattering.

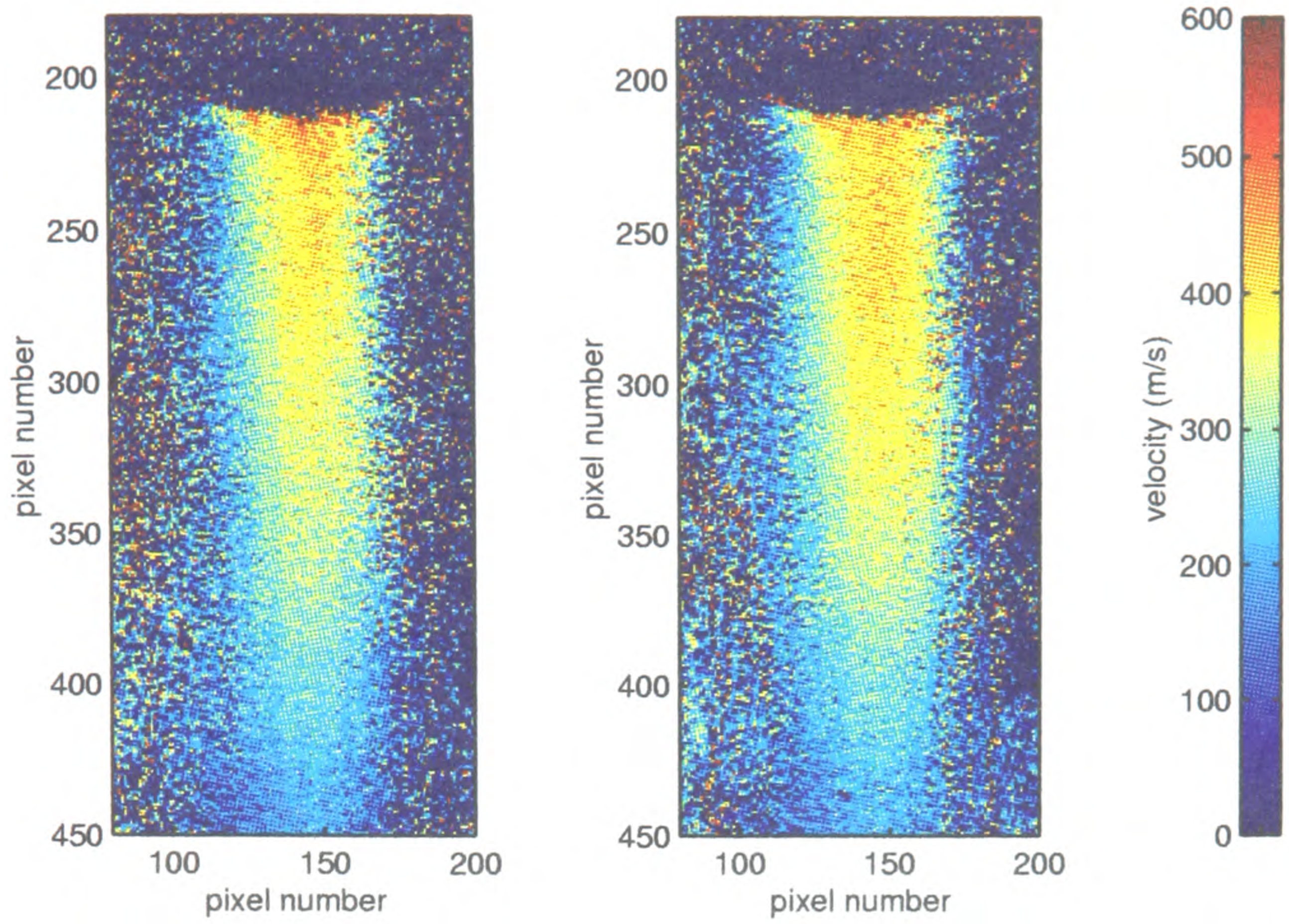


Figure 5.28 Time-integrated velocity field measured for 3.1 mg of 4.7 μm polymer microspheres delivered from the conical nozzle, with (*left*) normal image processing and (*right*) an image of a drug-free jet used in place of the black image.

For both the conical and Mach 2.5 nozzles, the discrepancy between velocity maps is slight. The maps generated by subtraction of a drug-free jet image show velocities that are generally higher by around 5%. This is considered to be an acceptable level of error. Although it would be possible to incorporate this correction technique as a routine part of the DGV measurement procedure, this is felt to be a potentially dangerous approach, since the mechanisms underlying the problem of spurious light scattering, and its correction, are not fully understood. Because the effects of spurious scattering on measured velocity are apparently small, velocity results are prepared and presented without any attempt to cancel the effects, other than the physical precautionary measures described in preceding paragraphs. Instead, a warning is given here that measured velocity fields are representative of particles other than the drug particles of interest, and that the velocity of the drug particles may be expected to be 5 to 10% greater than measured values.

5.5.2 Effect of Imaging System Aperture

In the context of optical imaging, the term *aperture* refers to a physical constriction in an optical system which limits the amount of light which can be processed by the system. A variable diameter aperture is built in to most photographic compound lenses, as a means of controlling the total light flux through the lens. The effective aperture of the DGV imaging system in this research has been controlled by varying the aperture of the front lens while the other two lenses are left fully open. Aperture is often expressed non-dimensionally as f-number, with the notation f/8, for example, denoting an aperture diameter equal to 1/8 of the focal length.

Since available light for time-integrated DGV experiments is effectively limited by laser power, it is desirable to select the largest possible aperture in order to pass on as much light as possible to the CCD camera, hence maximising the signal-to-noise ratio in the image. The choice of a large aperture also has deleterious consequences — reduction in depth of field is a familiar example in conventional photography. In the present DGV work, large front lens aperture has been found to have other unwanted effects, which have a strong impact on the quality of the discriminated image, leading to inaccuracy in the measured velocity map. This section describes a series of experiments designed to investigate these effects, and to help decide on a suitable f-number value for velocimetry. The problem first became apparent in images of dots test cards, and its effects on velocity fields were studied later.

When using the time-integrated DGV apparatus, it was noticed that at small values of front lens f-number (i.e. large aperture), discriminated images of dots alignment cards contained a blur which could not be removed by adjustment of the front lens focus. This aberration was further investigated in images of a test card consisting of white dots on a black background, as used for alignment calibration, alongside a blank white region containing some printed text. Frames of this test card are shown in ***Figure 5.29*** (front lens aperture f/3.5) and ***Figure 5.30*** (f/5.6). The aberration is barely discernible in these images, and is best appreciated by a comparison of the imaged text “4.7 mm dot pitch” in the two discriminated images.

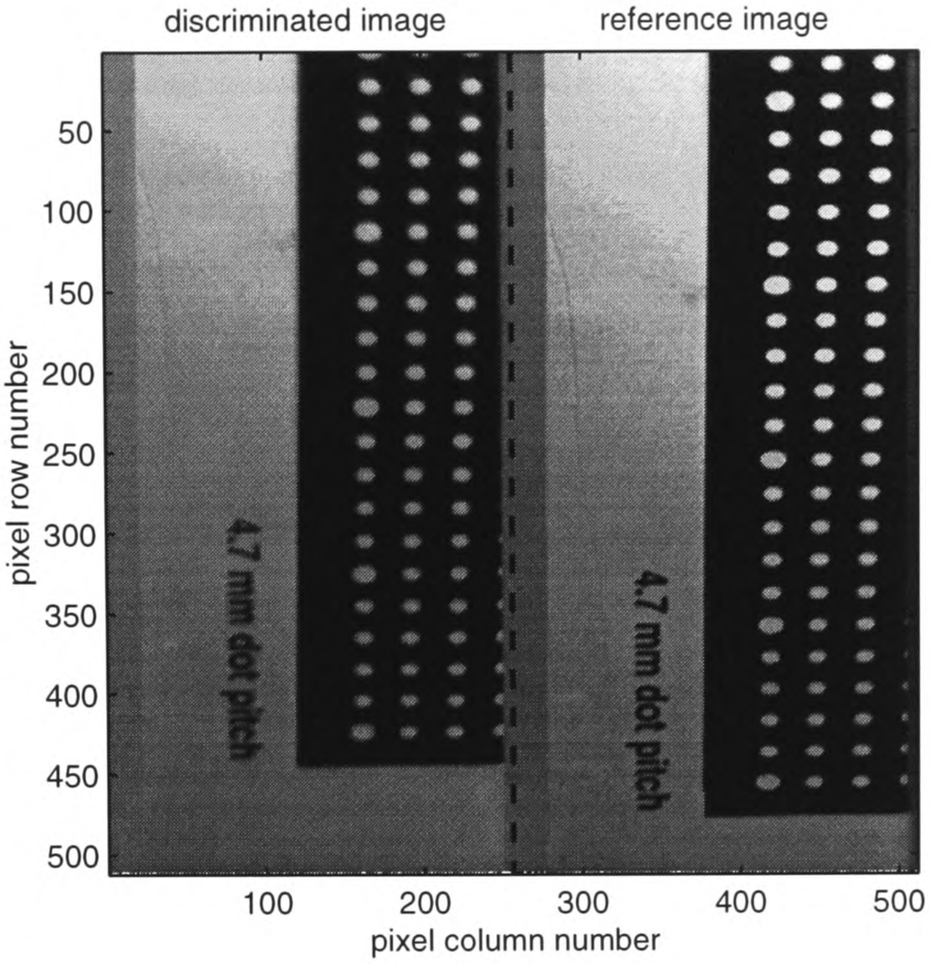


Figure 5.29 Frame of a test card taken at front lens f-number f/3.5.

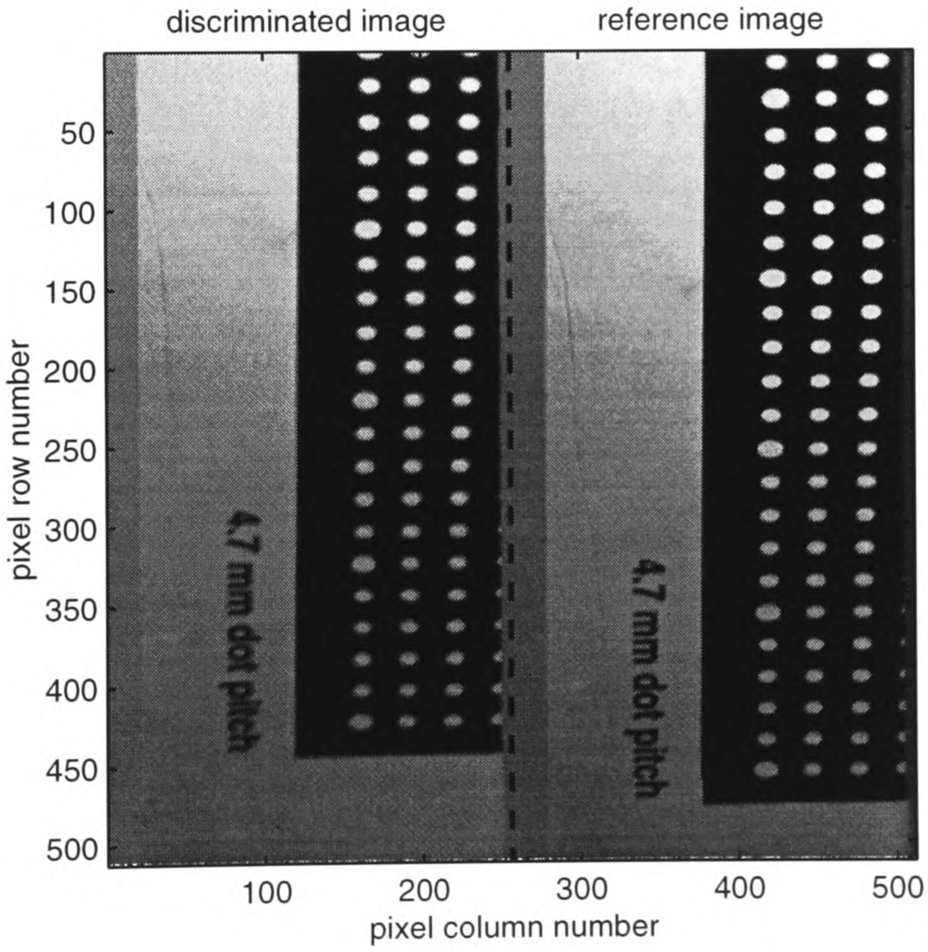


Figure 5.30 Frame of a test card taken at front lens aperture f/5.6. The test card, viewing geometry and illumination are the same as for the frame shown in ***Figure 5.29***.

The blur may be more clearly visualised in line plots of image intensity, as shown in ***Figure 5.31*** and ***Figure 5.32***. These graphs show the variation of intensity along pixel rows and columns which pass through approximate dot centres, in both reference and discriminated images of the test card. In the preparation of these graphs, the images were normalised in two stages. First, black and normalisation frames were applied in the usual way. They were then normalised by a scalar value, derived from a mean intensity over a large region of the (already normalised) image, to approximately equalise intensity levels between the reference and discriminated images, and facilitate comparison. The graphs on both a pixel row (***Figure 5.31***) and a pixel column (***Figure 5.32***) suggest that in the reference image, aperture has little effect on the form of the image. In the discriminated images, however, the intensity level in small image features (dots) is reduced and the sharpness of image intensity gradients (e.g., at the boundary of the blank white region) is lost as f-number is reduced to 3.5. Moreover, intensity in the black areas between dot rows is raised at f/3.5. This information suggests that at f/3.5, an aberration exists which smears out bright features, predominantly in the vertical direction.

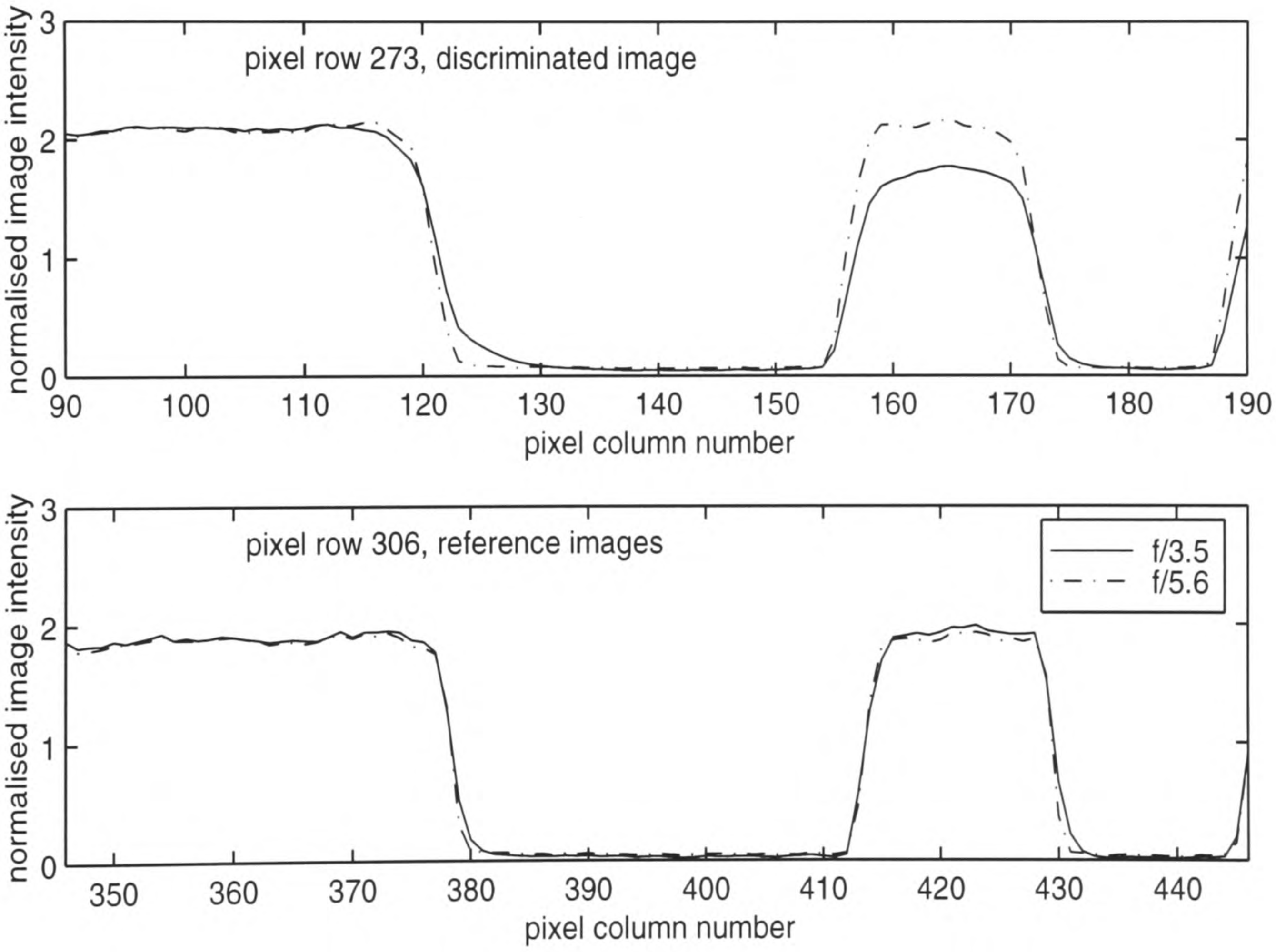


Figure 5.31 Intensity profiles on (horizontal) pixel rows passing through dot centres in images shown in the two preceding figures.

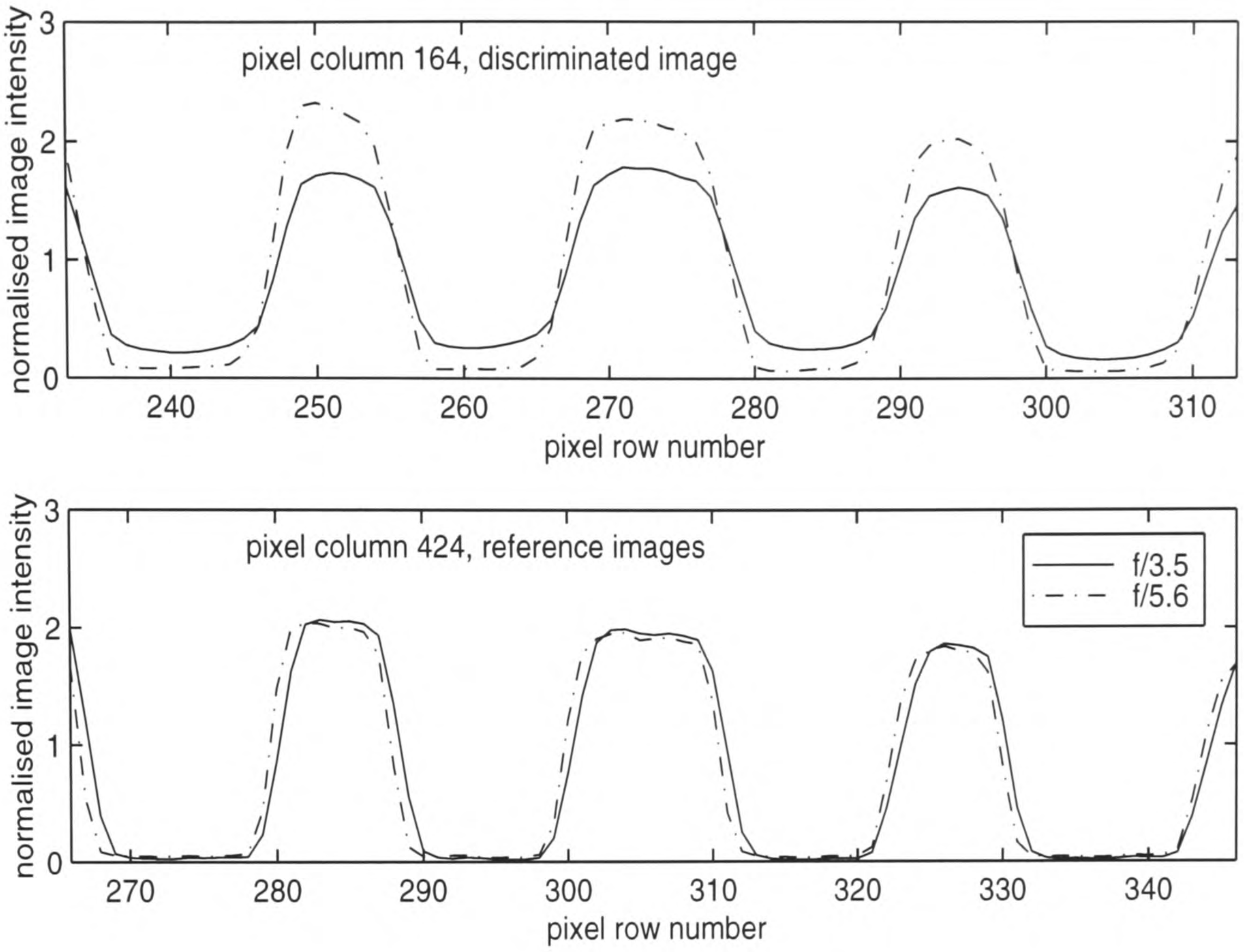


Figure 5.32 Intensity profiles on (vertical) pixel columns passing through dot centres in the test card frames shown in ***Figures 5.29*** and ***5.30***.

If all elements of the DGV experiment function correctly, the normalised reference and discriminated images of any stationary object should be identical, as the reflected light is not Doppler-shifted. Differences between the reference and discriminated images of the test card, illustrated by the experiment described here, would be interpreted through the image processing algorithms as the effect of a non-zero velocity field. Similarly, in a velocimetry experiment, the aberration could be expected to cause misrepresentation of the light intensity field in an image of moving light-scattering particles, and consequent velocity errors. This was confirmed by measuring the velocity of the jet from a Mach 2.5 nozzle device, loaded with 1.0 mg of 15.5 μm microspheres, at various f-numbers. Resulting velocity maps are shown in ***Figure 5.33***. The measured exit velocity decreases by about 100 m/s as f-number increases from 3.5 to 4 and 5.6. At f/8 and f/11, because of the limited available light, raw image intensity and signal-to-noise ratio are severely degraded. If the velocity maps for these apertures are heavily smoothed to suppress the noise, the reported velocities are not substantially different from those measured at f/5.6. It appears, therefore, that the aberration effect is insignificant at apertures of f/5.6 or less, but that the use of f/8 or f/11 is not feasible, in view of the limited available light. On the basis of these observations, it was decided that subsequent measurements on Mach 2.5

nozzle jets should be carried out at a front lens aperture of f/5.6, with the transfer and camera lenses fully open.

In a programme of similar experiments with the conical nozzle, on the other hand, there was no apparent effect of f-number on measured velocity, because of the gentler variation of image intensity in the conical nozzle flow-fields. The maximum available aperture, f/3.5, was therefore used for routine time-integrated velocity measurements of flow from the conical nozzle. In any case, image intensity was severely restricted by available laser power in work with this nozzle, and a higher f-number was not a viable option. For the majority of time-resolved measurements (all of which was carried out with the conical nozzle) an aperture of f/8 was used. No aberrations were detectable in images of dots alignment cards taken with the imaging system used in time-resolved measurements, and there was ample illumination energy to operate at much less than maximum laser output, even at high f-numbers.

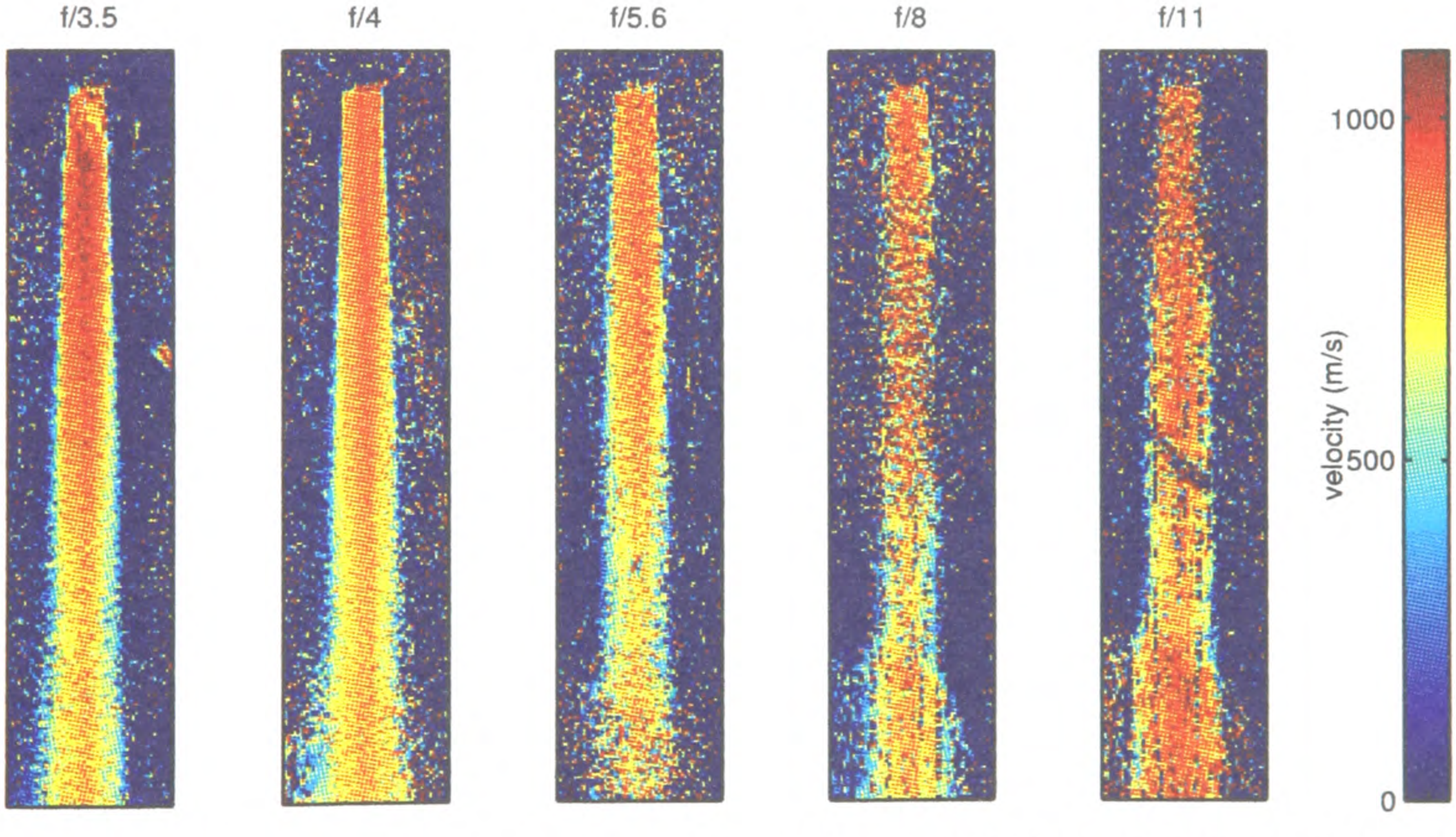


Figure 5.33 Time-integrated velocity maps of the jet from a Mach 2.5 nozzle device with 1.0 mg of 15.5 μm microspheres, measured at various values of front lens f-number.

The aberration discussed in preceding paragraphs occurs only in discriminated images. It is therefore likely to be due to some interaction of the collimated light bundle with the iodine cell's glass structure. The diameter of the collimated light bundle (in the imaging system used for time-integrated DGV) is shown in ***Table 5.1*** for a range of f-number values. The internal diameter of the iodine cell is 47 mm, and is larger than the light bundle diameter in all cases, indicating that there should be no interaction of the bundle with the cylindrical wall of the cell. However, the end windows of the cell are not perfectly flat near their periphery. A visibly distorted image of the background can be seen when looking through the outer parts of the cell, along its axis. This distortion is probably due to deformation of the windows which occurred when they were fused to the wall at high temperature in manufacture. This deformation is the most probable cause of the imaging problems at low f-number. This explanation is consistent with the fact the imaging system used in time-resolved measurements performs well at all f-numbers. It has an iodine cell of similar dimensions, but lenses of a smaller absolute aperture than the system used in time-integrated work. Hence, the collimated beam is smaller.

f-number of front lens	3.5	4	5.6	8	11
diameter of collimated light bundle (mm)	35	28	22	16	13

Table 5.1 Variation of the diameter of the collimated light bundle in the DGV imaging system, with aperture of the front lens, while the other lenses are fully open.

Operation of the DGV imaging system at less than maximum aperture does not take full advantage of the light-gathering capability of the lenses, which is a valuable asset in measurements where scattered light is restricted. For future work, therefore, it is recommended that the problem be eliminated at source by installing an iodine cell with a larger diameter.

5.5.3 Velocimetry Inside the Nozzle

For time-resolved measurements, a conical nozzle was built in perspex, with the aim of at least qualitatively visualising the shape of the particle cloud as it moved down the nozzle. Of course, velocity measurements inside the nozzle were desirable, but this was not expected to be straightforward. Since the nozzle wall forms a curved window between the viewing system and the image measurement region, distortion and other optical aberrations might corrupt the measurement by causing misalignment of the reference and discriminated images. Reflections or glare off the nozzle's interior surface might also conceivably be a problem. Scattered light might enter the imaging system after a reflection off the nozzle wall, and would be indistinguishable from light gathered directly from the particles. Since the initial scattering direction of multiply reflected light would be unknown, its frequency shift would be an unknown function of velocity.

§6.4 includes discussion of the success of velocity measurements inside the nozzle, based on empirical assessment of measurement results. In this section, experiments are described in which the effects of imaging aberrations due to viewing through the nozzle wall were investigated independently of velocity measurement experiments. An alignment calibration was carried out in the absence of a nozzle in the usual way. A second dots test card, trimmed to fit, was then placed inside the (horizontal) nozzle in a vertical plane containing the nozzle centreline. The nozzle, containing the card, was located so that the card lay in the same plane as the original dots card used for alignment calibration. Images of the card inside the nozzle were then taken using the DGV imaging system, and warped using the alignment calibration data generated earlier. The resulting warped reference and discriminated images are shown in ***Figure 5.34***.

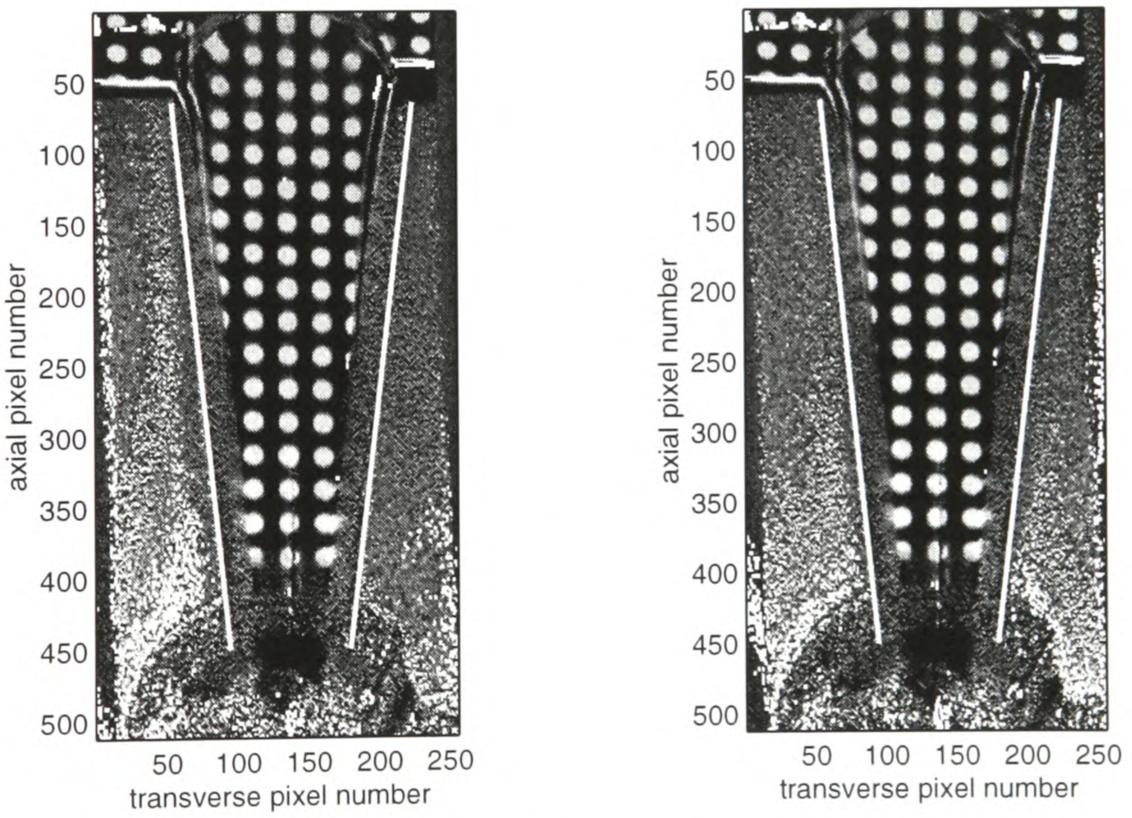


Figure 5.34 Warped discriminated *(left)* and reference *(right)* images of a dots test card positioned inside a perspex conical nozzle. The superimposed oblique white lines indicate the position of the exterior nozzle wall.

Severe aberrations, in the form of blurring of the dot shapes, are apparent near the nozzle rim and near the walls, particularly near the upstream end of the nozzle. Some slight blurring in the axial direction is also apparent for all dots near the downstream end of the nozzle. The important question, however, is whether the ability of the warp algorithm to align the two images to each other has been degraded. In ***Figure 5.35***, the warped reference and discriminated images are compared precisely in axial profiles of image intensity through the centres of various dots on the centreline, and transverse profiles through a row of dots approximately 7 mm from the exit plane. All the dots have jagged profiles, probably due to the surface finish on the perspex. Excellent alignment, to within less than 0.15 pixels, is displayed in transverse profile on the edges of the dots away from the wall. Axial alignment of the dots is poorer, with typical errors of 0.2 to 0.4 pixels. Very close to the wall, the dot images lose all definition.

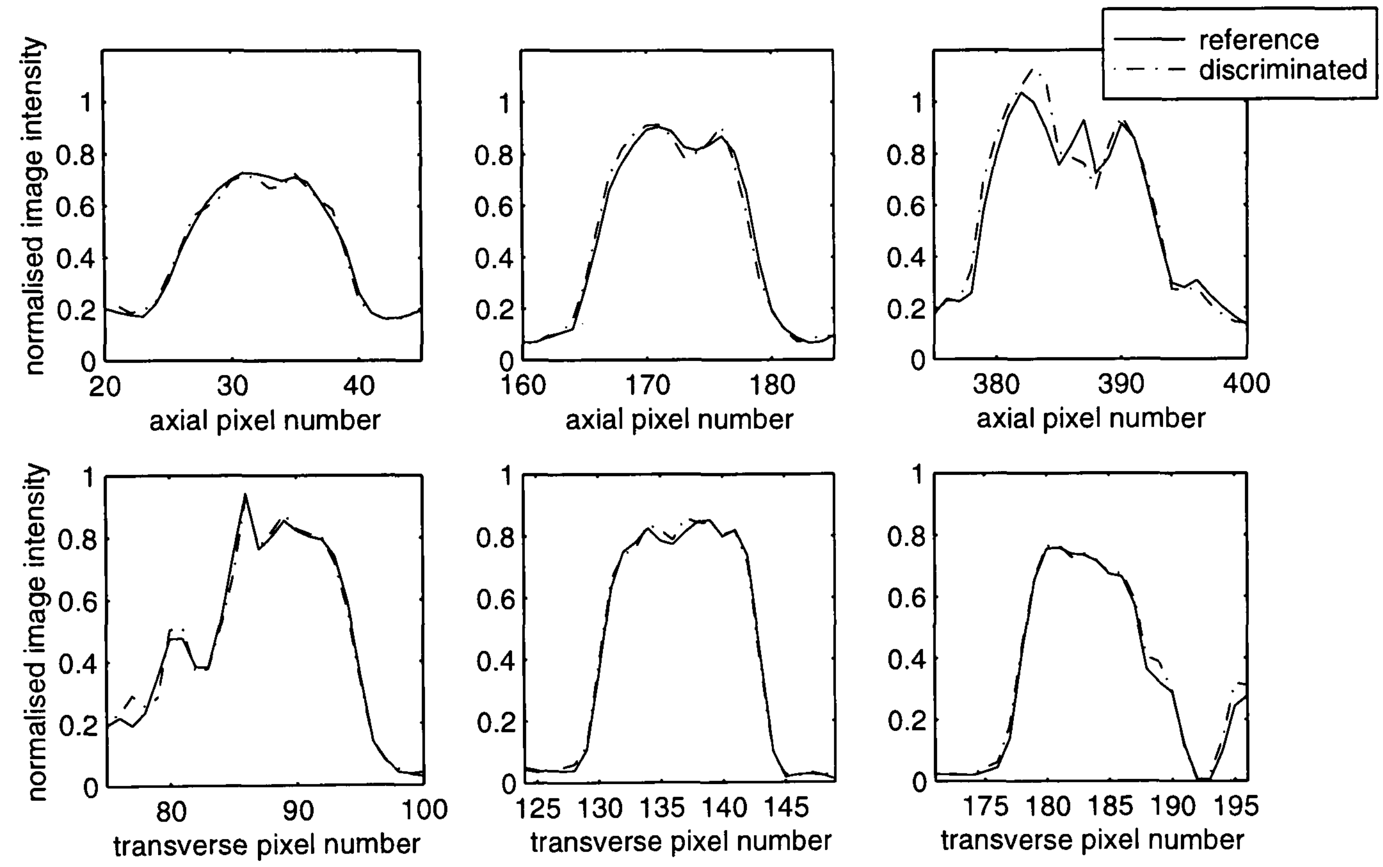


Figure 5.35 Plots of image intensity in warped images (shown in ***Figure 5.34***) of a dots card positioned inside a perspex nozzle. The top row contains plots along the nozzle centreline. The bottom row contains plots across the nozzle cross-section at axial pixel number 149.

For comparison, an equivalent set of reference and discriminated intensity profiles from warped images of a set of dots, imaged without an interposed curved window, are shown in ***Figure 5.36***. Alignment is so accurate that the edges of the reference and discriminated dot images are almost indistinguishable. Actual gaps between the edges are less than 0.1 pixels everywhere. It seems, therefore, that alignment quality is degraded by optical aberrations due to the nozzle wall. However, images are not entirely corrupted, and the effect may be tolerable.

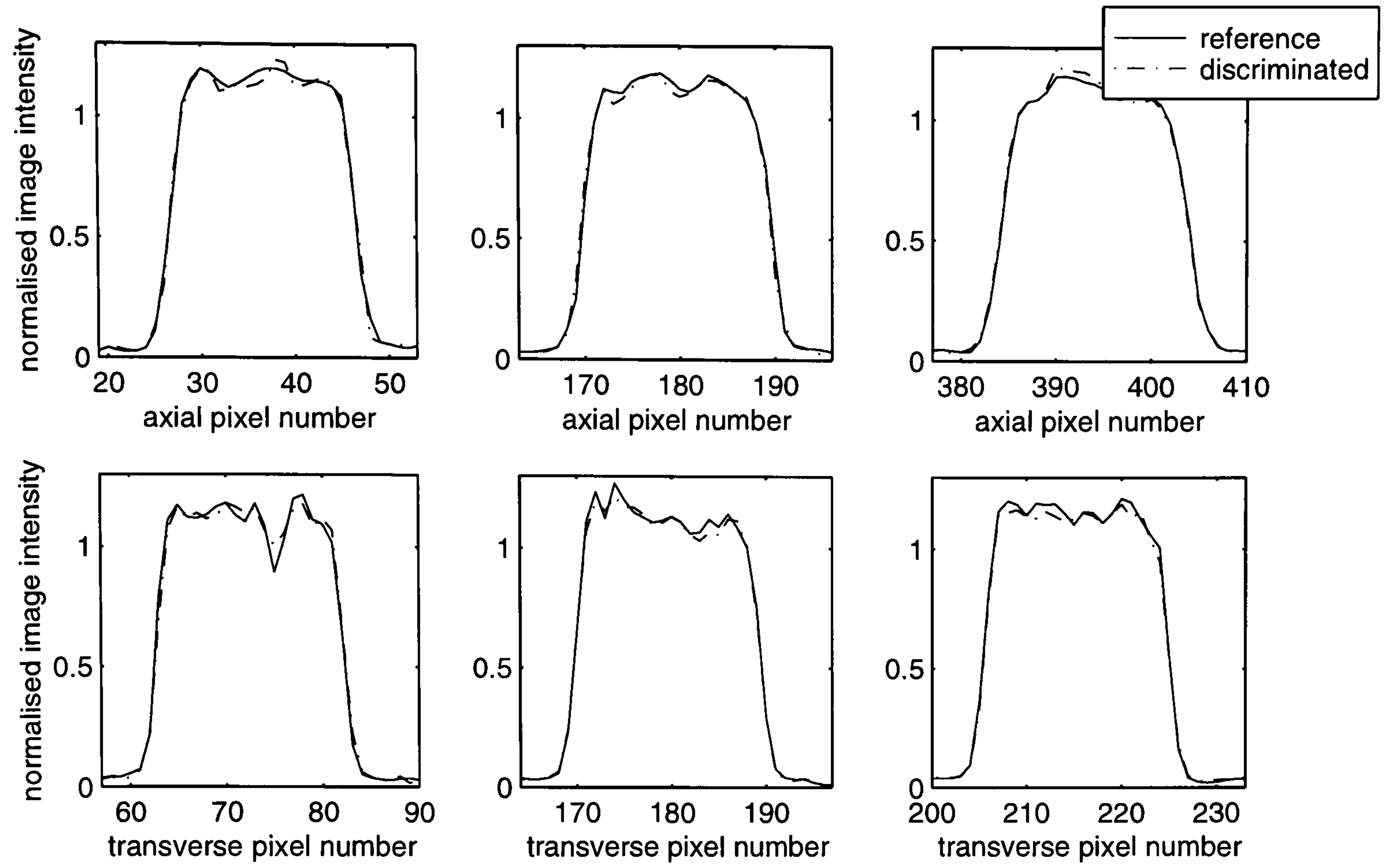


Figure 5.36 Plots of image intensity along pixel columns *(top)* and pixel rows *(bottom)* in warped images of a dots card imaged without any interrupting window.

5.5.4 Interpretation of Time-Integrated Velocity Measurements

Time-integrated DGV has been used extensively in this research to make measurements in strongly unsteady short-duration flows from drug delivery devices. Each such experiment yields a single measured velocity field which is representative of the complete history of the unsteady particle velocity field. It cannot be taken for granted that the measured velocity map represents a simple average, over time or over particles, of the time-varying velocity field. This is why the term *time-integrated*, rather than the more specific *time-averaged*, is used in this thesis as a shorthand for DGV based on continuous illumination and long exposure times. Some mathematical analysis has been carried out in an attempt to gain insight into the physical meaning of the time integrated velocities, and it is presented in this section.

Smith *et al.* [54] (who refer to DGV by a synonymous term, *Absorption Filter Planar Doppler Velocimetry*) touched on this issue. Having carried out 100 instantaneous velocity measurements of a turbulent fluctuating sonic jet, using a pulsed Nd:YAG laser, they generated two time-integrated velocity fields: one was a simple time average of the measured instantaneous velocity fields, and the other was based on the average of 100 pairs of raw instantaneous flow-field images. The two resulting velocity fields were similar in regions of steady flow, but differed substantially in the shear layers, where strongly unsteady turbulent fluctuations were significant.

In the velocity field which results from a time-integrated DGV experiment, the velocity reported at any pixel is a function of the measured transmission ratio for that pixel. This is calculated in image processing as the ratio of the pixel's intensity in the warped discriminated image to the intensity at the corresponding pixel in the warped reference image. The pixel intensity values represent the accumulated light energy which arrived at that pixel over the entire camera exposure period. Ultimately, time-integration takes place physically, in this way, on every pixel.

This is expressed in ***Eq. 5.22***, where $\bar{T}$ is the measured time-integrated transmission ratio at an arbitrary pixel. $I(t)$ is the normalised, black-corrected value of the intensity of light scattered onto the pixel in the reference image at any time, and is a function of the particle's size and the local illumination intensity. τ is the camera exposure time. $T(V(t))$ is the true iodine cell transmission ratio at which light destined for the chosen pixel at time t is processed by the iodine cell; it depends only on the velocity V of the particles scattering light onto the pixel at that time.

$$\bar{T} = \frac{\int_0^{\tau} T(V(t))I(t)dt}{\int_0^{\tau} I(t)dt} \tag{5.22}$$

This equation will now be rewritten as a summation of the contributions of N individual particles which pass through the space corresponding to an arbitrary pixel in the final velocity map. The subscript i will be used to identify individual particles; if s is the length in the local flow direction of the flow-field region corresponding to the pixel, then particle i spends a time s/V_i scattering light onto the pixel. The measured transmission ratio is then given by ***Eq. 5.23***.

$$\bar{T} = \frac{\sum_{i=1}^{N} T(V_i)I_i\frac{s}{V_i}}{\sum_{i=1}^{N} I_i\frac{s}{V_i}} \tag{5.23}$$

The numerator and denominator represent the sum over the flow history of all particle contributions to the discriminated image and the reference image, respectively. This equation is a fully general description of the time-integration process. It shows that the time-integrated transmission ratio is certainly not, in general, equivalent to a time average, or an average over all particles. It is biased towards particles which scatter brightly, and particles with low velocity. The implications for measured velocity are further complicated by the non-linear dependence of transmission ratio on velocity.

Inspection of ***Eq. 5.23*** immediately reveals a trivial result. If all particles have the same velocity, then regardless of variations in scattering intensity, the uniform transmission ratio (and hence velocity) will be reported correctly. This is the situation in time-integrated DGV of steady flows.

To proceed with this model towards some understanding of how a measured velocity field relates to the true unsteady field, some restrictive assumptions are required. First, it is assumed that all particles are the same size, which is approximately true for much of the work (on polystyrene spheres) described in this thesis. At a given pixel, the only influence on scattered light intensity is then the variation of illumination intensity across the thickness of the laser light sheet. However, if the particle concentration is such that a very large number of particles passes through the space corresponding to the pixel of interest in a short time, so that their velocity is effectively constant, they may be modelled as a collection of particles with a representative mean scattering intensity value. Intensity I_i may then be assumed constant for all particles. To show this rigorously, the summations in ***Eq. 5.23*** can be partitioned into M sums over short periods of effectively constant velocity V_j, in each of which N_j particles scatter onto the pixel of interest. This gives:

$$\bar{T} = \frac{\sum_{j=1}^{M} \sum_{i=1}^{N_j} T(V_j) I_i \frac{s}{V_j}}{\sum_{j=1}^{M} \sum_{i=1}^{N_j} I_i \frac{s}{V_j}} \tag{5.24}$$

s is a constant for the pixel, and cancels out; V_j and $T(V_j)$ are constants associated with each time period, and can be moved outside the sums over i. This gives:

$$\bar{T} = \frac{\sum_{j=1}^{M} \frac{T(V_j)}{V_j} \sum_{i=1}^{N_j} I_i}{\sum_{j=1}^{M} \frac{1}{V_j} \sum_{i=1}^{N_j} I_i} \qquad (5.25)$$

Defining $\bar{I}$ as the average scattering intensity for a large number of particles distributed across the light sheet, and N as the total number of particles (the sum of the N_j), the above expression reduces to ***Eq. 5.26*** below. $\bar{I}$ is constant for each pixel, and is independent of time and particle velocity.

$$\overline{T} = \frac{\bar{I} \sum_{j=1}^{M} N_j \frac{T(V_j)}{V_j}}{\bar{I} \sum_{j=1}^{M} \frac{N_j}{V_j}} = \frac{\sum_{i=1}^{N} \frac{T(V_i)}{V_i}}{\sum_{i=1}^{N} \frac{1}{V_i}} \qquad (5.26)$$

This can be further simplified by modelling transmission ratio as a linear function of velocity, as in ***Eq. 5.27***. This is a reasonable approximation in certain regions of the transmission profile, or if the range of velocities is small, but in some cases it may be quite inaccurate.

$$T(V_i) = a + bV_i \qquad (5.27)$$

Substituting this into ***Eq. 5.26***, and rearranging, the time-integrated measured velocity is given by:

$$\bar{V} = \frac{N}{\sum_{N} \frac{1}{V_i}} \ . \qquad (5.28)$$

This equation again highlights the disproportionate influence of low-velocity particles on the measured velocity value. It states that the measured velocity is the reciprocal of the average, over all particles, of the reciprocal of velocity. In cases where particles with a small range of velocities contribute to the pixel, this approaches an average over particles.

Without knowledge of the time-varying velocities and image intensities, it is impossible to make a precise assessment of the relationship between time-integrated measurement of velocity and the true velocity history. However, by assuming particle velocity distributions, it would be possible to examine the effects of various distributions on the time-averaged measurements, through computations based on the analysis presented above. The effects of an assumed particle size distribution could also be incorporated by calling on Mie theory to describe the variation of scattered light intensity with particle size.

The analysis presented in this section highlights the difficulty of interpreting time-integrated velocity data. The strongest conclusion it offers is that although the time-integrated velocity is representative of all particles, it is biased towards large particles and slow particles, and is also affected by the non-linearity of the iodine cell transmission characteristic. When the frequency of scattered light is in a near-linear region of the iodine cell profile, and the range of particle velocities is small, the measured velocity approaches an average over all particles. However, the data obtained through time-resolved DGV can be compared with time-integrated velocity measurements to enable a direct empirical assessment of the significance of biasing effects in this application of DGV. Such an analysis is presented in §6.5.2, and shows that the time-integrated velocity measurements do approximate the particle-averaged velocity in the measurements described in this thesis.

5.6 Summary

The DGV method has been described in this chapter, explaining the basic physical principles of Doppler shift and the use of an absorption filter, and showing how these ingredients have formed the basis of two working whole-plane velocimetry systems in Oxford University. The hardware and software involved in this implementation have been described in outline. One of the Oxford systems uses a continuous-wave Argon ion laser to provide time-integrated measurements, and the other system, with an injection seeded pulsed Nd:YAG laser, can be used for time-resolved measurements.

Some features of the present application of DGV to flow-fields in powder drug delivery have influenced the specific details of the experimental method in this research. A viewing geometry was chosen which determined that the axial component of velocity is measured. Very high velocities have required oblique viewing angles, and frequency measurement across the widest range available from the iodine cells.

The presence of light-scattering material in the flow-field, other than drug particles, posed a problem in time-integrated DGV. It was eventually solved by excluding air from the drug delivery device's helium reservoir, as far as possible, and in the case of the conical nozzle, by flushing clean air across the nozzle exit during a shot. Residual levels of unwanted light scattering still biased measured velocities, reducing them by approximately 5%. Also in time-integrated work, an imaging aberration due to the iodine cell, accentuated by the small scale of the flow-fields under study, was found to affect the measurement of flows from the Mach 2.5 nozzle when the imaging system aperture was less than f/5.6. This constraint on f-number exacerbated the difficulty of low levels of available light in all the time-integrated work.

Mathematical analysis of the precise meaning of the time-integrated velocity measurements of unsteady flows has been carried out. It has been shown that the time-integrated velocities do not, in general, correspond to a true time-average, but they approximate an average over all particles, if the particles are uniformly sized and the range of their velocities is not large.

Finally, the configuration of the DGV apparatus for the various groups of experiments undertaken in this research is summarised below in ***Figure 5.37*** and ***Table 5.2***. The imaging system and drug delivery device were placed in a horizontal plane, with the illuminating laser light directed into the drug delivery device in a vertical sheet. The major parameters which describe the configuration of each experiment are the viewing angle θ, the width (measured in a vertical direction) and thickness (horizontal and normal to the illumination direction) of the light sheet, and the CCD exposure time.

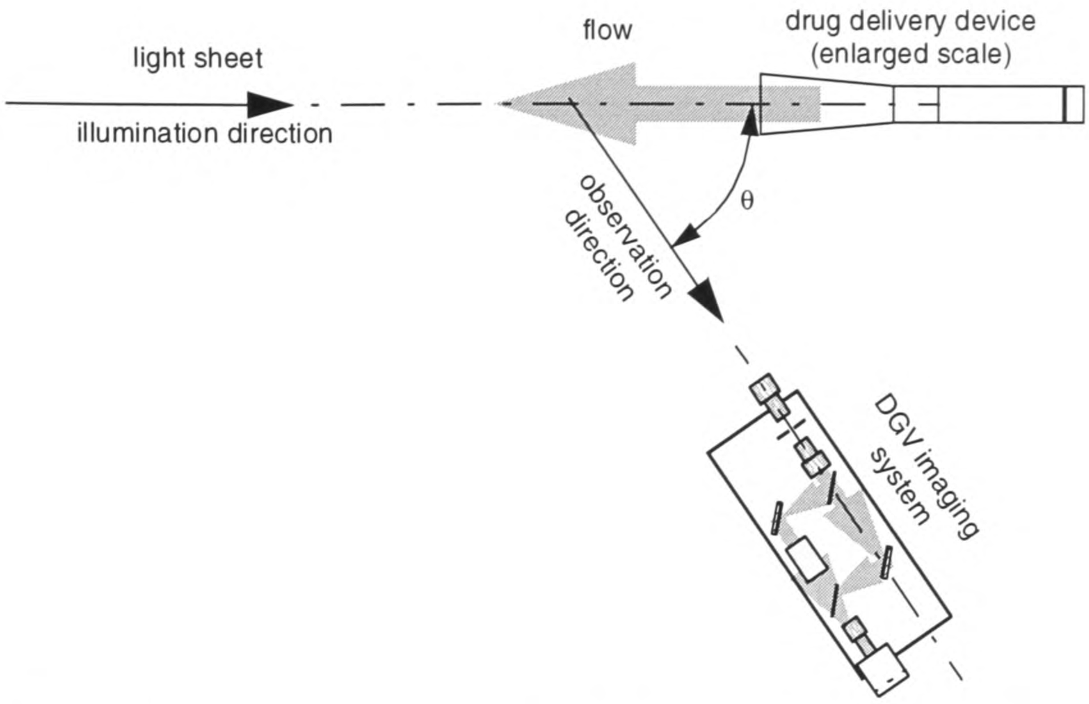


Figure 5.37 A schematic plan view of the DGV apparatus as arranged for measurements of the particle velocity field in flow from drug delivery devices.

	time-integrated		time-resolved (conical nozzle only)
	conical nozzle	contoured nozzle	
viewing angle (θ)	71°	46°	45°
light sheet width	25 mm	15 mm	50 mm
light sheet thickness	1 mm	1 mm	0.5 mm
CCD exposure time	0.8 – 1.0 s	0.8 – 1.0 s	0.5 – 2.0 s
illumination time	> 1 s	> 1 s	10 ns
front lens aperture	f/3.5	f/5.6	f/8

Table 5.2 A summary of the major features of the configuration of the DGV apparatus.

Chapter 6 Doppler Global Velocimetry: Results

DGV has provided a wealth of data on the axial velocity fields of particles in flow from drug delivery devices. It has shown itself to be a powerful technique in this context, extracting detailed information from flows which present a troublesome environment for measurement. This information constitutes a quantitative description of the most important mechanical operating characteristic of these devices, the particle velocity distribution. More importantly, it is central to the understanding of drug delivery flow-fields which has emerged through this research. These statements are particularly true of the information obtained through time-resolved DGV, which has revealed aspects of the flow-field that could not previously have been assessed.

The results of DGV measurements are presented and discussed in this chapter. Selected measured velocity fields are shown in full, and summary information is presented for all measurements. The emphasis is on the understanding which can be gathered from the data, rather than on an exhaustive presentation of the data. The chapter begins with a list of the devices and device configurations tested. Results are presented in three groups. Time-integrated results for the contoured nozzle are presented and discussed in §6.2, followed by time-integrated results for the conical nozzle in the next section. The time-resolved results for the conical nozzle are presented along with a detailed discussion of the flow-field in §6.4. Each class of results is accompanied by an uncertainty estimate, calculated for a representative case. Finally, the time-integrated and time-resolved results are compared, and the relationship between them is re-examined. The findings of the DGV experiments are summarised at the end of the chapter.

6.1 Test Conditions

All the measurements described in this chapter were carried out at a fill pressure of 60 bar, with 20 μm polycarbonate membranes, a 5 ml (nominal) cylinder, and a 0.2 ml (nominal) rupture chamber, while the particle material, particle size and payload mass were varied. Three grades of polymer microspheres, three grades of lidocaine, and Aerosil silica powder were used in test payloads. The full range of payload conditions tested are listed in ***Table 6.1***.

		Mach 2.5 nozzle			conical nozzle		
		1 mg	3 mg	5 mg	1 mg	3 mg	5 mg
polystyrene microspheres	4.7 μm	4	5	3	3, ***53***	5	3
	15.5 μm	5	6	3	1	3	3
	26.1 μm	6	4	2	—	1	3
lidocaine	20-38 μm	3	2	3	2, ***13***	3	3
	38-53 μm	3	3	2	3	3	3
	75-106 μm	3	2	3	4	3	2
silica		3	4	2	***20***	—	—

Table 6.1 Numbers of successful DGV velocity measurements obtained for each configuration of the drug delivery device and payload. Figures in ***bold italics*** denote time-resolved measurements; figures in plain text are for time-integrated measurements.

In the time-integrated work, repeated measurements were made for a range of payload configurations. In time-resolved DGV, measurements for one payload configuration — 1.0 mg of 4.7 μm polystyrene microspheres — were carried out in detail. A less detailed programme of time-resolved experiments was carried out for payloads of lidocaine and silica. The configuration of the experimental apparatus for all DGV measurements is summarised at the end of Chapter 5.

6.2 Time-Integrated Measurements: Contoured Nozzle

In this section, the results of particle velocity measurements in the flow from a drug delivery device with a Mach 2.5 contoured nozzle are presented and discussed. Some selected whole measured velocity fields are shown in full in §6.2.2. First, an explanation is given of the formats and conventions used in this chapter to represent time-integrated velocity data in a visual form. An estimate of the uncertainty of these measurements is detailed in §6.2.3. The complete data set is summarised in terms of exit velocities in §6.2.4, illustrating trends in velocity as particle size and payload mass are varied.

6.2.1 Format for Graphical Presentation of Results

A range of time-integrated DGV results is presented in §6.2.2. At the centre of each figure is a visualisation of one velocity field in which axial velocity values are represented by colour, ranging from dark blue for low velocity to dark red for highest velocity. The colour mapping is shown explicitly on a legend in each figure. As a consequence of the image processing procedure, the velocity field appears as if viewed from a direction normal to the measurement plane. The velocity field is shown with scales in mm, where zero corresponds to the centreline (in the transverse coordinate) and the nozzle exit plane (in the axial coordinate). Because the flow was viewed obliquely from upstream in experiments (as shown in ***Figure 5.37***), the nozzle wall obscures a small region of the measurement plane just downstream of the exit. The nozzle axis is vertical in all images, with flow from top to bottom.

Although there is no measurable or meaningful velocity outside the fringes of the gas-particle jet, the DGV software reports a non-physical velocity for these regions, effectively based on very slight apparent light scattering due to camera noise. In the presented velocity maps, spurious background velocity readings are suppressed by assigning zero velocity to background pixels, which are defined as pixels in which the reference image intensity is less than a small threshold value. This has been emphasised visually by assigning the colour white to pixels of negative or near-zero velocity.

A centreline velocity profile and four cross-sectional profiles are also included in each figure. The centreline profile shows the variation of velocity along the pixel column closest to the centreline, as well as a mean over the 9 pixel columns surrounding the centreline. The label at the top right corner of each cross-section plot shows the axial coordinate (measured downstream from the nozzle exit plane) at which the section is taken, and the baseline of each graph is aligned with that axial coordinate in the longitudinal profile and the velocity map.

6.2.2 Representative Results

Examples of measured particle velocity fields for flow from the Mach 2.5 nozzle are presented in this section. The results shown have been chosen to reflect the range of test conditions and of measurement quality which occur in the data set. Measurement quality, in this context, refers to noise levels in the final velocity maps, and was dictated by image intensity, which varied with payload type.

Results for polystyrene microspheres are shown in ***Figures 6.1*** to ***6.6***. The structure of the particle jet in all six of these images is consistent and simple. Near the nozzle exit, the particle jet is quite uniform, with little or no variation of velocity across its width. Further downstream, the jet becomes wider as a thin layer of low velocity develops at its periphery. On the centreline, the velocity remains approximately constant for a distance of 20 to 30 mm from the exit. It begins to fall off with further increase in distance from the nozzle exit. The cross-sectional velocity distribution then loses the uniform character which is apparent at locations near the exit, and a more gradually varying profile appears.

The measured exit velocity for polystyrene microspheres ranges from 830 m/s to 1030 m/s. These values lie between the velocity of approximately 554 m/s expected for gas behind the starting shock (derived from static pressure data in §4.1.3), and the 1450 m/s for quasi-steady inviscid supersonic flow, quoted in ***Table 2.1***. The range of measured velocities does not point definitively to either flow regime as the major mechanism for particle acceleration in this device. It is consistent with significant roles for both quasi-steady supersonic flow and shock tube flow; it is also consistent with the dominance of quasi-steady supersonic flow in which particles lag the gas flow substantially. However, the data seem to rule out the shock tube flow as the sole mechanism for particle acceleration.

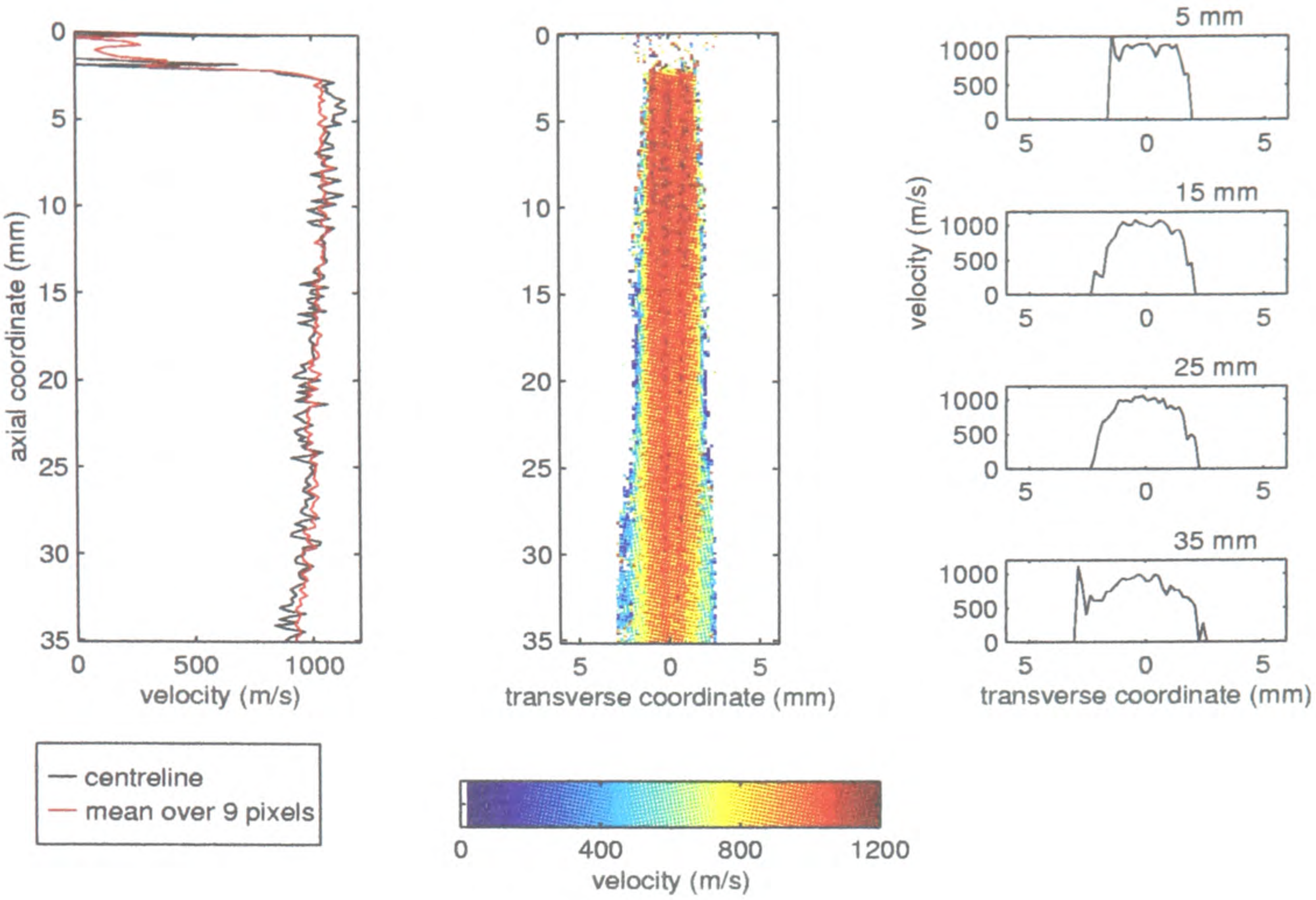


Figure 6.1 Time-integrated velocity field of 1.0 mg of 4.7 μm polymer microspheres delivered from a Mach 2.5 nozzle.

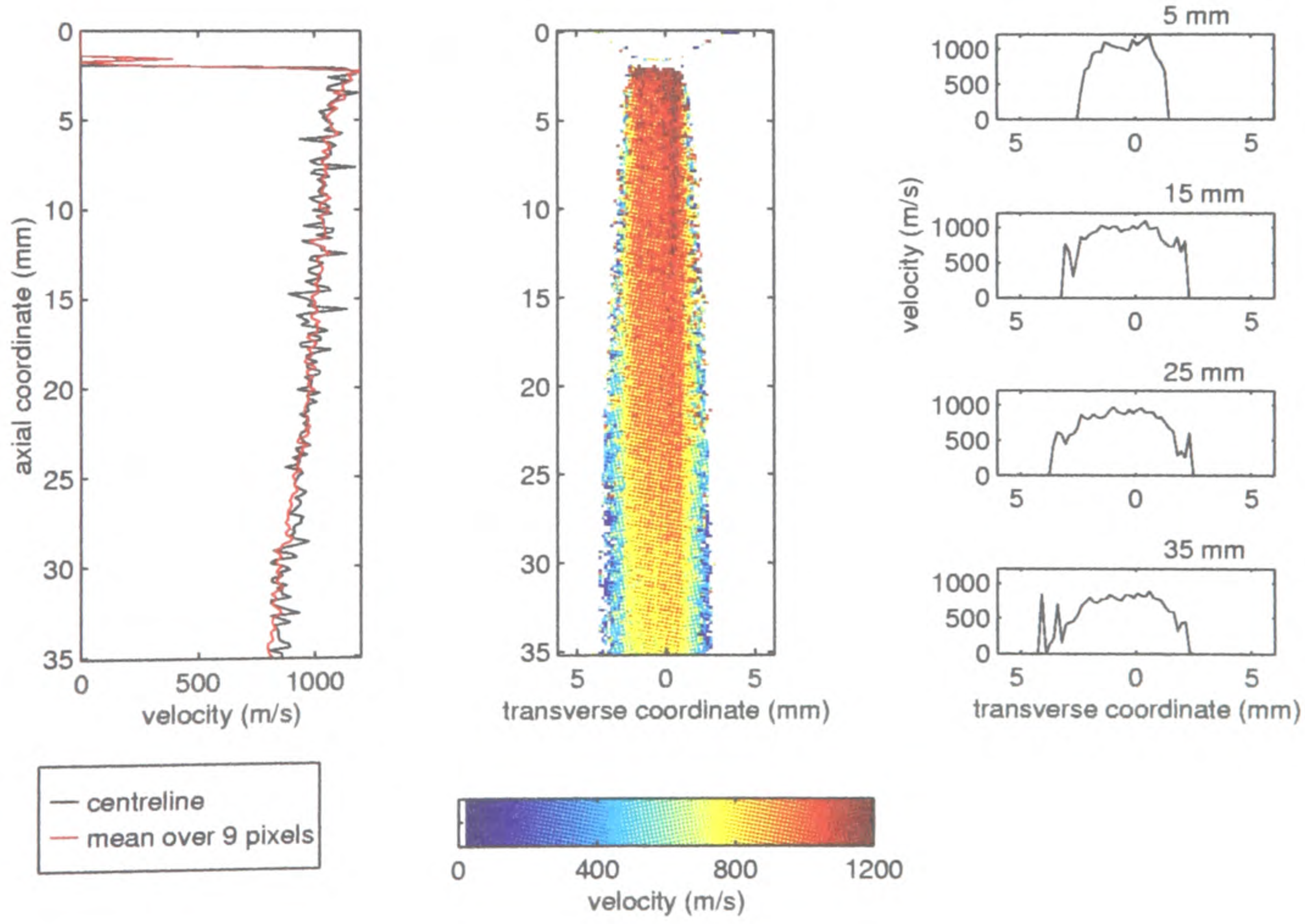


Figure 6.2 Time-integrated velocity field of 1.0 mg of 4.7 μm polymer microspheres delivered from a Mach 2.5 nozzle.

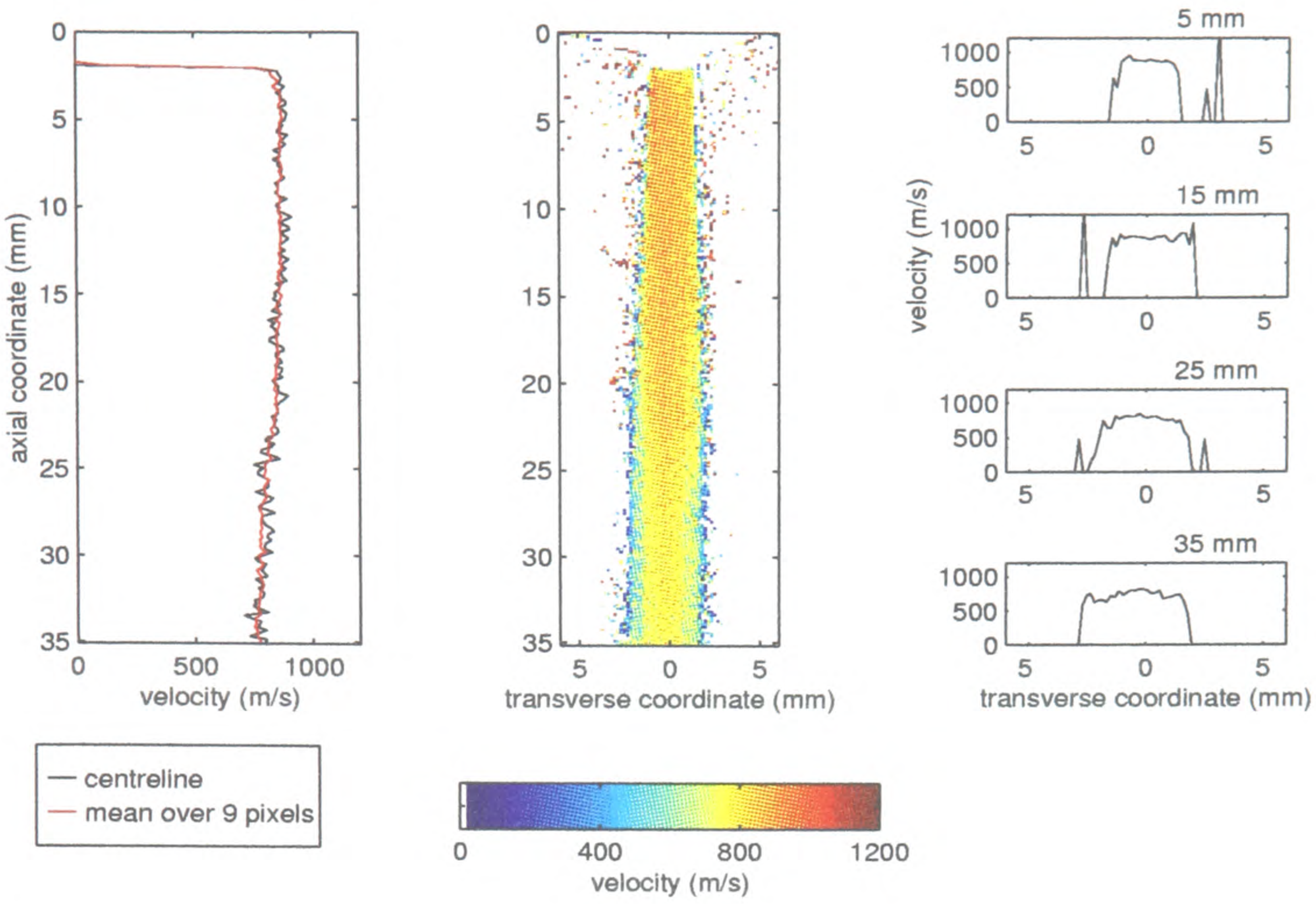


Figure 6.3 Time-integrated velocity field of 3.0 mg of 15.5 µm polymer microspheres delivered from a Mach 2.5 nozzle.

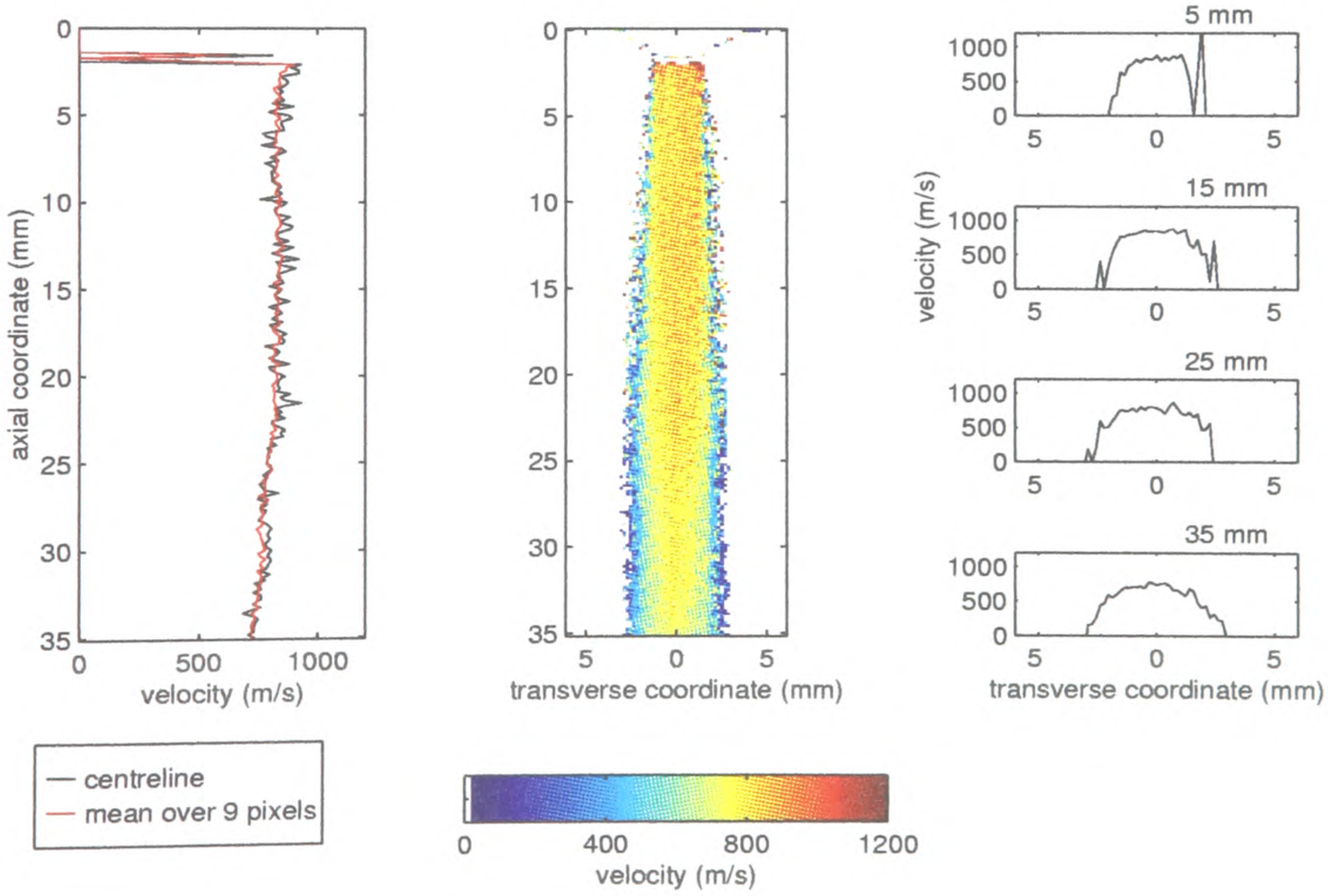


Figure 6.4 Time-integrated velocity field of 5.0 mg of 15.5 µm polymer microspheres delivered from a Mach 2.5 nozzle.

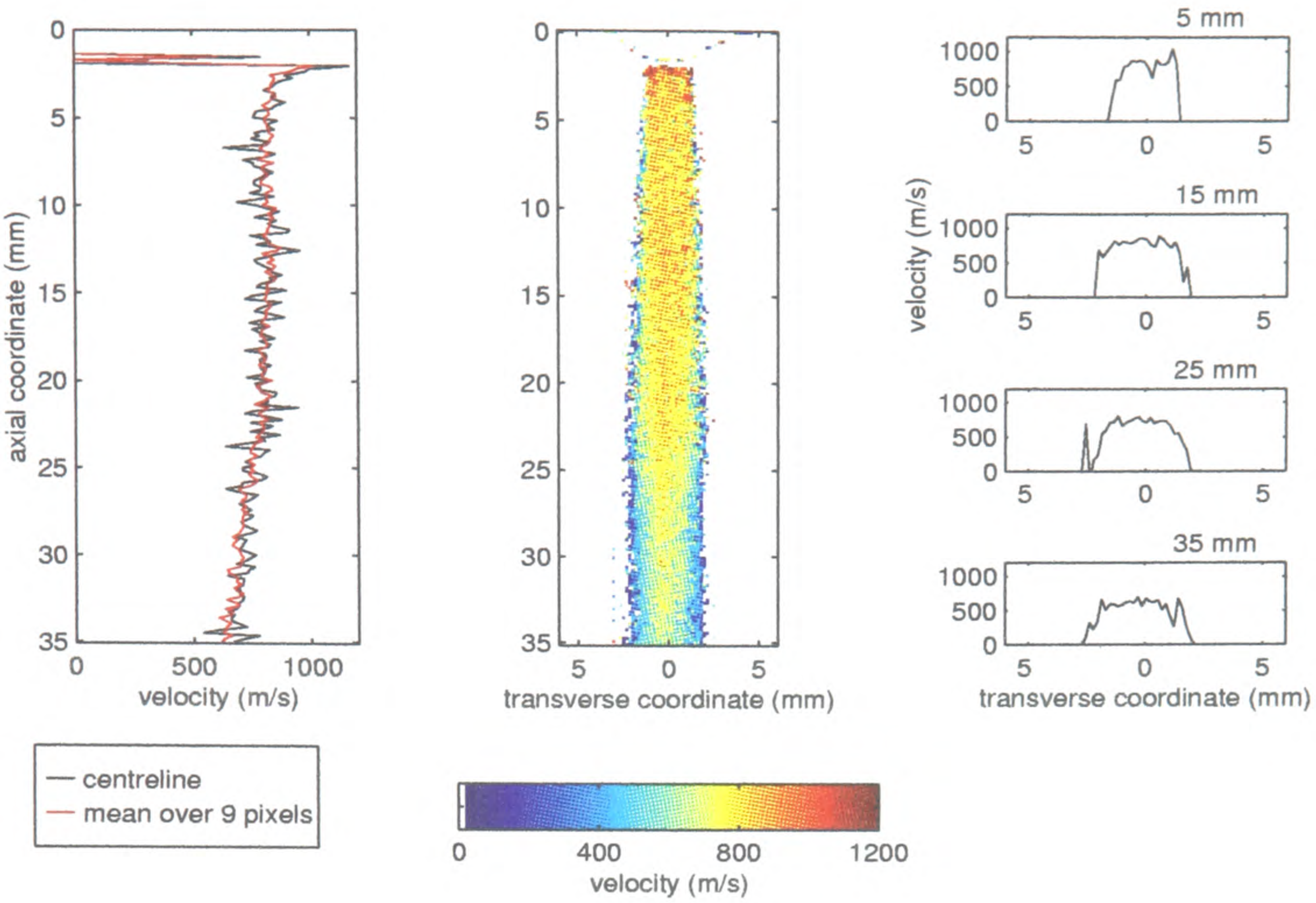


Figure 6.5 Time-integrated velocity field of 1.0 mg of 26.1 μm polymer microspheres delivered from a Mach 2.5 nozzle.

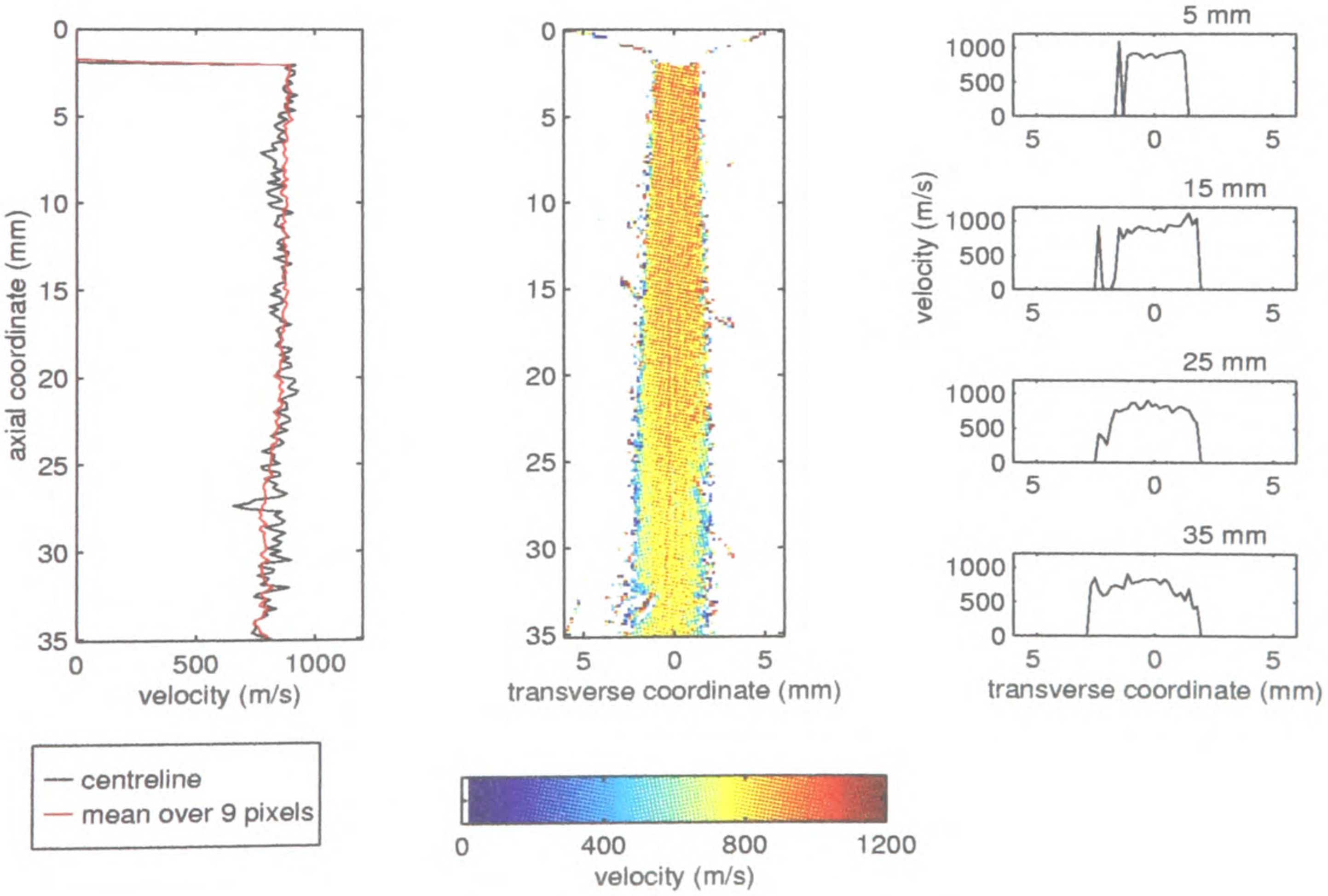


Figure 6.6 Time-integrated velocity field of 3.0 mg of 26.1 μm polymer microspheres delivered from a Mach 2.5 nozzle.

The measurement results shown in ***Figures 6.1*** and ***6.2*** were both obtained for 1.0 mg payloads of 4.7 μm spheres, and are presented as an indication of data repeatability. These particular results display excellent consistency. Two to six measurements were carried out for each microsphere payload configuration. In the summarised exit velocity data presented in §6.2.4, it can be seen that the scatter in exit velocity at each condition is between 50 m/s and 160 m/s, corresponding to a maximum of 18% of the mean measurement value.

Figures 6.7 and ***6.8*** show two typical velocity maps obtained for lidocaine payloads. The velocity fields of lidocaine particles are similar to those of polystyrene microspheres, with continuously decreasing velocity in the downstream and radial directions, outside a relatively uniform core. Spreading of the jet is more pronounced than in the case of microsphere payloads, and the overall velocities are somewhat lower.

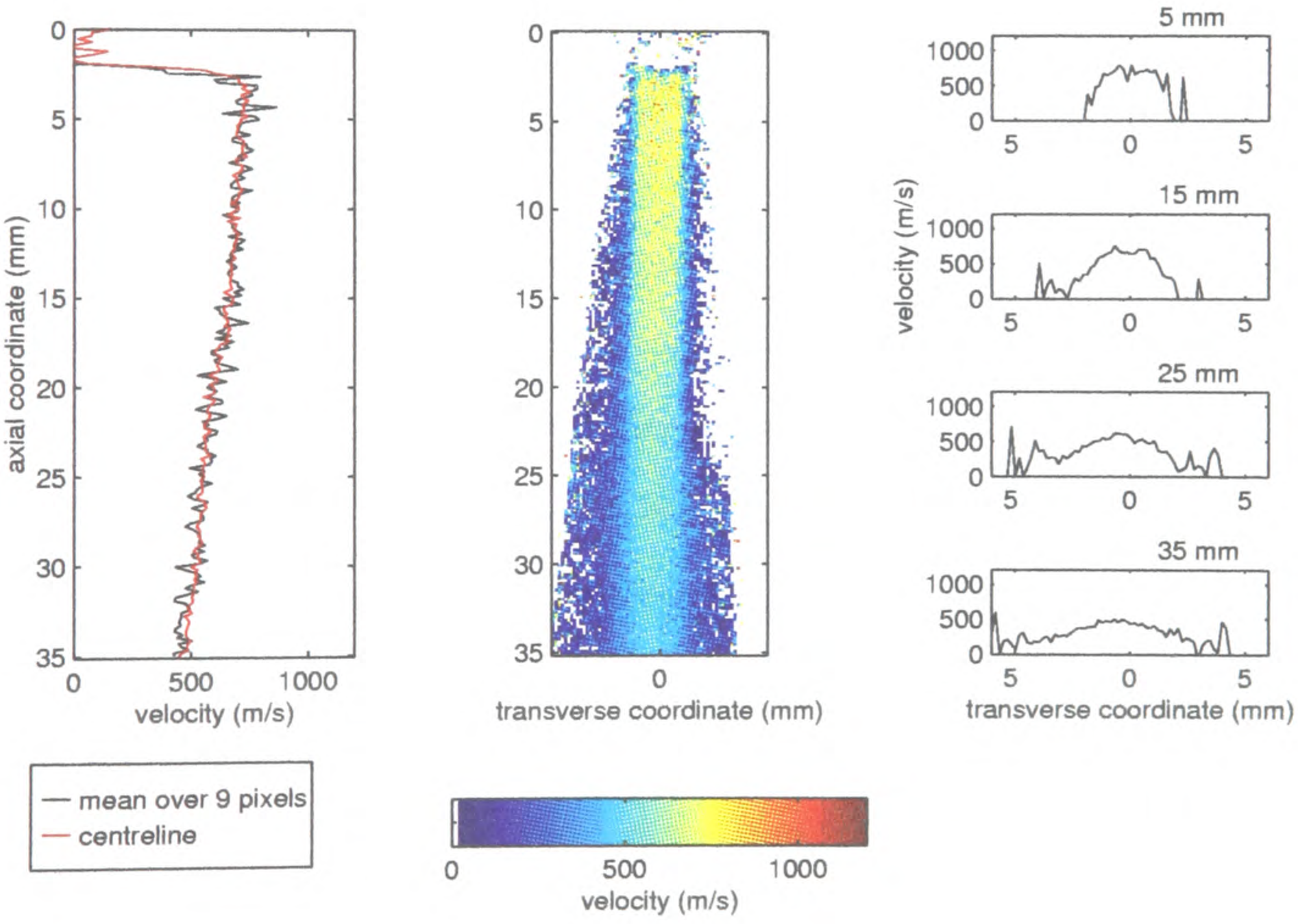


Figure 6.7 Measured velocity map for 5.1 mg of 20-38 μm lidocaine particles, delivered from a Mach 2.5 nozzle.

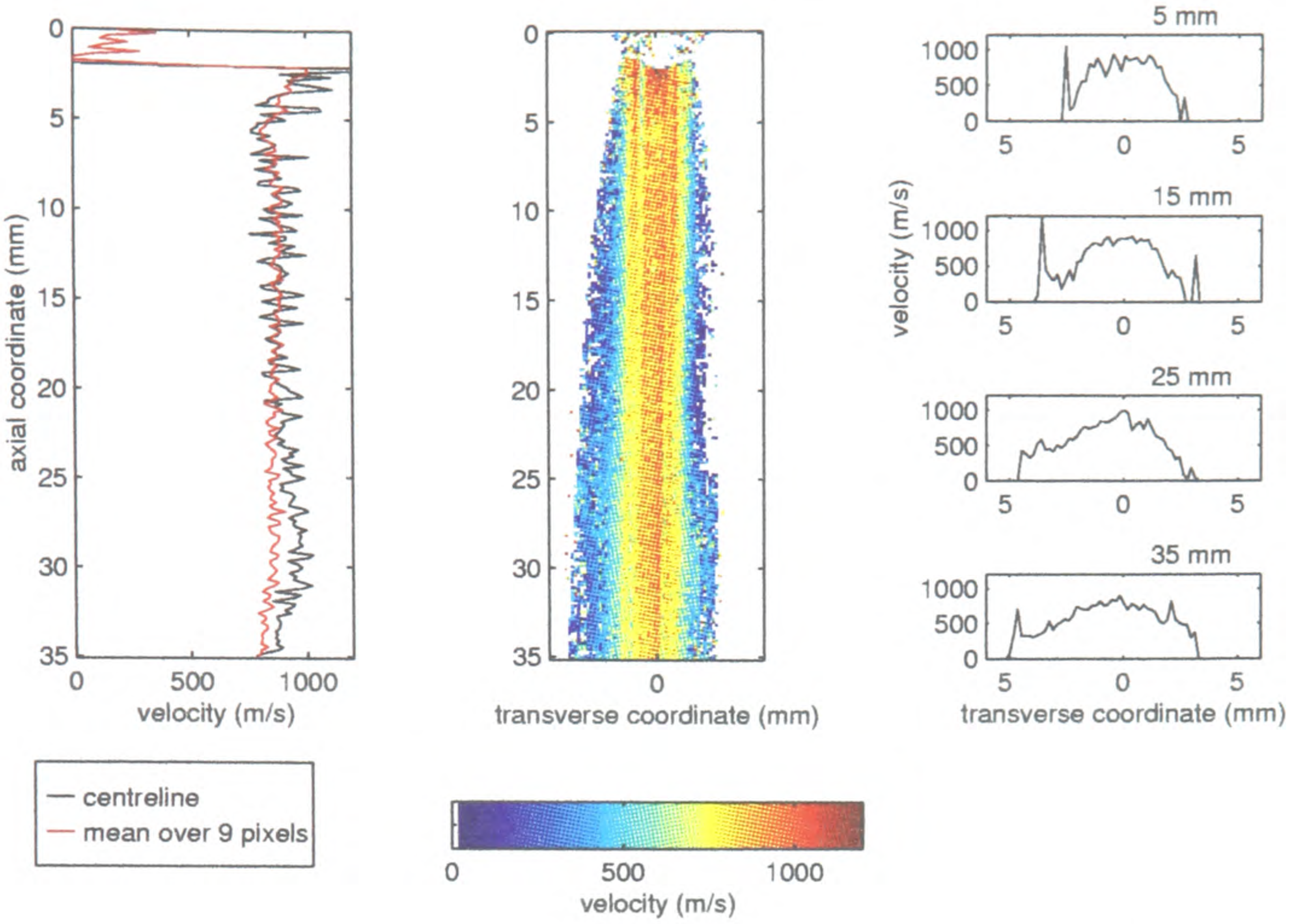


Figure 6.8 Measured velocity map for 3.1 mg of 75-106 μm lidocaine particles, delivered from a Mach 2.5 nozzle.

Two velocity maps measured for silica payloads are shown in ***Figure 6.9*** and ***Figure 6.10***. Exit velocities are higher than those observed in microsphere or lidocaine payloads. These maps differ from results for other particle materials in one striking feature — they contain a region of low velocity around the jet centreline, surrounded by comparatively high velocities. The region is approximately 1 mm wide and is located on the jet centreline, beginning approximately 5 mm from the nozzle exit plane, and extending downstream over approximately 2 mm. In the case shown in ***Figure 6.9***, which was chosen as an extreme illustration of the phenomenon, the minimum velocity in the region (in a median-filtered version of the data) is 680 m/s, occurring at an axial location where the maximum velocity is approximately 1160 m/s. In ***Figure 6.10***, a more typical case, the central minimum velocity is 810 m/s, with an outlying maximum of 990 m/s. The most pronounced low velocity features occur when payload mass is small, and the weakest are observed with large payloads.

Central regions of low velocity have also occasionally been observed in lidocaine jets from the contoured nozzle, but were later found to be spurious results of an imaging aberration. However, examination of the raw flow-field images for silica, in the light of experience gained with the problematic lidocaine measurements, has shown that the same aberration is certainly not responsible for the low-velocity features seen in silica flow-fields.

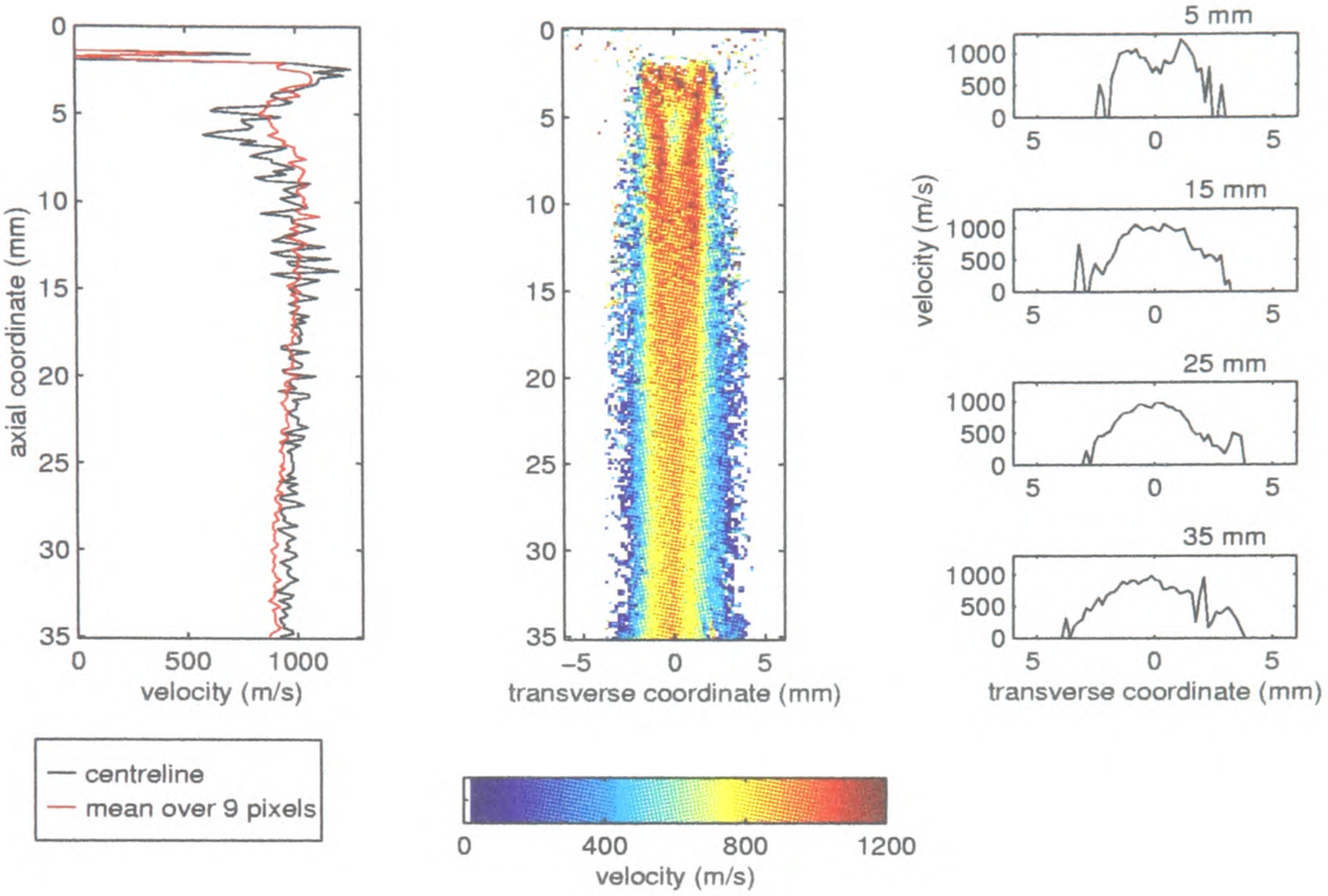


Figure 6.9 Measured velocity map for a payload of 1.0 mg of silica particles, delivered from a Mach 2.5 nozzle.

If it is not spurious, the observed low velocity feature could be explained as part of the structure of an overexpanded supersonic jet [11]. When a supersonic nozzle exit flow is at a lower pressure than the exterior gas, it is compressed to the exit pressure through oblique shocks which extend downstream and inwards from the nozzle rim. Gas decelerates on passing through these shocks, and accelerates again further downstream in the expansion fans which result from reflection of the oblique shocks at the jet boundary. The low velocity zone observed in DGV results could be the region between the oblique shocks and the expansion fans in an overexpanded jet. This interpretation is consistent with the observation that the measured low velocity feature becomes less pronounced when payload mass is increased. Any gas flow feature is visible to the time-integrated DGV experiment only through its effects on the particle flow-field, and only for as long as particles flow. Thus, with a large payload (if it is assumed that larger payloads are delivered over longer times), an evolving flow feature (such as an element of an overexpanded jet structure) would be integrated over a long period, and would effectively be blurred out. With a small payload and correspondingly short-lived particle flow, an unsteady feature would change relatively little over the image integration period, and would be well-defined in the final images. The assumption that the duration of particle flow increases with payload mass has not been tested under these experimental conditions. However, optical particle detection tests, described in §4.3, showed that this was the case for the smallest polystyrene spheres in the conical nozzle. If the observed low-velocity zone is indeed a genuine flow feature, it follows that the silica particles make excellent flow tracing particles, as their nominal mean size of 16 nm suggests.

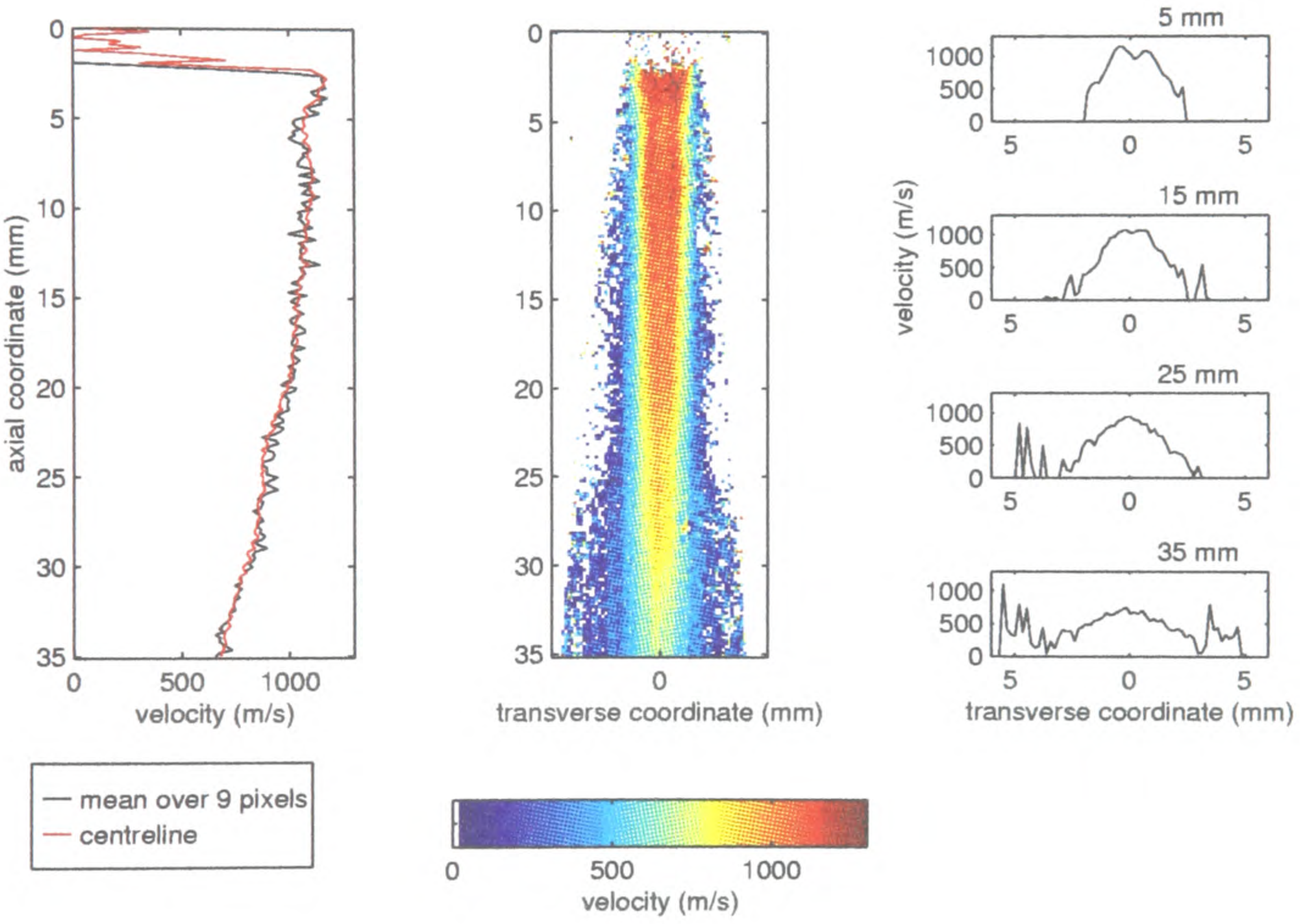


Figure 6.10 Measured velocity map for a payload of 5.0 mg of silica, delivered from a Mach 2.5 nozzle.

6.2.3 Uncertainty Estimate

Using the method laid out in §5.3, the measurement uncertainty has been estimated for a velocity measurement of the flow from a contoured nozzle with a payload of 3.0 mg of 15.5 μm polystyrene microspheres. The computed uncertainty distribution is shown as a colour map in ***Figure 6.11*** alongside the corresponding velocity field. The contributions of the three main categories of error — iodine cell transmission profile, misalignment and shot noise — are shown in ***Figure 6.12***.

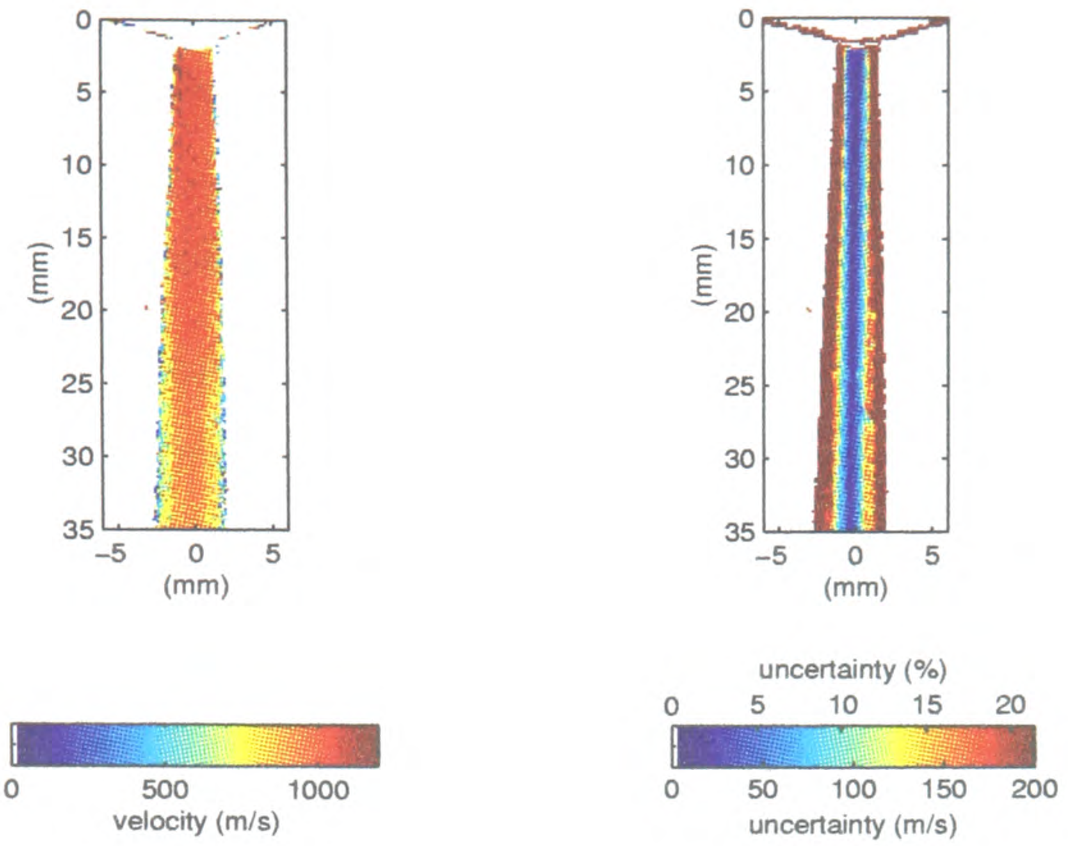


Figure 6.11 Measured velocity field (*left*) for a 3.0 mg payload of 15.5 μm microsphere delivered from the Mach 2.5 contoured nozzle, with the estimated distribution of uncertainty (*right*), expressed in terms of absolute velocity and as a percentage of the maximum velocity measured in this flow.

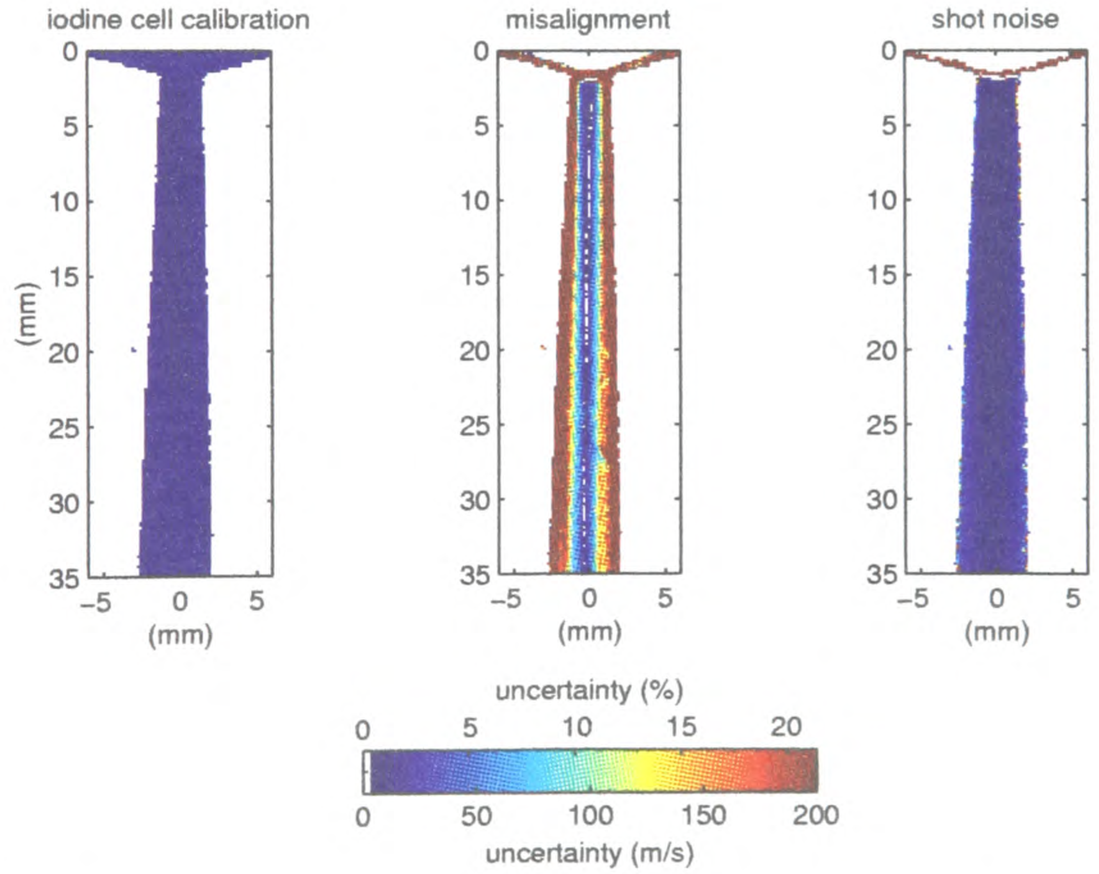


Figure 6.12 The uncertainty distribution in a velocity measurement of flow from the contoured nozzle, broken down into contributions in three categories. Uncertainty is expressed in terms of absolute velocity and as a percentage of the maximum velocity measured in this flow.

On the nozzle centreline, the uncertainty is less than 30 m/s, or 3% of the maximum velocity. Away from the centreline, uncertainty rises rapidly, exceeding 200 m/s on the periphery of the measurement region. The breakdown of error sources in ***Figure 6.12*** shows that the uncertainty off the centreline is due almost entirely to misalignment, which is accentuated by the high image intensity gradients associated with the jet boundary.

6.2.4 Effects of Payload Configuration

The influence of particle type, particle size and payload mass on the particle velocity field is summarised in ***Figures 6.13*** to ***6.15***. They show the exit velocities measured for polystyrene microspheres, lidocaine and silica, respectively, for payload masses ranging from 1 mg to 5 mg. For the microspheres and for lidocaine, the velocity values shown are the median over a 1.2x1.2 mm region (9x9 pixels) located on the centreline 5 mm downstream of the nozzle exit. For silica, each value plotted is the maximum velocity on a cross-section 5 mm downstream of exit in a velocity field which has been subjected to 5x5 pixel median filter. A maximum velocity, rather than a centreline velocity, was chosen as the basis of summary data for silica because of the central low-velocity feature.

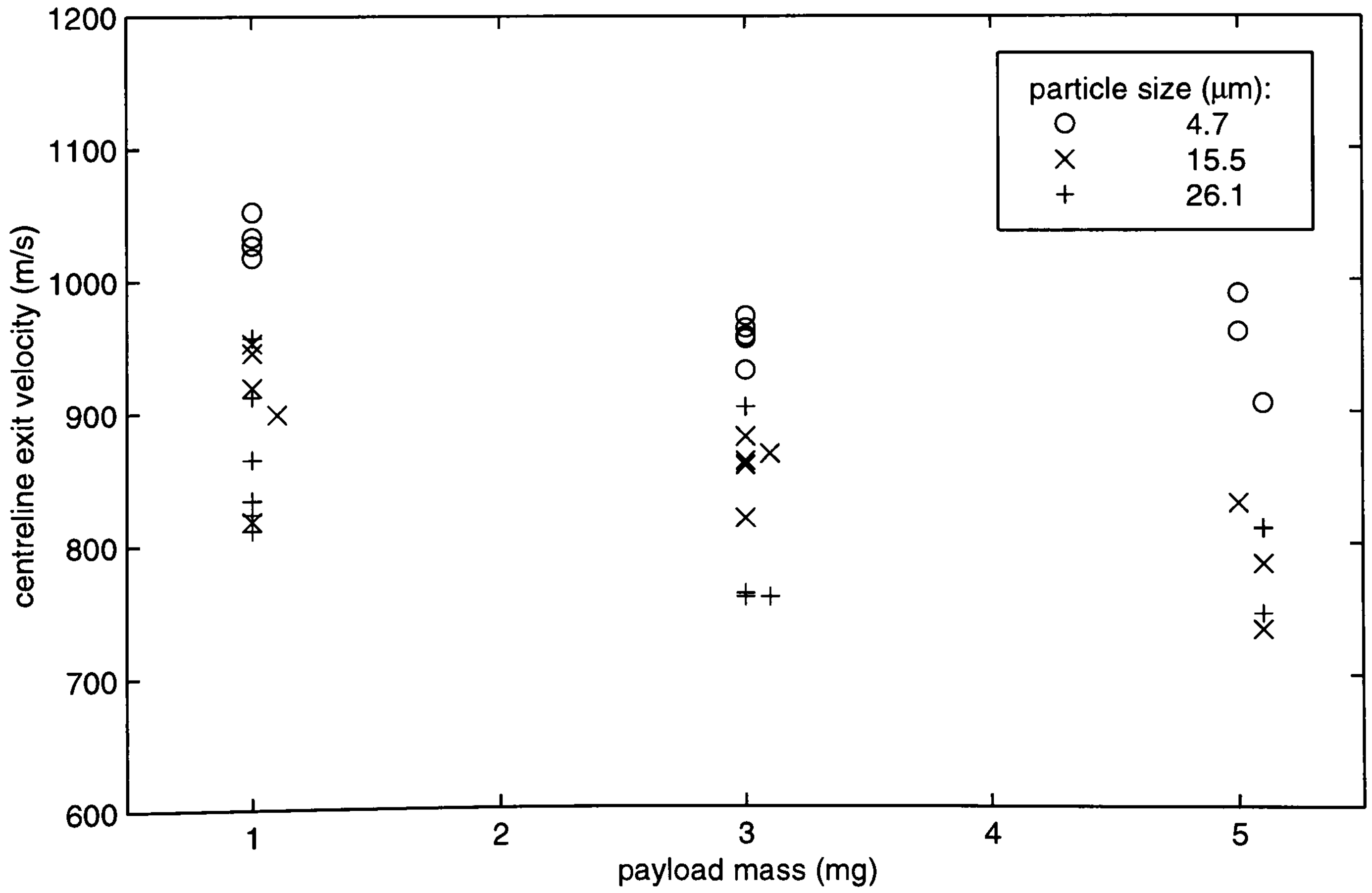


Figure 6.13 Particle velocity on the nozzle centreline 5 mm downstream of the exit, averaged over a 1.2 mm square region, for various pay loads of polymer microspheres, delivered from a contoured nozzle.

In ***Figure 6.13***, two trends are apparent in the exit velocity of microspheres. Exit velocity is higher for the 4.7 μm particles than for both grades of larger particles, and decreases as payload mass is increased. These findings are in line with a simple understanding of the gas-particle interaction. As noted in Chapter 2, the smallest particles are expected to accelerate most rapidly. Specifically, acceleration is inversely proportional to diameter, if drag coefficient is constant. Hence, the smallest particles attain the highest velocity when accelerated from rest over the distance from the drug cassette to the nozzle exit. There is no clear distinction between the velocities of 15.5 μm and 26.1 μm spheres. However, their diameters differ by a factor of only 1.7, while the 15.5 μm particles are 3.3 times larger in diameter than the 4.7 μm particles.

The measured fall-off of velocity with particle size suggests that the 15.5 μm and 26.1 μm particles, and perhaps also the 4.7 μm particles, are travelling at a velocity significantly less than that of the gas at exit. If, for example, the 15.5 μm particles were in equilibrium with surrounding gas at the nozzle exit, then no further increase in velocity would be expected when particle diameter was reduced to 4.7 μm.

The decrease in particle velocity with increasing payload mass suggests that the particle loading of the gas flow is high enough to exert a significant influence on the gas dynamics. The acceleration of the particles must be draining a large fraction of the momentum of the gas. This is to be expected, since a payload of 5 mg is a significant mass in comparison with the 50 mg or so of compressed gas stored in the device, of which only a fraction might be available to interact with the drug.

As well as providing information on the fluid mechanics of the drug delivery system, these results for polystyrene microspheres in the Mach 2.5 contoured nozzle have demonstrated that the time-integrated DGV method is sensitive to changes in the flow associated with payload mass and particle size. They therefore represent a significant confirmation of the experimental technique in this application.

Figure 6.14 is a summary of the variation of exit velocity with payload mass and nominal particle size for lidocaine payloads. The velocities range from about 640 to 870 m/s. The highest velocities observed for lidocaine are comparable to those measured for 26.1 μm polymer microspheres. There is no significant correlation between velocity and particle size or payload mass.

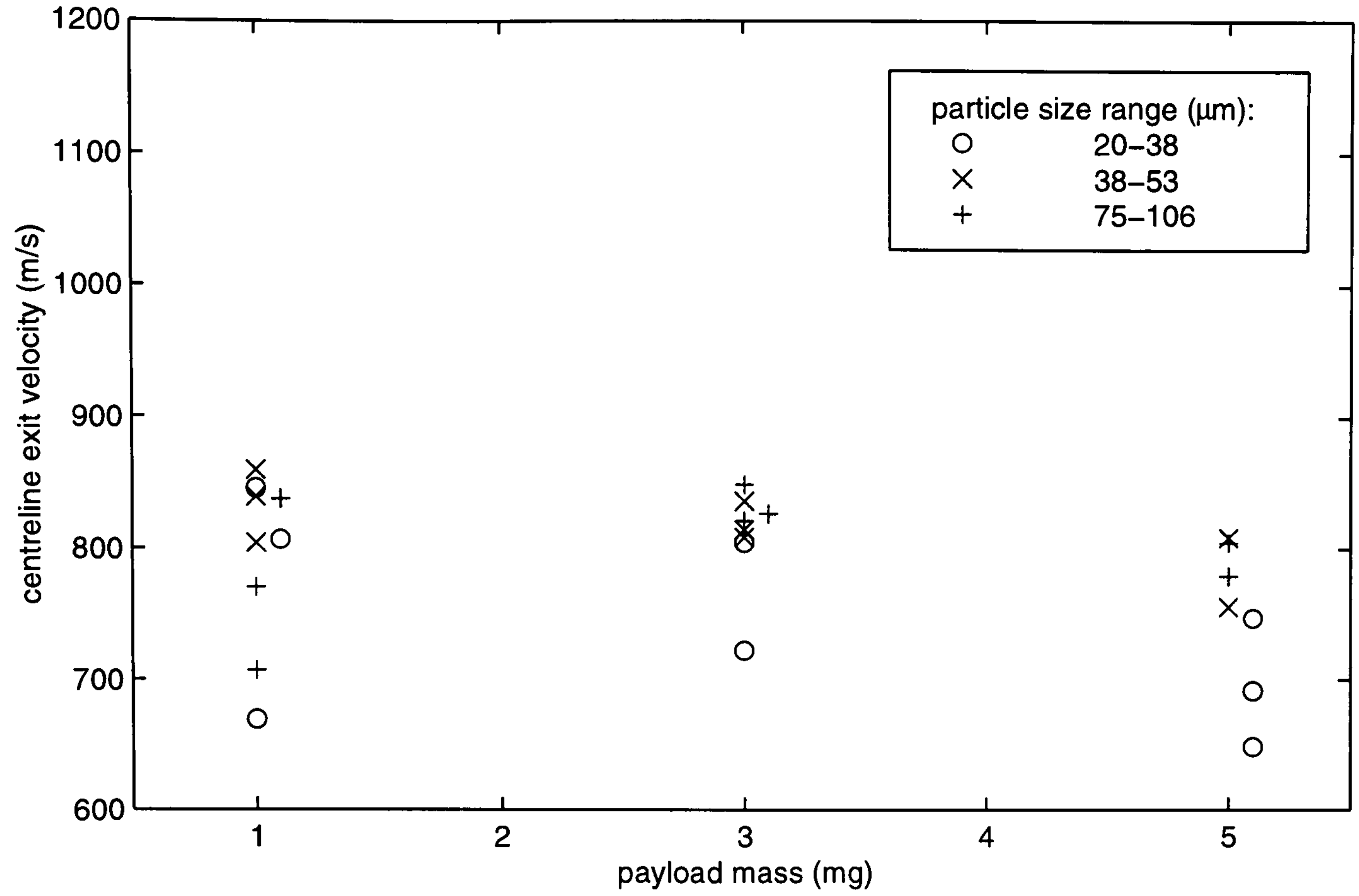


Figure 6.14 Particle velocity on the nozzle centreline 5 mm downstream of the exit, averaged over a 1.2 mm square region, for various payloads of lidocaine, delivered through a contoured nozzle.

The fact that velocity does not decrease with increasing particle size suggests that lidocaine particles do not preserve their initial size throughout device operation, but break up in the flow. The break-up of particle agglomerates in shock waves has been documented by researchers in the field of aerosol generation [58,59]. This phenomenon may also underlie the greater spreading of lidocaine particle jets, in comparison with microsphere jets. Fragmentation could introduce very fine particles (significantly less than 4.7 μm in size) to the flow, which could respond to the gas acceleration quickly enough to become involved in mixing between the jet and the ambient air.

This logic may seem to contradict the observation that lidocaine velocities are on a par with the velocities of the largest microspheres. However, the non-uniform particle size of lidocaine may help to resolve this discrepancy. As explained in §5.5.4, time-integrated DGV measurements are biased towards the largest and slowest particles in the flow. Also, since many lidocaine particles are long and slender, their largest dimension may be considerably larger than the nominal size (which is based on sieve mesh size). Thus, although extremely small particles may make their presence felt through their activity in the mixing layer, the measurements in the bulk of the flow-field will be dominated by any large, slow particles which emerge intact from the nozzle.

The observation that velocity does not fall off significantly with increasing payload size is more difficult to explain. It suggests that in the case of lidocaine, the interaction between particles and gas is too complex to be understood in terms of a simple momentum balance, perhaps because of the influence of the wide particle size distribution. Again, measurement bias based on particle size and velocity may be clouding the true velocities. It was because of these areas of uncertainty with lidocaine, aggravated by batch-to-batch variations, that polystyrene microspheres were originally adopted as a drug model.

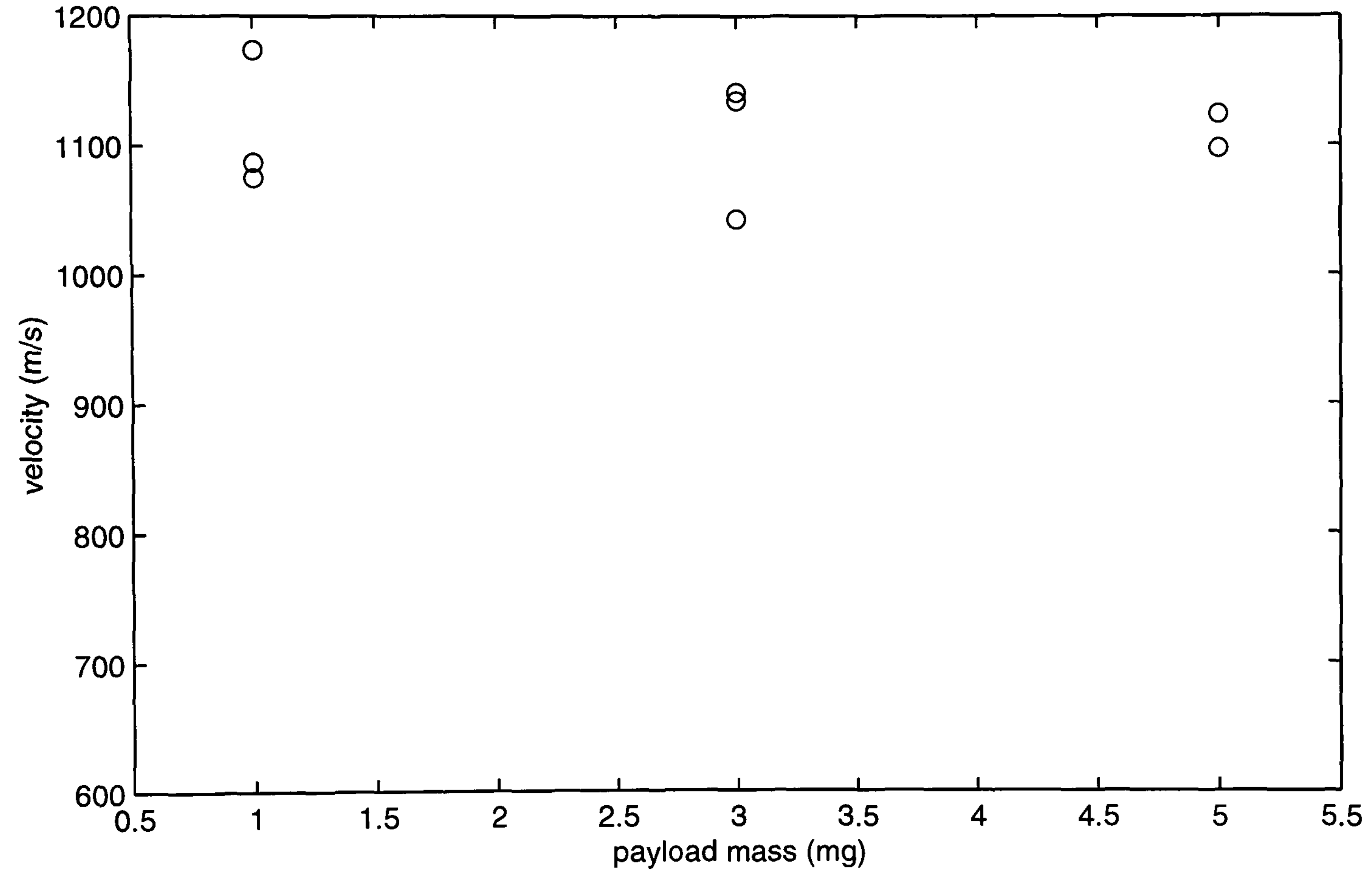


Figure 6.15 Maximum particle velocity on a transverse section 5 mm downstream of the nozzle exit plane, in a velocity map processed with a 5x5 pixel median filter, for various payloads of silica delivered from a contoured nozzle.

A summary of the maximum (as opposed to centreline) exit velocities measured for silica particles is presented in ***Figure 6.15***. The velocities lie between 1040 and 1170 m/s, and are apparently independent of payload mass. The highest velocities are considerably higher than any velocities measured with other particle types, suggesting again that silica follows the gas velocity more effectively. These results indicate that helium can attain velocities of at least 1200 m/s at the nozzle exit.

6.3 Time-Integrated Measurements: Conical Nozzle

The programme of time-integrated DGV on flow from the conical nozzle covered a range of payload configurations, in a manner similar to the experiments on the Mach 2.5 contoured nozzle. Typical results are presented in §6.3.1, using the format described in §6.2.1 for the contoured nozzle. An estimate of measurement uncertainty, and a discussion of the variation of exit velocity with payload mass and particle size, are presented in §6.3.2 and §6.3.3, respectively.

6.3.1 Representative Results

A representative range of velocity maps measured for microsphere payloads delivered from a conical nozzle is shown in ***Figures 6.16*** to ***6.21***. The low intensity of scattered light in flows from the conical nozzle, and its implications for measurement quality, have already been mentioned in §5.4.3. The velocity map in ***Figure 6.16*** is derived from raw images which were amongst the brightest; in the measurement shown in ***Figure 6.20***, image intensity was barely acceptable. Raw images with maximum intensity below a certain threshold were rejected.

For microspheres, the exit velocities range from approximately 300 m/s to 500 m/s. In most cases, velocity is highest on the centreline close to the nozzle exit, and falls off gently with downstream distance from the nozzle, or remains roughly constant for some distance before beginning to decay. Velocity varies smoothly across the width of the jet, with a maximum in the centre. The uniform central core, characteristic of jets from the Mach 2.5 nozzle, is not present in these flows. Instead, flow varies smoothly and continuously across the nozzle exit.

The consistency between ***Figures 6.16*** and ***6.17***, measurements of the velocity of 3.0 mg payloads of 4.7 μm spheres, is typical of the repeatability observed in the velocity of microspheres delivered from the conical nozzle. Like the measurements of contoured nozzle exit velocity, these data are scattered over ranges of 40 to 120 m/s. Expressed as a percentage of the exit velocity, the greatest scatter is 32%.

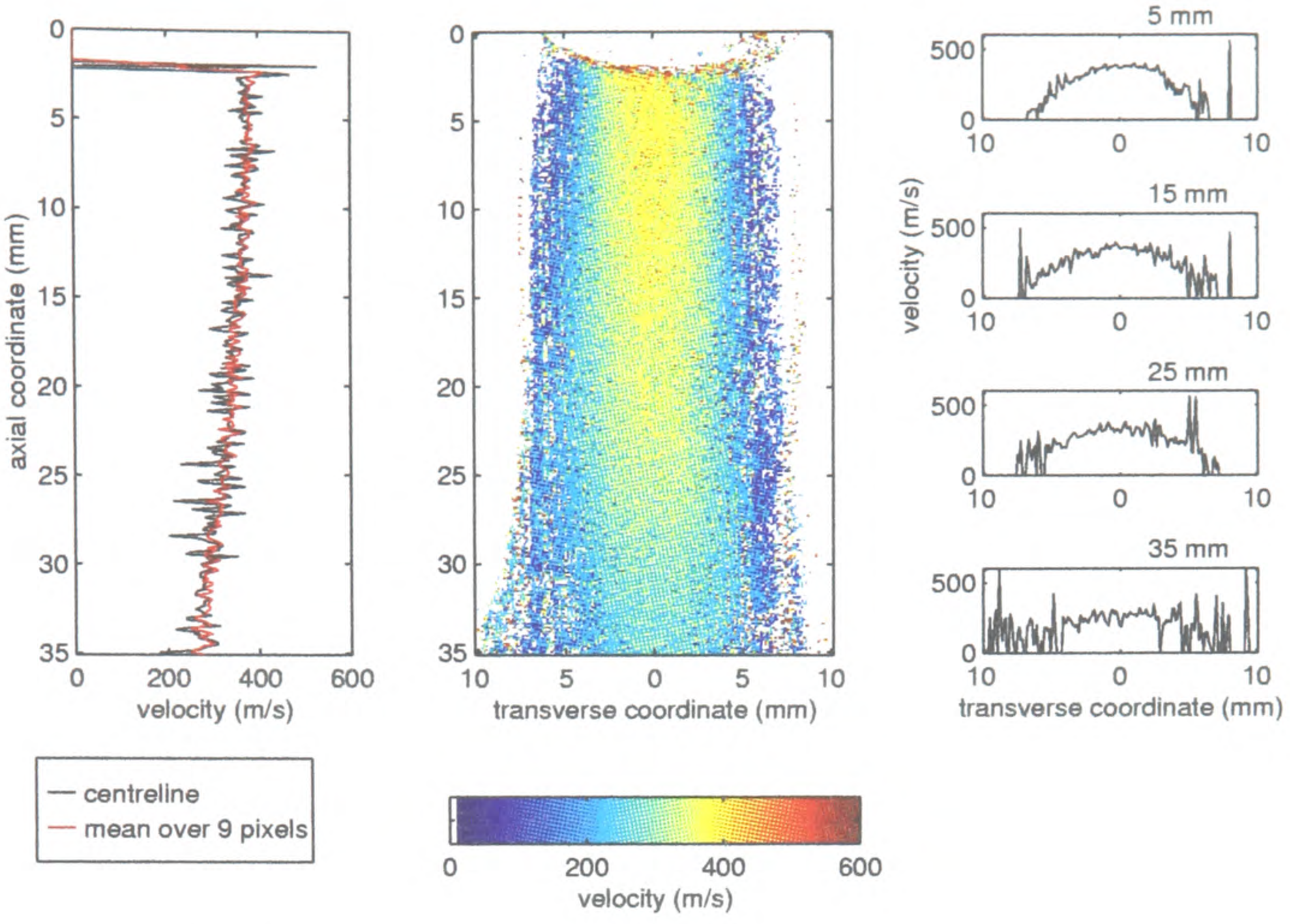


Figure 6.16 Measured velocity map for 3.0 mg of 4.7 µm polymer microspheres, delivered from a conical nozzle.

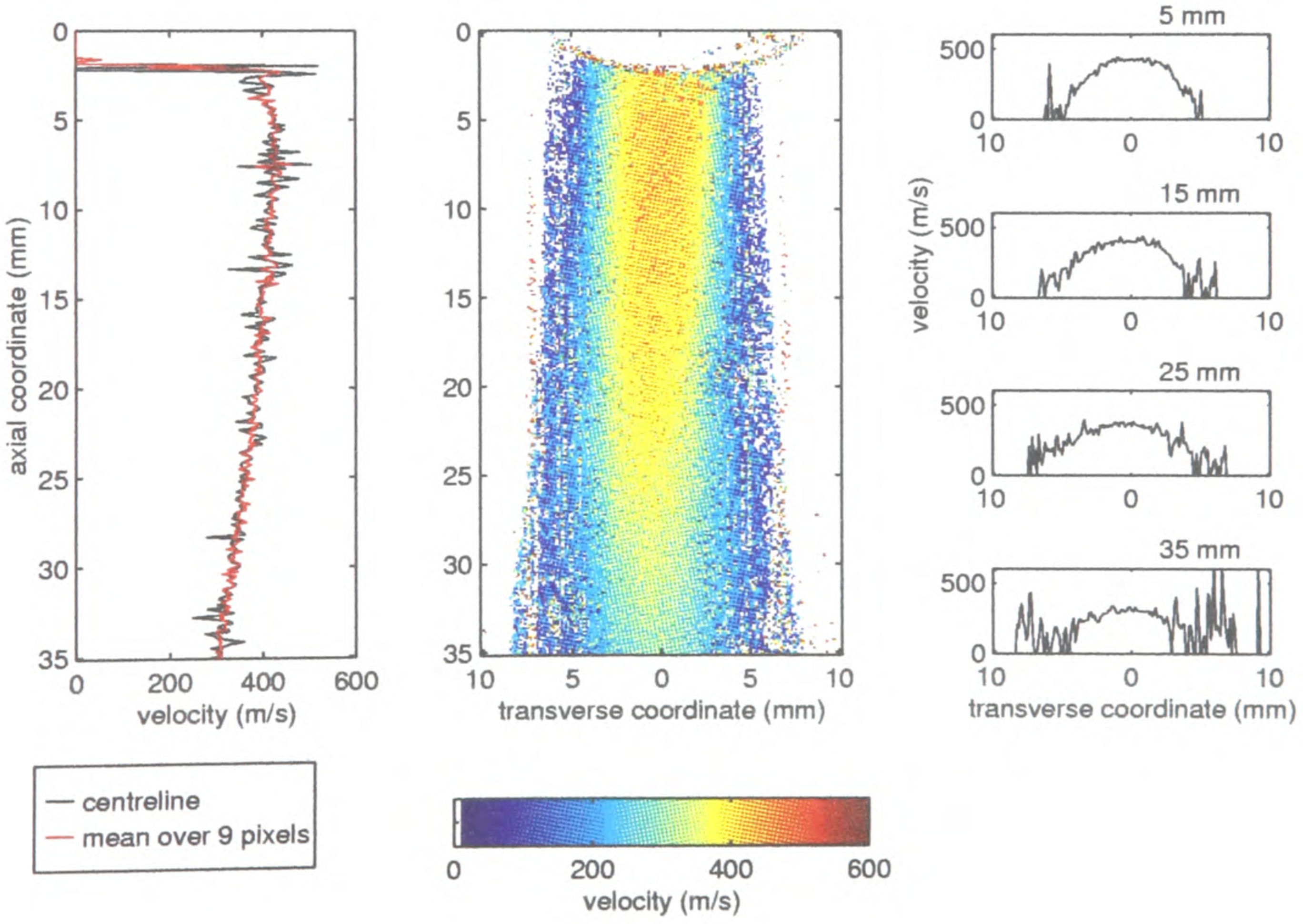


Figure 6.17 Measured velocity map for 3.0 mg of 4.7 µm polymer microspheres, delivered from a conical nozzle.

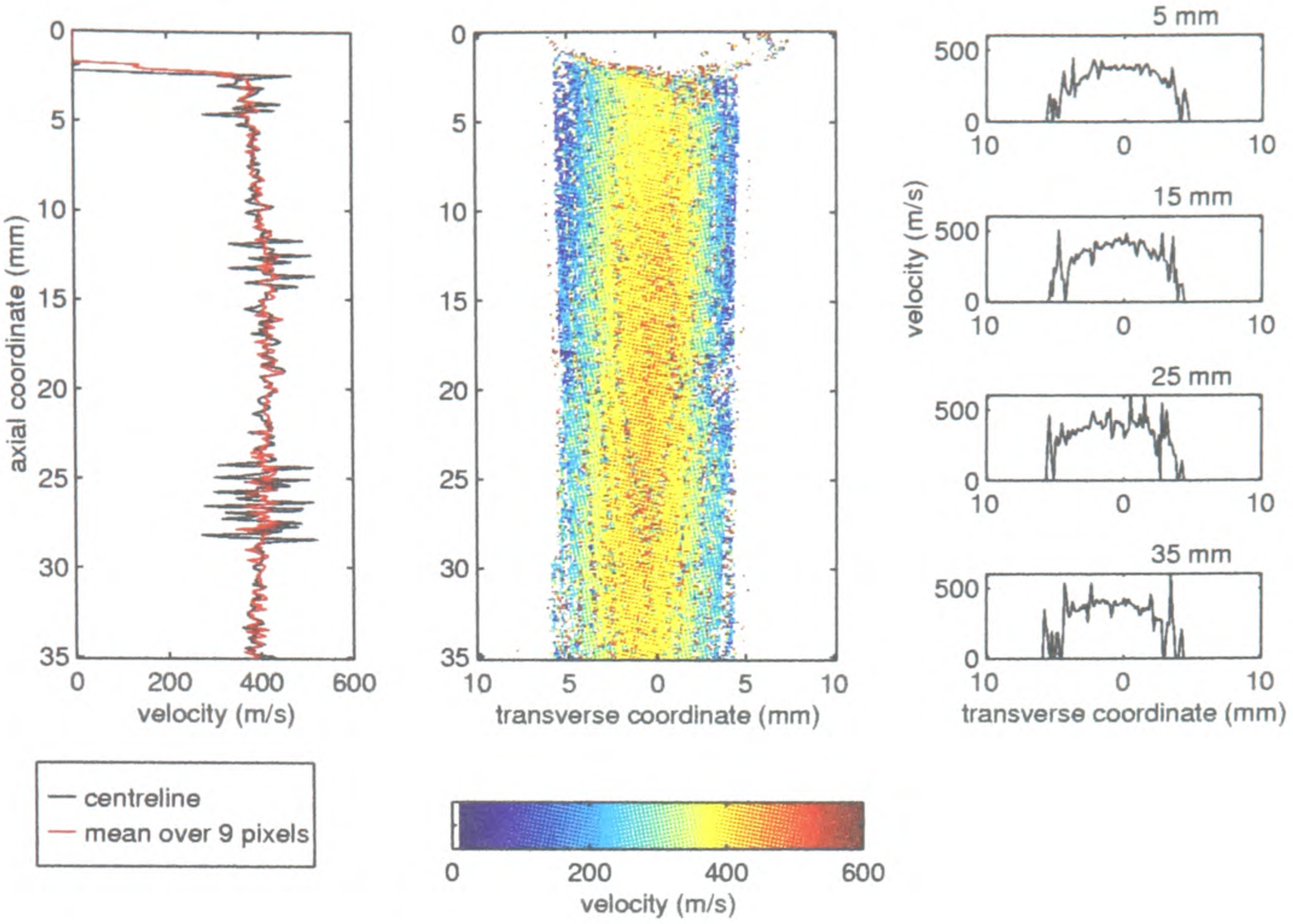


Figure 6.18 Measured velocity map for 1.0 mg of 4.7 μm polymer microspheres, delivered from a conical nozzle.

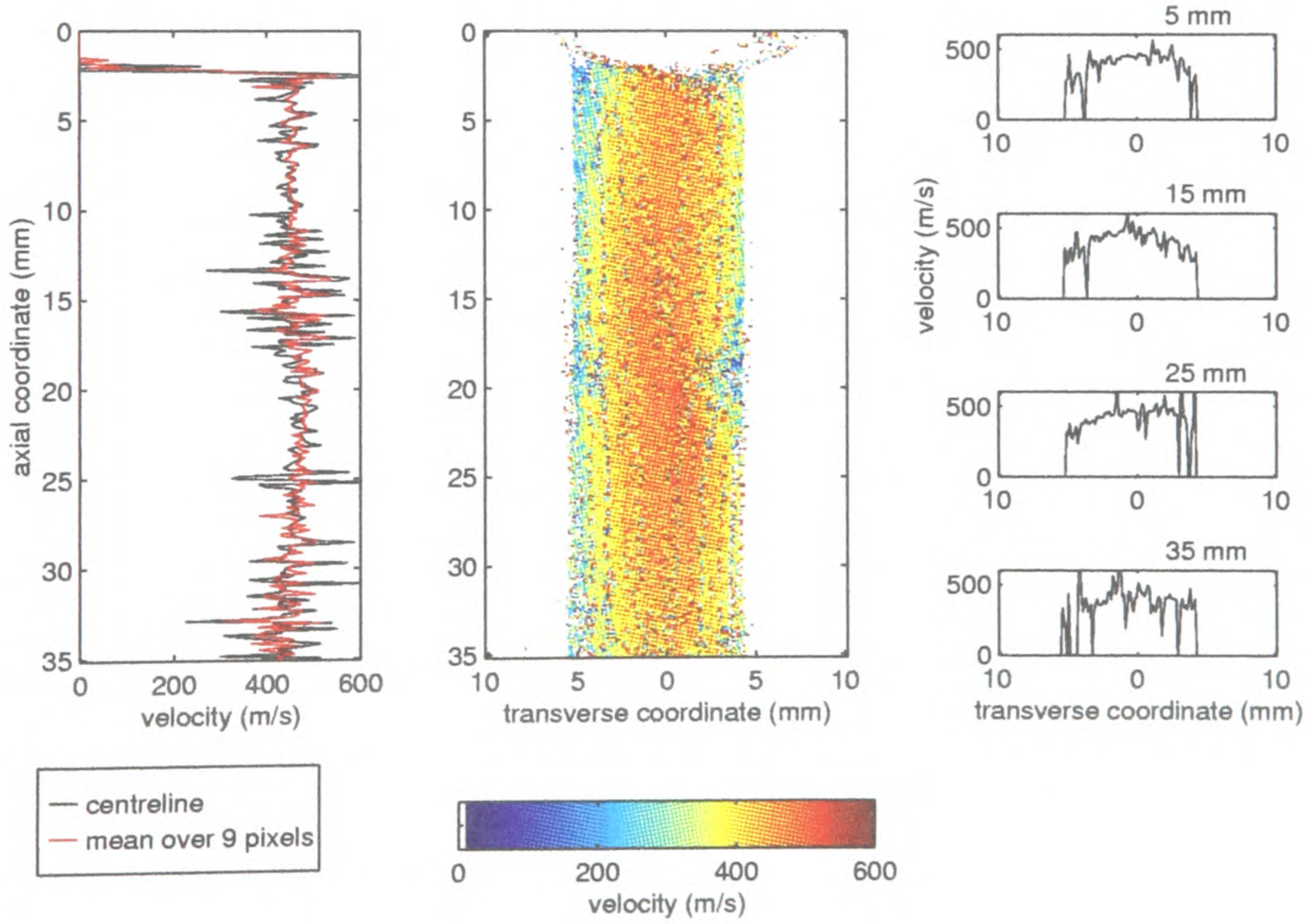


Figure 6.19 Measured velocity map for 5.0 mg of 15.5 μm polymer microspheres, delivered from a conical nozzle.

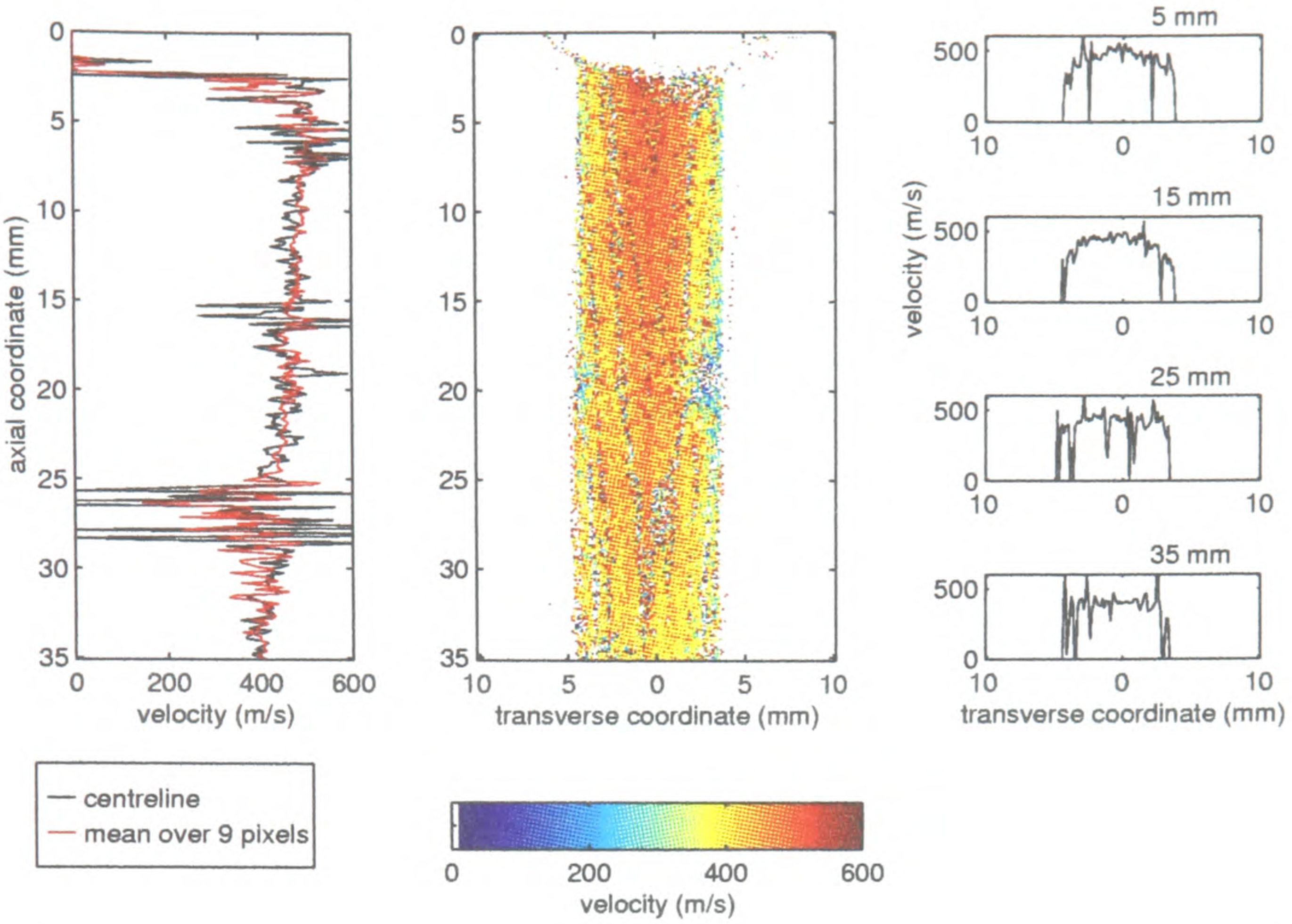


Figure 6.20 Measured velocity map for 5.0 mg of 26.1 μm polymer microspheres, delivered from a conical nozzle.

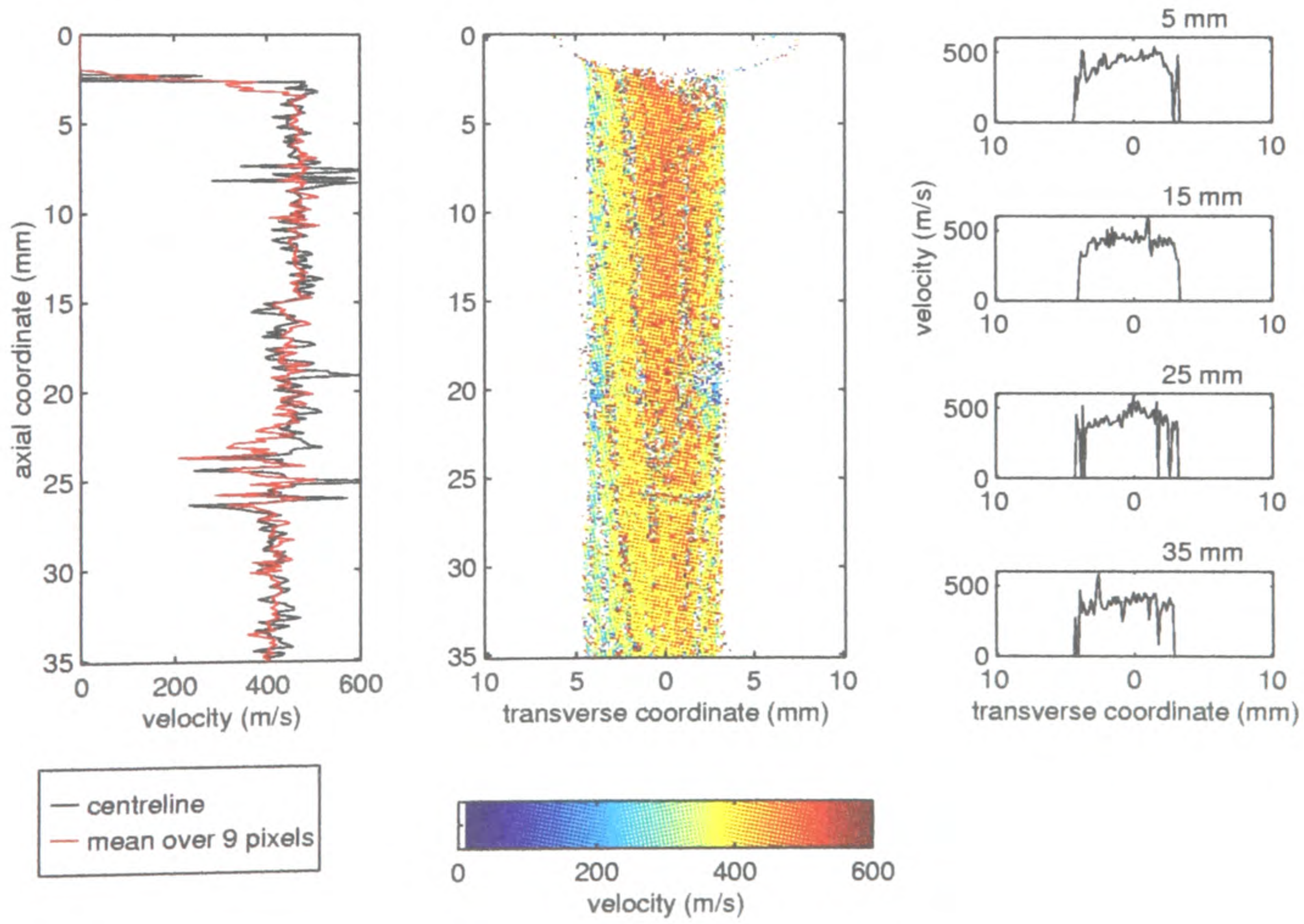


Figure 6.21 Measured velocity map for 3.0 mg of 26.1 μm polymer microspheres, delivered from a conical nozzle.

In one of the cases shown, a measurement for 1 mg of 4.7 µm microspheres (***Figure 6.18***), the maximum velocity occurs well downstream of the nozzle exit. This behaviour is not common to repeats at this operating condition. It is probably due to a particularly severe instance of light scattering by material other than payload particles, a phenomenon which is discussed at length in §5.5.1. ***Figure 5.28*** in that section demonstrates that this spurious signal is strongest in the neighbourhood of the nozzle exit, and that its result is a depression of measured velocity in that area.

An unusual form of noise appears in the velocity maps based on relatively dim images, such as the map shown in ***Figure 6.20***. This noise is restricted to long, narrow bands, no wider than two or three pixels. Examination of the raw data associated with these velocity maps shows that the noise originates in the course of the process of warping and interpolation in both the reference and discriminated images, and that the bands on which it is distributed coincide with contours of constant intensity in the warped images. The causes of this noise are not currently known. However, it appears only in velocity fields based on very dark images and apparently does not have any global effect on the measurement results.

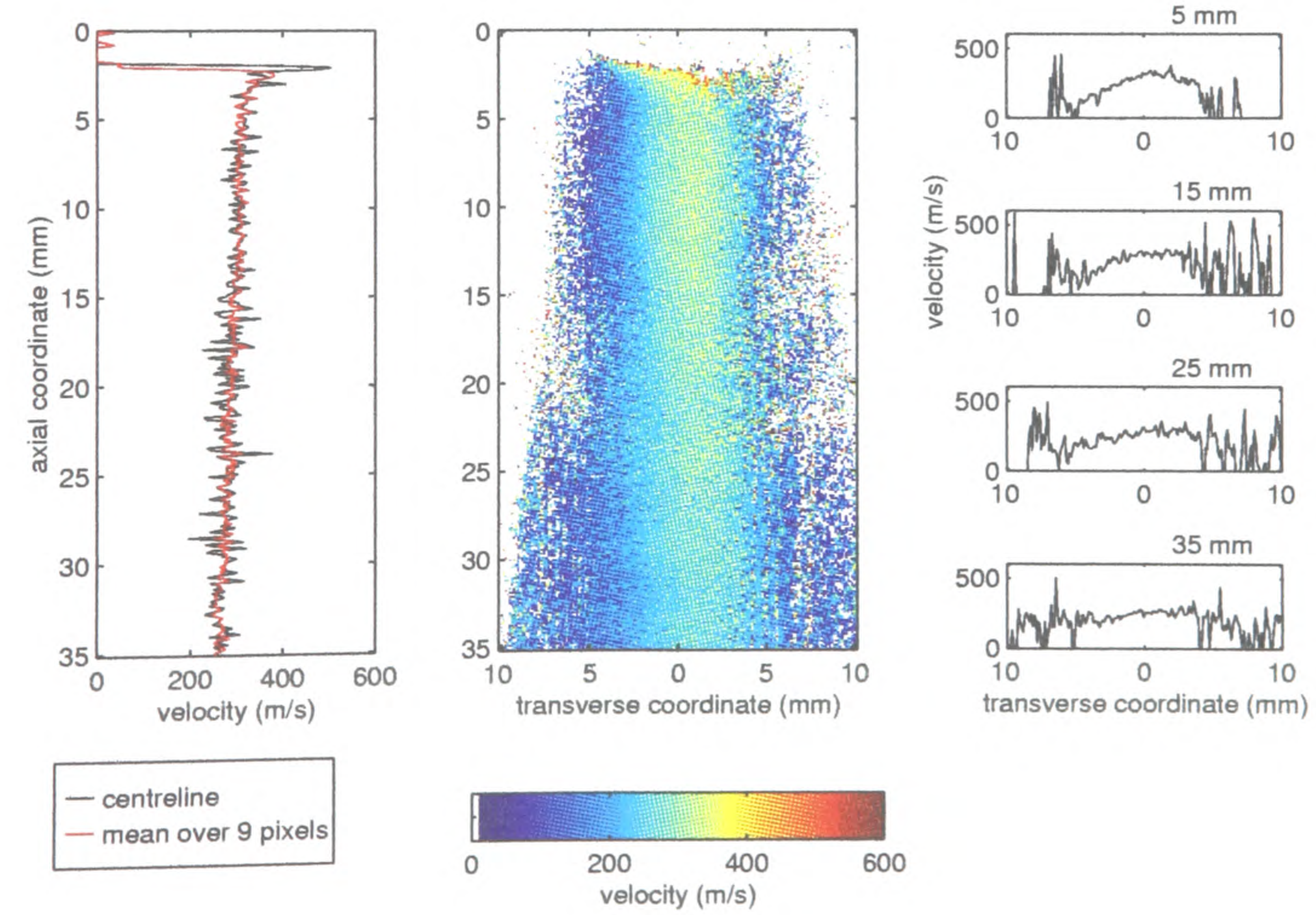


Figure 6.22 Velocity map measured for 5.0 mg of 20-38 µm lidocaine, delivered from a conical nozzle.

Examples of velocity maps measured for lidocaine are presented in ***Figures 6.22*** and ***6.23***. They are similar in overall form to the fields measured for microsphere payloads. It can be seen in ***Figure 6.22*** that the jet spreads to a greater width than the microsphere jets, again suggesting the presence of fine particles which accelerate quickly enough to follow the local gas flow to some extent into a mixing layer at the jet boundary.

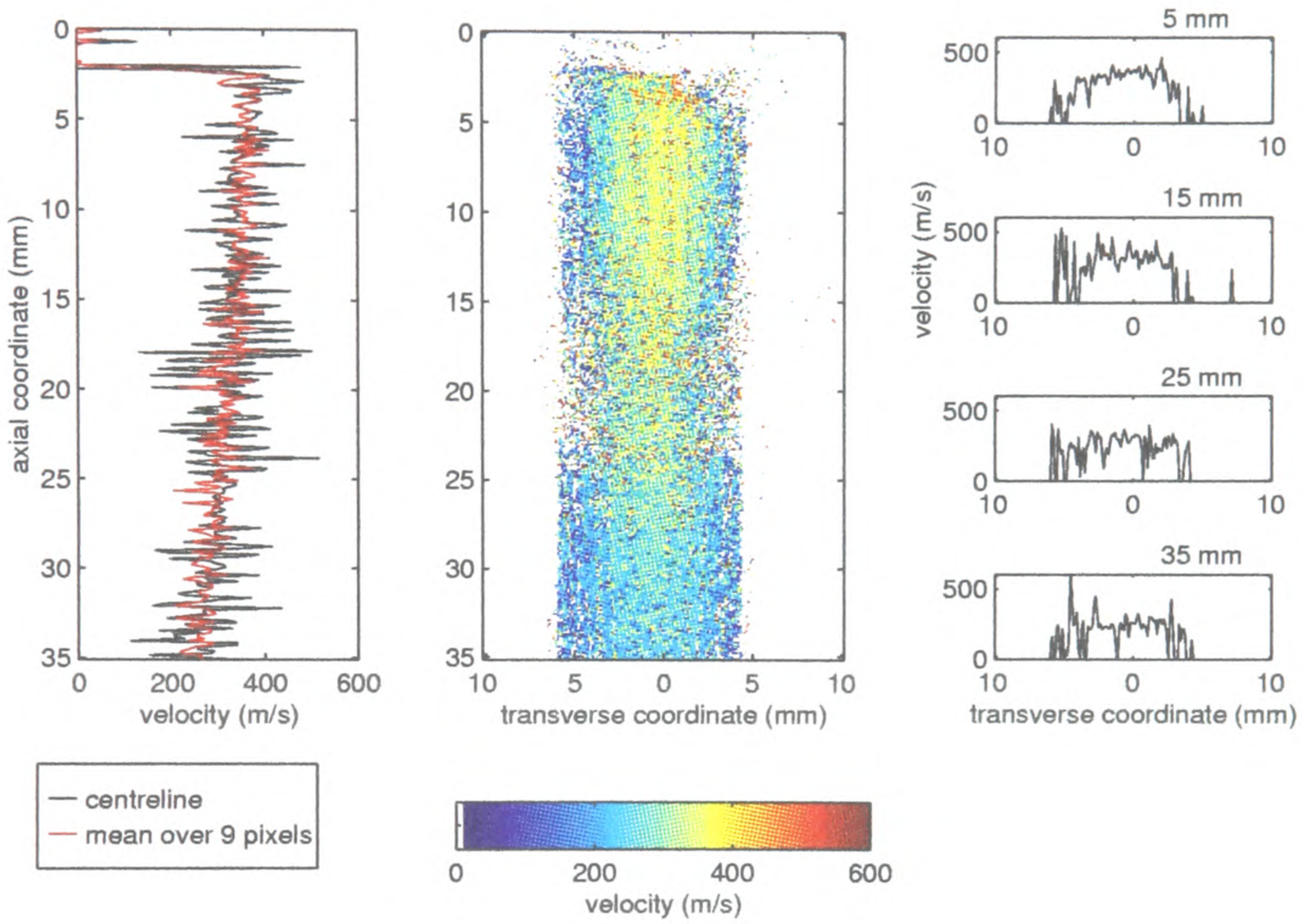


Figure 6.23 Velocity map measured for 1.0 mg of 75-106 μm lidocaine, delivered from a conical nozzle.

6.3.2 Uncertainty Estimate

Uncertainty in velocity measurement of flows from the conical nozzle was calculated by the method laid out in §5.3 for a representative case, in which the device was loaded with 5.0 mg of 38–53 μm lidocaine. The computed uncertainty distribution is shown as a colour map in ***Figure 6.24*** alongside the corresponding velocity field. The contributions of iodine cell calibration uncertainty, misalignment and shot noise to overall uncertainty are compared in ***Figure 6.25***.

The uncertainty in measured centreline velocities is approximately 15 m/s (4% of the centreline exit velocity), and is due largely to shot noise. Near the edges of the particle jet, where image intensity gradients are higher, misalignment becomes significant, and uncertainty grows to 50 m/s on the fringes of the measurement region. The relative uncertainty in centreline velocity (4%) is slightly higher than in the case of the Mach 2.5 contoured nozzle (approximately 3%), reflecting the reduced intensity and signal-to-noise ratio of the raw images. However, the uncertainties at off-centreline locations are greater for the contoured nozzle (over 20% as opposed to 15% for the conical nozzle), since misalignment error in that case is exacerbated by the smaller scale and consequent high intensity gradient of the jet images.

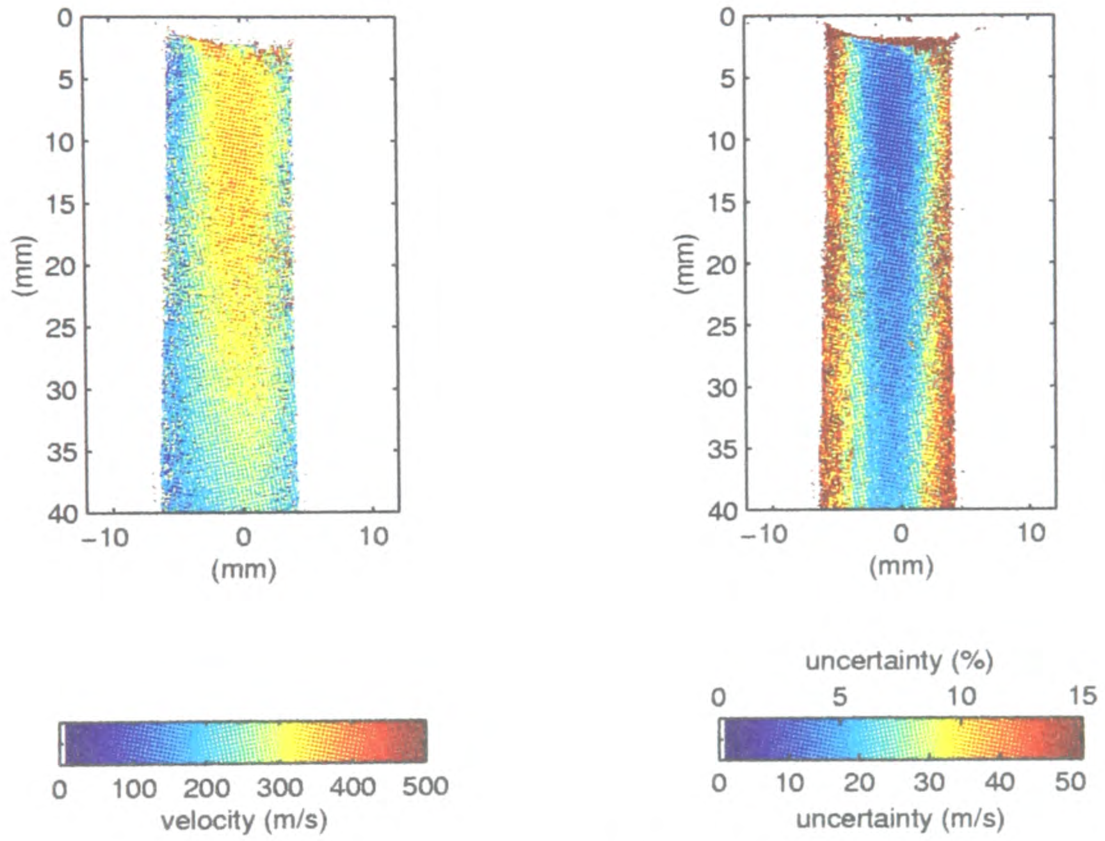


Figure 6.24 Measured velocity field for a 5.0 mg payload of 38–53 μm lidocaine delivered from the conical nozzle, with the estimated distribution of uncertainty on the right. Uncertainty is given both as an absolute velocity and as a percentage of the highest measured velocity in this flow-field.

In DGV, errors in measured frequency shift are absolute in character; that is, the absolute magnitude of the uncertainty is largely independent of the magnitude of the frequency shift being measured. However, the above comparison of uncertainties in the conical and contoured nozzle measurements, formulated in terms of percentages of peak velocity, is appropriate because the constant of proportionality relating frequency shift to axial velocity is different for the two sets of experiments. In each case, this sensitivity was adjusted (by setting the viewing angle) to accommodate the maximum expected velocity. This action effectively determined the value of the full scale velocity measurement, and scaled all velocity uncertainties in proportion to that value. The maximum measured frequency shift, on the other hand, was approximately the same in both cases.

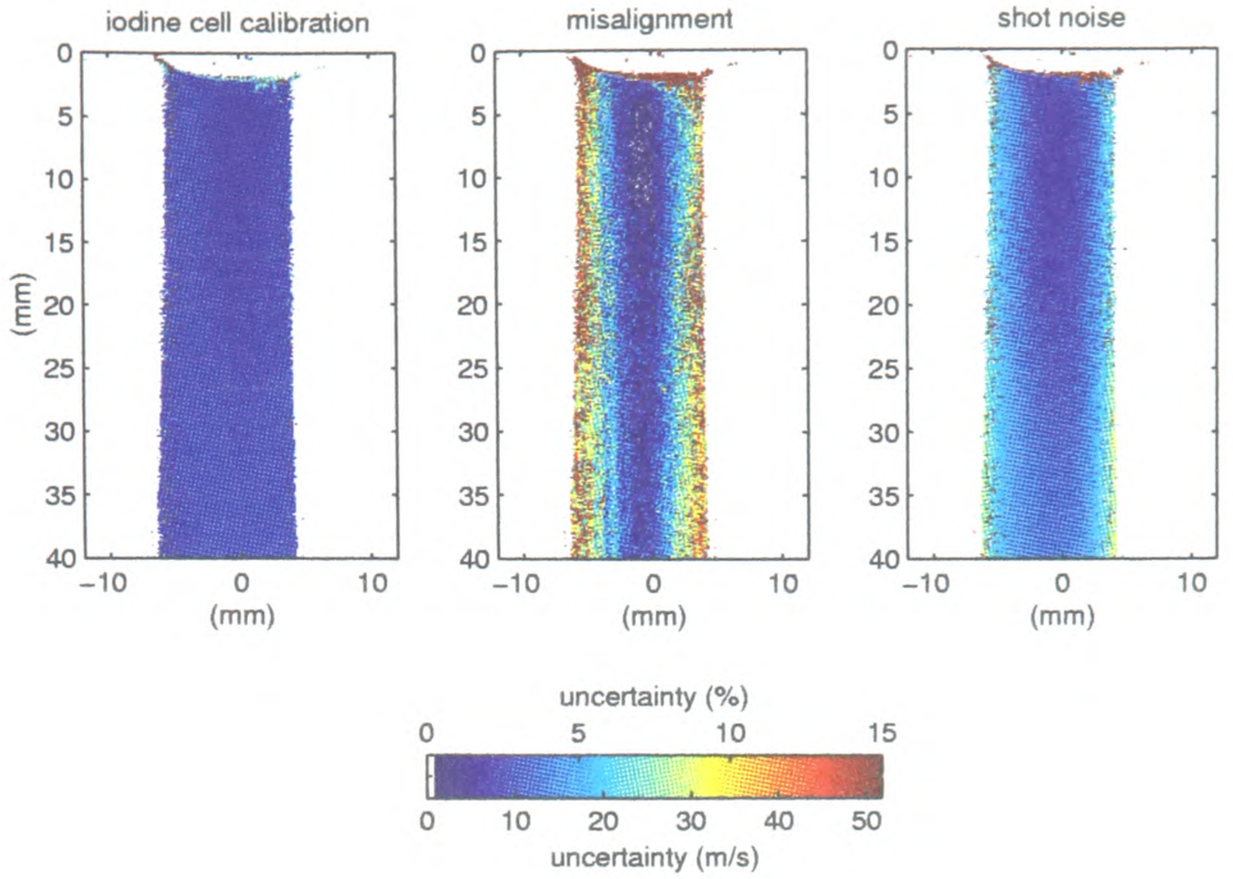


Figure 6.25 The uncertainty distribution in a velocity measurement of flow from the contoured nozzle, broken down into contributions in three categories, and expressed both as an absolute velocity and as a percentage of the highest measured velocity in this flow-field.

6.3.3 Effect of Payload Configuration

Figure 6.26 is a summary of the exit velocities measured for various microsphere payloads delivered from the conical nozzle. Specifically, these are the average velocities over a 17×17 pixel (or 2×2 mm) region centred 5 mm from the exit, on the centreline. The velocity values range from 310 m/s to 510 m/s. Velocity does not appear to vary with particle size or payload mass. This suggests that all particles, regardless of size, reach a velocity close to that of the gas before reaching the nozzle exit. It appears that particles have enough time in the conical nozzle to accelerate to the gas exit velocity, which is low in comparison with the contoured nozzle.

These data allow some discussion of the flow regimes which may be involved in particle acceleration. The estimated gas velocity behind the starting shock as it approaches the nozzle exit increases with payload mass from 195 m/s (no payload) to 225 m/s (with a 5 mg payload). The calculation is based on measured shock propagation speed, and is described in §4.4. This velocity is significantly lower than the measured particle velocities. It seems unlikely, therefore, on the basis of these measurements, that the shock tube flow is a dominant mechanism in particle delivery in this nozzle.

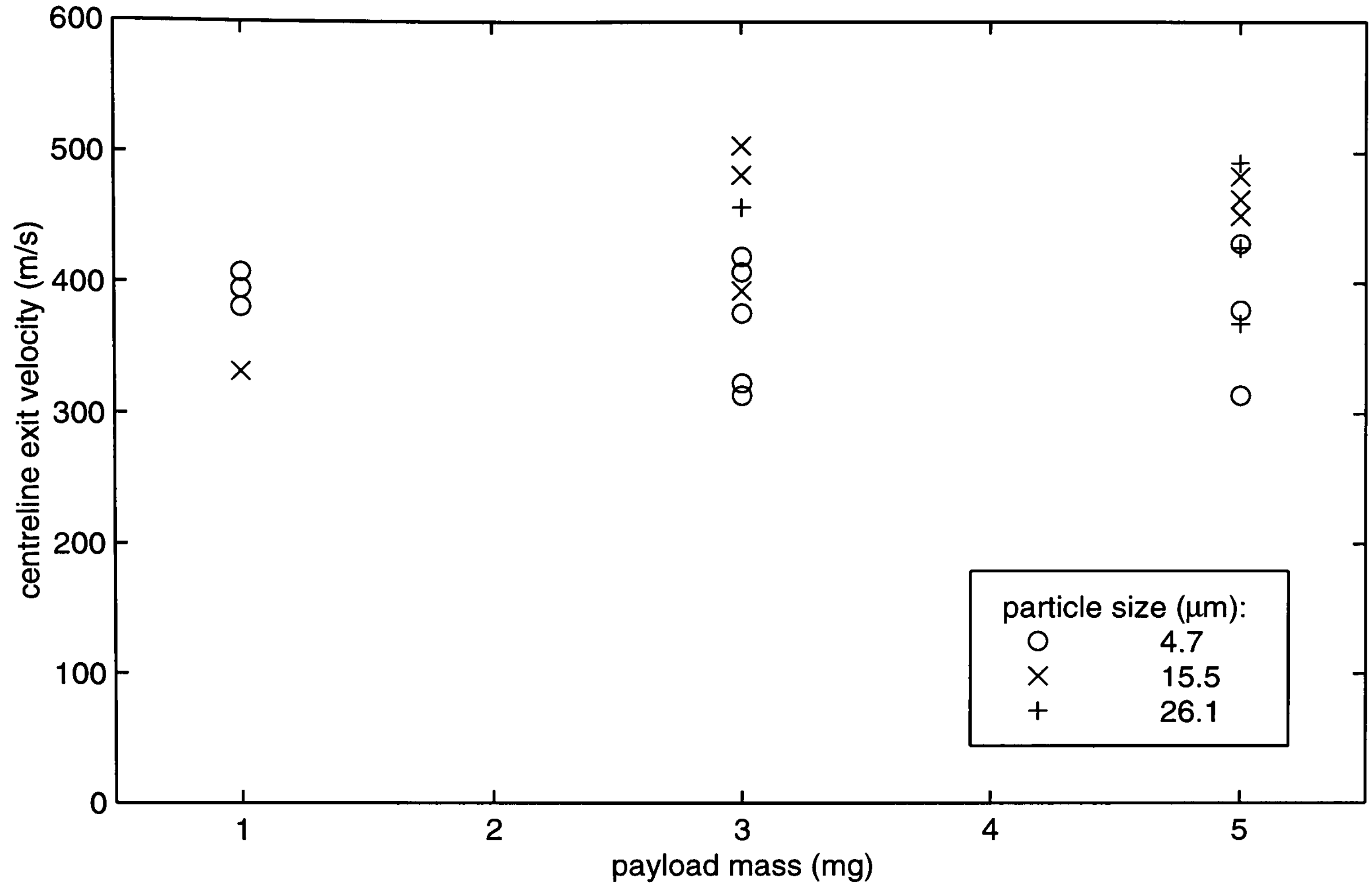


Figure 6.26 Polystyrene microsphere velocities on the centreline 5 mm downstream of the exit, averaged over a 2 mm square region for various payload masses, delivered from a conical nozzle.

On the other hand, exit velocities are a fraction of the hypothetical supersonic gas exit velocity (given in Chapter 2 as 1700 m/s), confirming the conclusion (stated in Chapter 4 on the basis of pressure measurements) that the operative total pressure upstream of the conical nozzle is not high enough to drive fully supersonic flow. Instead, it appears that the quasi-steady flow in the divergent section passes through a shock wave or a system of shocks and emerges from the nozzle in a subsonic state.

A summary of the measured exit velocity of lidocaine particles is shown in ***Figure 6.27***. Data for 1 mg payloads are very widely scattered, ranging from 150 m/s to 420 m/s. Measurements for 3 mg and 5 mg payloads range from 240 to 380 m/s. Again, velocity is not correlated with particle size or payload mass. The general observations made above, regarding the influence of the transient and quasi-steady flows on microsphere payload velocities, are also valid here. Velocities for lidocaine are slightly lower than for microspheres. This is surprising, since all the indications in the results for microspheres are that all tested particles, up to 26.1 μm in diameter, are close to equilibrium with the gas at exit.

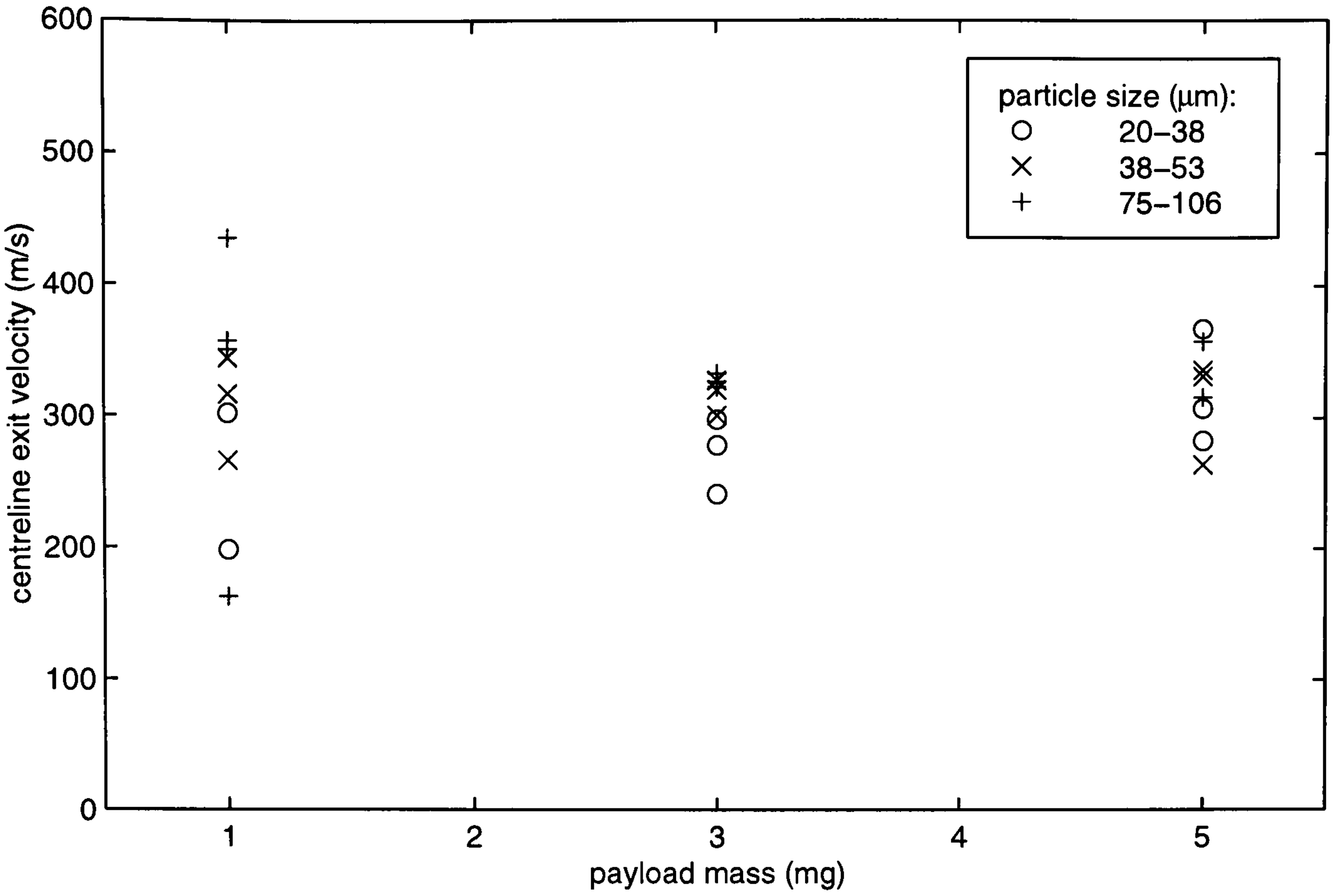


Figure 6.27 Lidocaine particle velocities on the nozzle centreline 5 mm downstream of the exit, averaged over a 2 mm square region for various payload masses, delivered from a conical nozzle.

6.4 Time-Resolved Measurements: Conical Nozzle

The most important outcome of the application of time-resolved DGV in this research was a detailed description of the unsteady drug particle velocity field at one operating condition. This data set is the central topic of this section, and is presented in full in §6.4.2 and interpreted in §6.4.3. It was acquired for a 1 mg payload of 4.7 μm particles, delivered into free air through a conical nozzle. An accompanying uncertainty estimate is presented in §6.4.4. This main set of time-resolved results is systematically compared with time-integrated results in a later section, §6.5. An independent check on the time-resolved DGV measurements, computed simply by analysing changes in the particle cloud's position from shot to shot, is described in §6.4.5. The main data set for 1 mg of 4.7 μm microspheres has been supplemented by less detailed time-resolved measurements of flows at other operating conditions of interest. These are discussed in §6.4.6. First, the format for the presentation of all time-resolved velocity maps is outlined in §6.4.1.

6.4.1 Format for Graphical Presentation of Results

The time-resolved velocity measurements in this thesis are presented in a consistent format. They are arranged in figures such as ***Figure 6.28***, consisting of two columns of six images each. The images in the left hand column are the measured velocity fields, mapped in colours which correspond to axial velocities from 0-1000 m/s. The colour code for the axial velocity component is indicated in a bar in the bottom left of the figure. Velocity data have been smoothed with a 3x3 median filter.

The images in the right hand column are the warped reference images — that is, images of the illuminated flow-field, formed in the DGV imaging system without the influence of the iodine cell, and geometrically corrected to remove perspective effects and any other distortion. The intensity levels in each reference image have been normalised to cancel the effects of variations from shot to shot in illumination pulse energy. Each pixel in a reference image represents the same region in space as the corresponding pixel in the velocity map. The reference images are included here because they correspond roughly to maps of the particle concentration distribution. As such, they are interesting in their own right, although their primary purpose is as a step in the process of velocity measurement. Each velocity map was derived from the warped reference shown image beside it (and a corresponding discriminated image), and is a representation of precisely the same instantaneous flow-field.

Both types of image are scaled and oriented in the same way, with the nozzle axis horizontal and flow from left to right. Horizontal scales, in mm, show the axial distance downstream of the nozzle exit, and vertical scales show transverse distance from the centreline. Roughly elliptical arcs are visible on either side of the exit plane in both types of image. These are images of the nozzle rim, projected onto the measurement plane at an angle of approximately 45° to the nozzle centreline. This is the direction from which the DGV imaging system viewed the scene (the important features of the measurement geometry are summarised at the end of Chapter 5). Since the nozzle used in this work was made in perspex, a region of the flow-field upstream of the nozzle exit is visible. The walls of the nozzle, where they are intersected by the light sheet, are clearly visible in the reference images. In the velocity maps, they have been artificially highlighted in thick black lines.

The text over the top left corner of each velocity map and reference image shows the time *t* at which the illuminating pulse was delivered, measured from the approximate instant of rupture of the downstream membrane. The images are laid out in order of increasing *t*.

Finally, it is emphasised again that the data were not gathered sequentially in one actuation of the drug delivery device. Every pair, comprising a velocity map and an associated reference image, was derived from a separate experiment, and represents an instant in the lifetime of one unique unsteady flow-field.

6.4.2 Velocity Field History for the Main Test Case

The results of the major series of measurements of the velocity field of 1.0 mg of 4.7 μm spheres are shown in ***Figures 6.28*** to ***6.32***. The first of the 30 instantaneous velocity measurements was taken 82 μs after bursting of the downstream membrane, when the first particles to emerge were approximately 10 mm upstream of the nozzle exit. Measurements were carried out at increasing intervals over a period of 3.8 ms, after which particle delivery almost entirely ceased.

In the early results (*t* = 82–192 μs), the progress of the first particles to emerge from the device can be observed in vivid detail. There are two distinct regions of particle behaviour. The first particles to emerge from the device have a velocity of 200 to 400 m/s in most instances. They travel in a cloud which has an approximately planar front as it approaches the nozzle exit (*t* = 82–103 μs). Immediately after passing through the exit plane, the cloud front begins to develop a convex curved shape, and widens as at advances. At 159 μs, the leading cloud is too wide to be fully contained in the 24 mm width of the measurement region. The back (upstream) face of the cloud is usually also convex in shape, so that the cloud resembles a crescent.

Behind the leading cloud, a stream of particles flows at a higher velocity of 650-800 m/s. This stream is some 10 mm wide, significantly less than the exit diameter. Velocity is low on the fringes of this jet, but there is a central core, typically 8 mm wide, of near-uniform high velocity. In contrast with the leading cloud, the width of the high-speed stream does not vary with distance from the nozzle exit. It flows into the centre of the back face of the leading cloud.

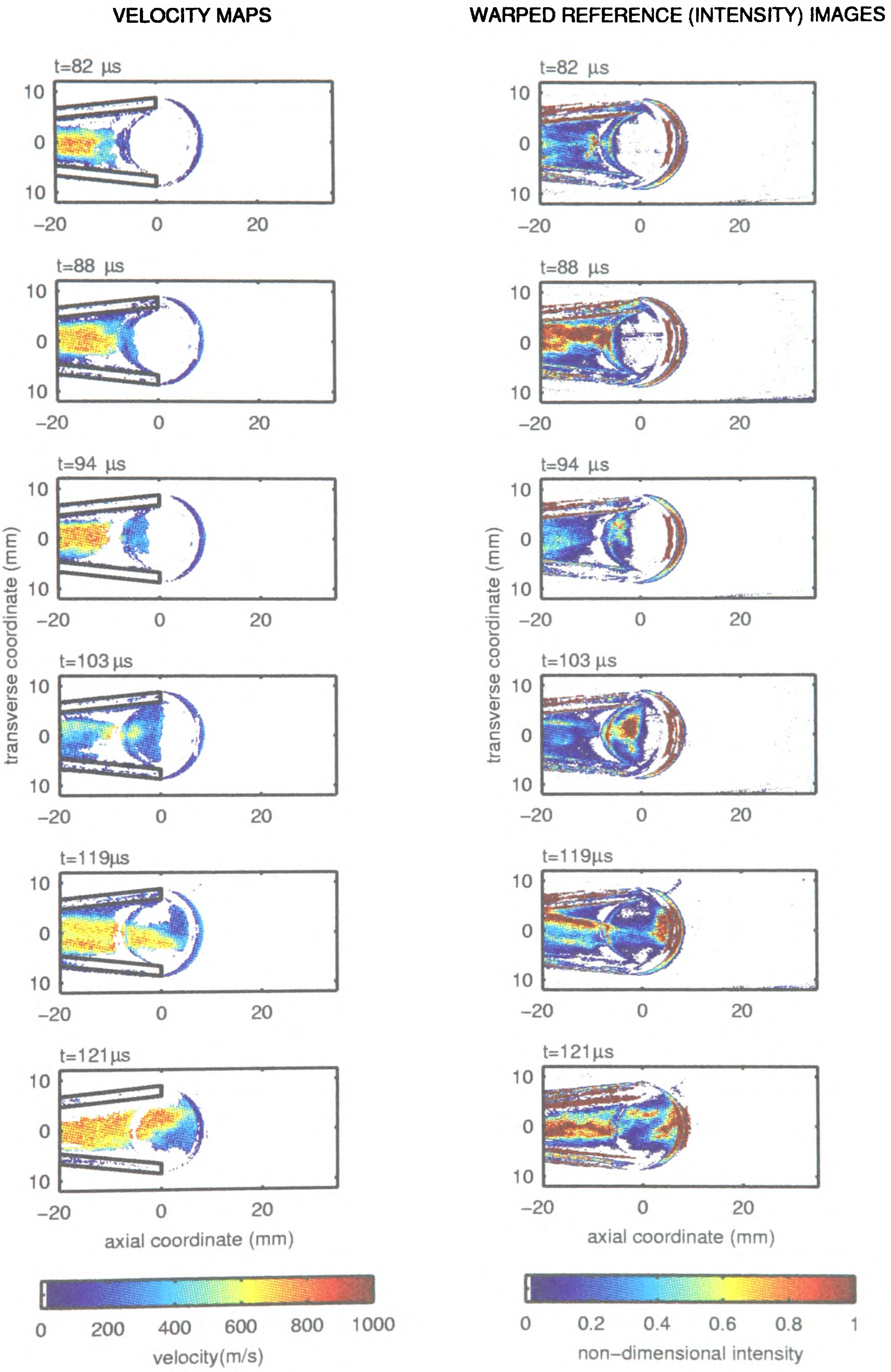


Figure 6.28 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of 4.7 µm microspheres, delivered from a conical nozzle (first in a series of five figures).

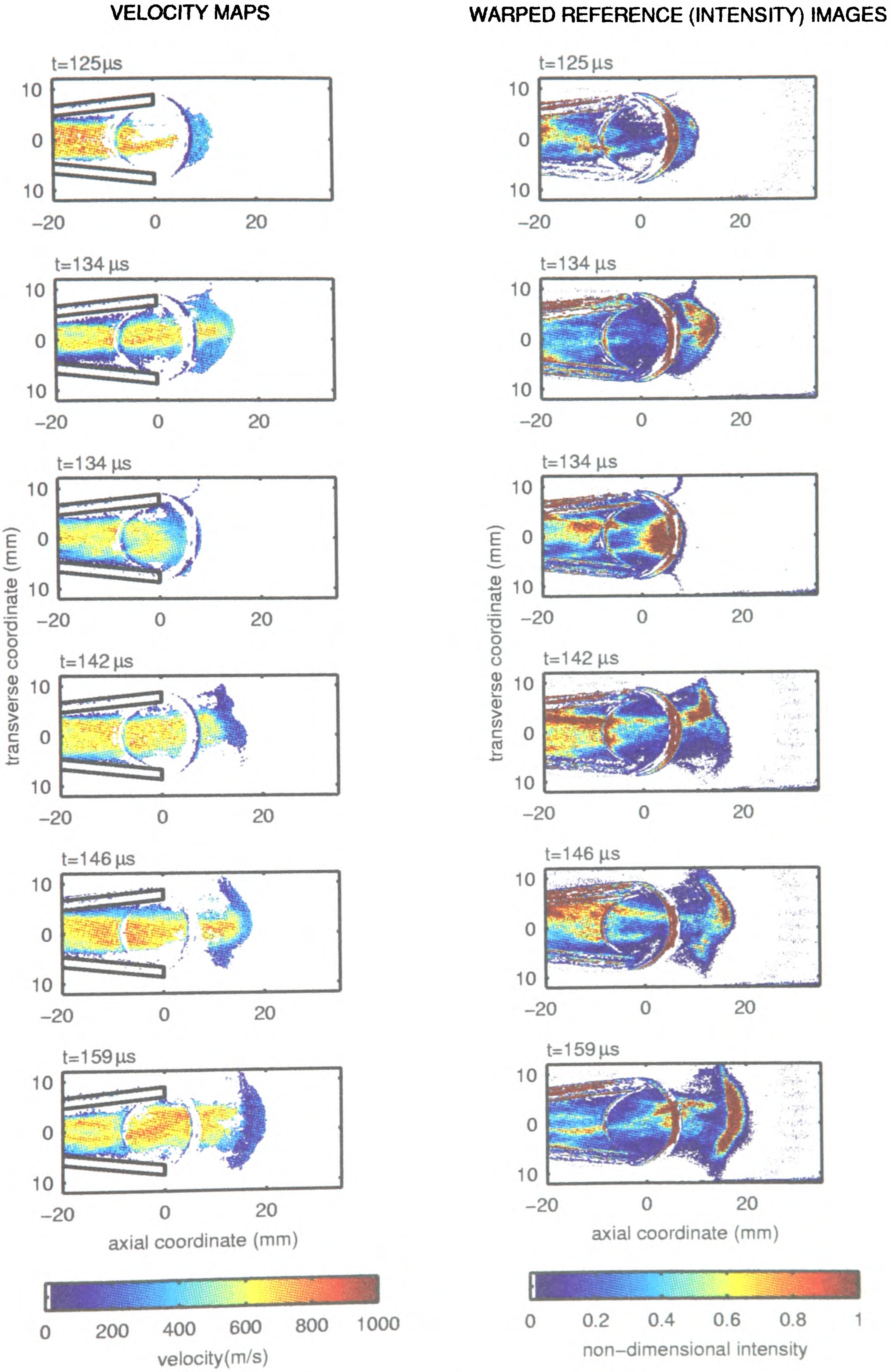


Figure 6.29 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of 4.7 µm microspheres, delivered from a conical nozzle (second in a series of five figures).

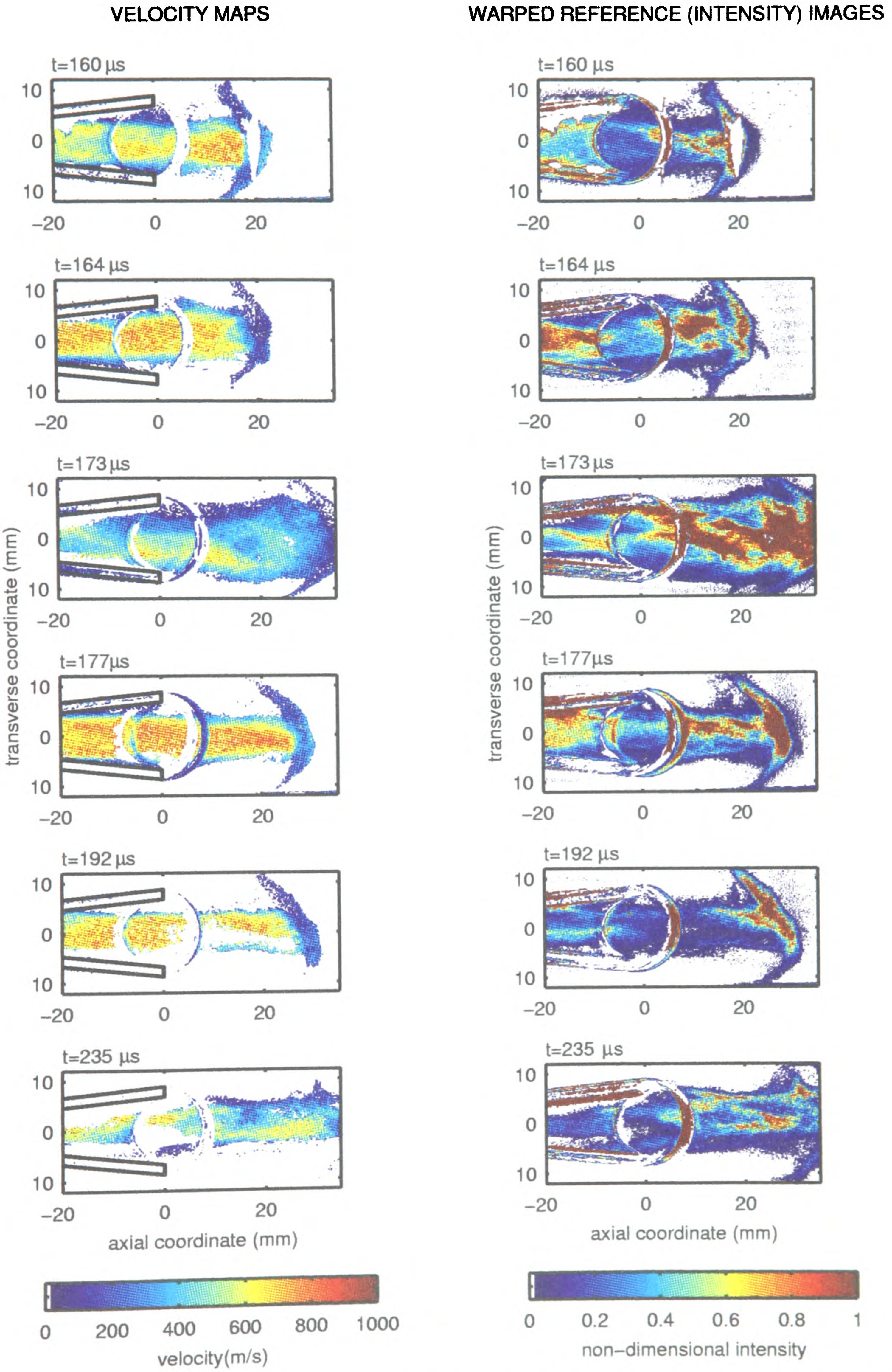


Figure 6.30 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of 4.7 µm microspheres, delivered from a conical nozzle (third in a series of five figures).

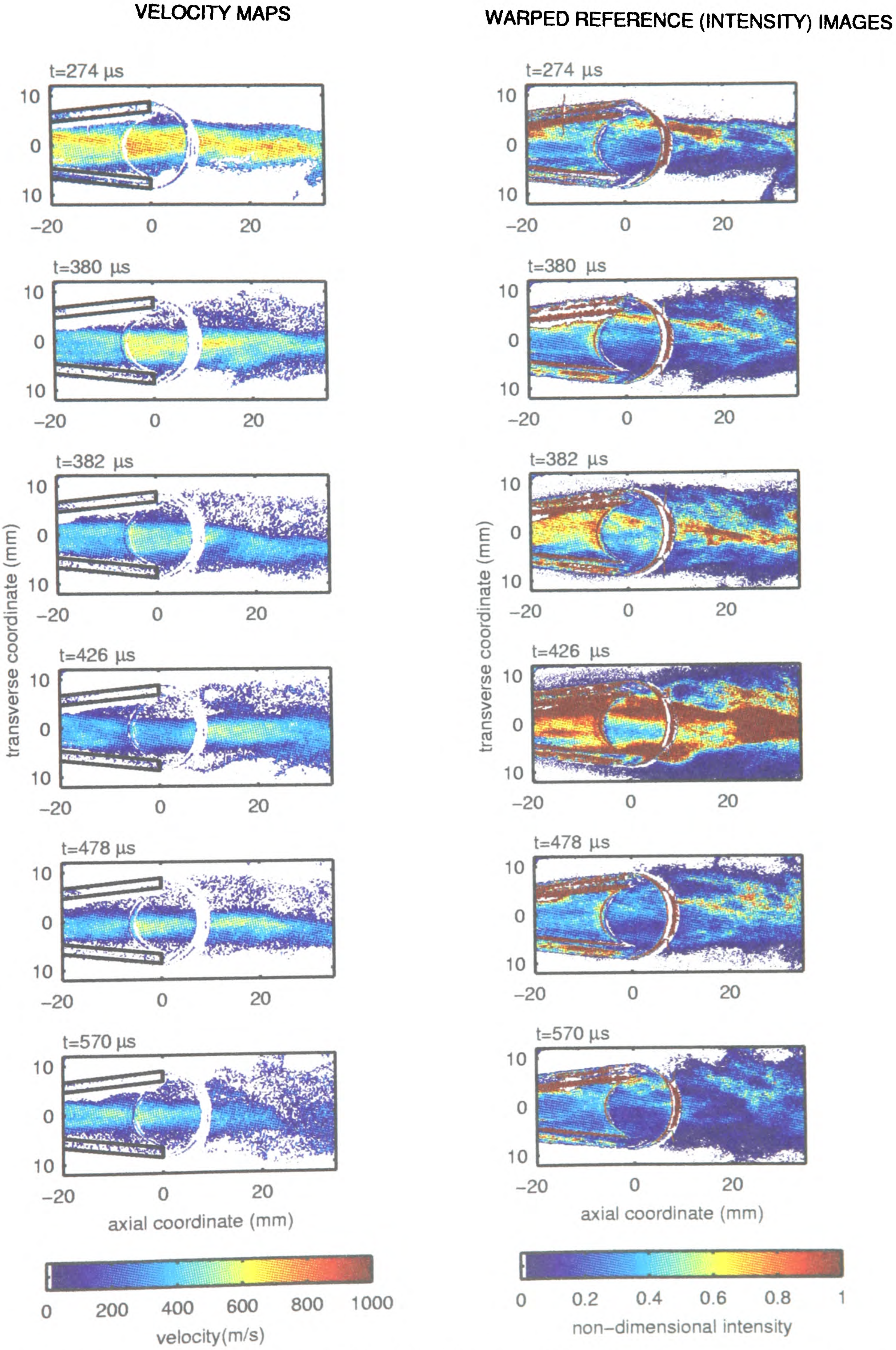


Figure 6.31 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of 4.7 µm microspheres, delivered from a conical nozzle (fourth in a series of five figures).

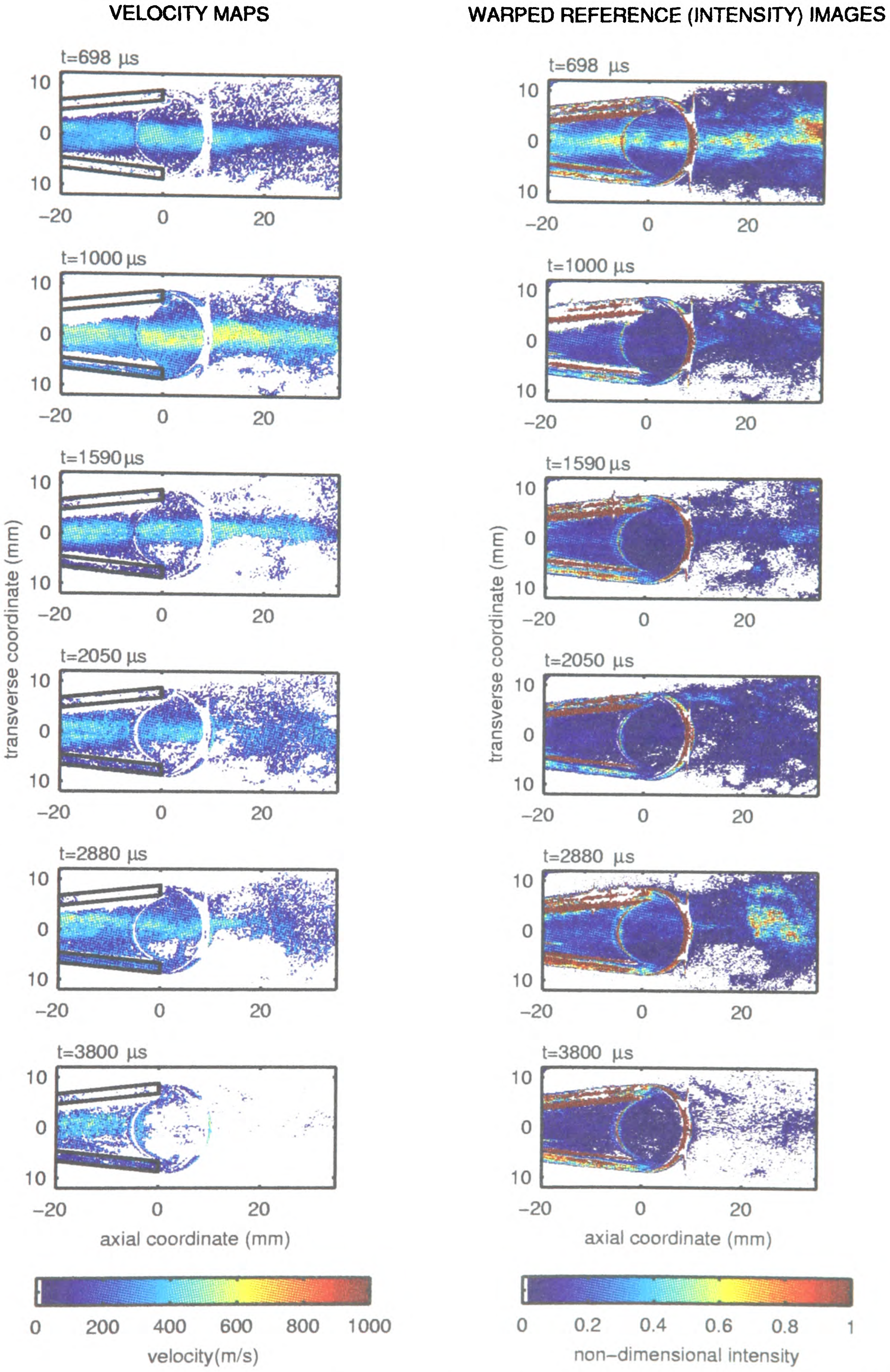


Figure 6.32 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of 4.7 μm microspheres, delivered from a conical nozzle (fifth in a series of five figures).

In the transition between the leading cloud and the high-speed stream, velocity varies smoothly between the characteristic values associated with the two main regions of the flow. Some cases are exceptional. At t = 134 μs, for example, the high-speed stream penetrates deep into the leading cloud, and velocity throughout the leading cloud is unusually high. In the velocity map for t = 146 μs, velocity in the central region of the cloud is slightly raised.

The reference images for the early period also display interesting structure. Almost invariably, particle concentration is very high in the centre of the leading cloud. Although comparable image intensity values do sometimes occur further upstream, the regions of high intensity within the cloud are consistently the largest by far in the entire field, with the sole exception of the image for t = 164 μs. Images of the high-speed stream are generally low in intensity, but are interspersed with isolated high-intensity features. Often, these take the form of long tendrils, oriented roughly parallel to the centreline. These are reminiscent of features in photographs shown by Anilkumar *et al.* [28] (see ***Figure 2.8***), taken in the later stages of a flow in which particles were accelerated from rest by the action of an expansion wave.

After the leading cloud has passed from view at $t \approx$ 200 μs, the high-speed stream remains essentially unchanged until $t \approx$ 300 μs. After that time, velocity decays appreciably, falling to around 400 m/s after 1 ms. Image intensity also falls off, and in the last image, taken at t = 3.8 ms, the powder is so sparse that particles are hardly detectable above background light levels.

Valid velocity data appear to have been obtained for locations inside the nozzle. In the majority of cases, there is no significant discontinuity in velocity in the neighbourhood of images of the nozzle rim. This indicates that the velocities measured inside the nozzle are not corrupted by reflection of the light sheet within the nozzle, or by imaging aberrations introduced by the curved nozzle wall. This is a promising result, not only for the current research, but for future applications of DGV to internal flows.

A few velocity maps do exhibit discontinuities — the map for t = 380 μs is the most prominent example. When discontinuities occur, they appear in the form of a jump in velocity across the upstream half of the image of the nozzle rim, which corresponds to the half of the nozzle cone lying further from the DGV imaging system. This suggests that reflection inside the nozzle is the culprit for the discrepancy. Of course, any argument based on an assessment of continuity across the nozzle rim images is valid only for the region near the nozzle exit. At points further upstream in the nozzle, the more severe wall curvature may have an impact, both through unwanted reflections of the light sheet and through image aberrations. Encouragingly, an independent experiment, described in §5.5.3, suggested that the impact of image aberration on alignment (though detectable) does not increase with distance upstream of the nozzle exit, for image features near the centreline.

6.4.3 Interpretation of Time-Resolved Velocity Measurements

The experimental results presented above provide key insights into the unsteady mechanisms of particle flow in the conical nozzle, which have provided the basis of a qualitative theory of the drug delivery flow-field. The time-resolved DGV measurements are complemented by experimental results in gas dynamics, described in Chapter 4, along with published literature and classical theory in gas dynamics, cited in Chapter 2. These various strands are brought together in this section.

In broad terms, the findings of time-resolved DGV suggest that the leading cloud of particles is associated with gas flow in the transient starting process, and that the subsequent high-speed stream is associated with the quasi-steady nozzle flow. This understanding is supported by considering the shape and velocity of the particle cloud or jet in the two observed regimes of flow.

It was explained in Chapter 2 that the primary shock wave, which leads the starting process, is expected to decelerate and weaken as it moves through the divergent part of the nozzle. Moreover, the shock and the flow which immediately follows it must fill the width of the nozzle. This point has been demonstrated simply by detection of the shock at wall-mounted pressure transducers, and in the published images of nozzle starting flows [24,21,22,23] referred to in Chapter 2. It is also physically reasonable, since it is difficult to conceive of any flow-field structure which would allow low-pressure atmospheric gas to exist in a gap between the starting shock wave and the nozzle wall, holding high-pressure shocked gas away from the wall. The large width of the leading cloud of particles, therefore, is an important characteristic by which it can be identified with the starting process.

The second distinctive characteristic of the leading particle cloud is its low velocity, in comparison with the particles which follow it in the long-term jet. This is consistent with a strong role for the starting process in acceleration of the leading particles. Based on measurements of shock propagation speed in the nozzle, it was stated in §4.4 that gas velocity behind the starting shock is 205 m/s in a flow loaded with 1 mg of particles. This is less than the average of the measured velocities of the leading particles inside the nozzle, which typically lie between 200 and 400 m/s. However, the discrepancy may be due to momentum acquired by the particles at an earlier stage, when the primary shock was in a narrower part of the nozzle and stronger. Alternatively, the leading particles may be gaining momentum from the high-speed stream behind them. This possibility will be revisited at a later stage in this discussion.

Arguments laid out in §2.1.2 and §2.2.2 indicate that in the quasi-steady flow left behind by the starting process, gas velocities are supersonic downstream of the throat as far as a system of shock waves. The shock wave(s) may trigger boundary layer separation from the nozzle wall. In the measurement region, the high-speed particle jet is straight, parallel to the nozzle axis and appears to bear no relation to the profile of the nozzle; this suggests that it is indeed separated. Depending on the character of the postulated shock system, gas velocities between subsonic and low supersonic values can be expected in the shocked flow. This is consistent with the highest measured particle velocities, at around 800 m/s. The uniformity of particle velocity in the axial direction suggests either that the particles are close to equilibrium with the gas, or that the local gas density is too low to effect any appreciable drag on the particles.

It is possible that the leading cloud contains particles which remain embedded in the transient starting flow from the moment the expansion wave passes over them in the drug cassette. However, the flow provides a mechanism by which particles can become involved with the starting process at a later stage. Particles which are entrained in the quasi-steady stream must come gradually closer to the low-speed starting process. After sufficient time elapses, these particles must approach and enter a region of relatively slow flow dominated by the starting shock. There, they encounter significant slip velocity, or collide with a dense particle cloud, and decelerate.

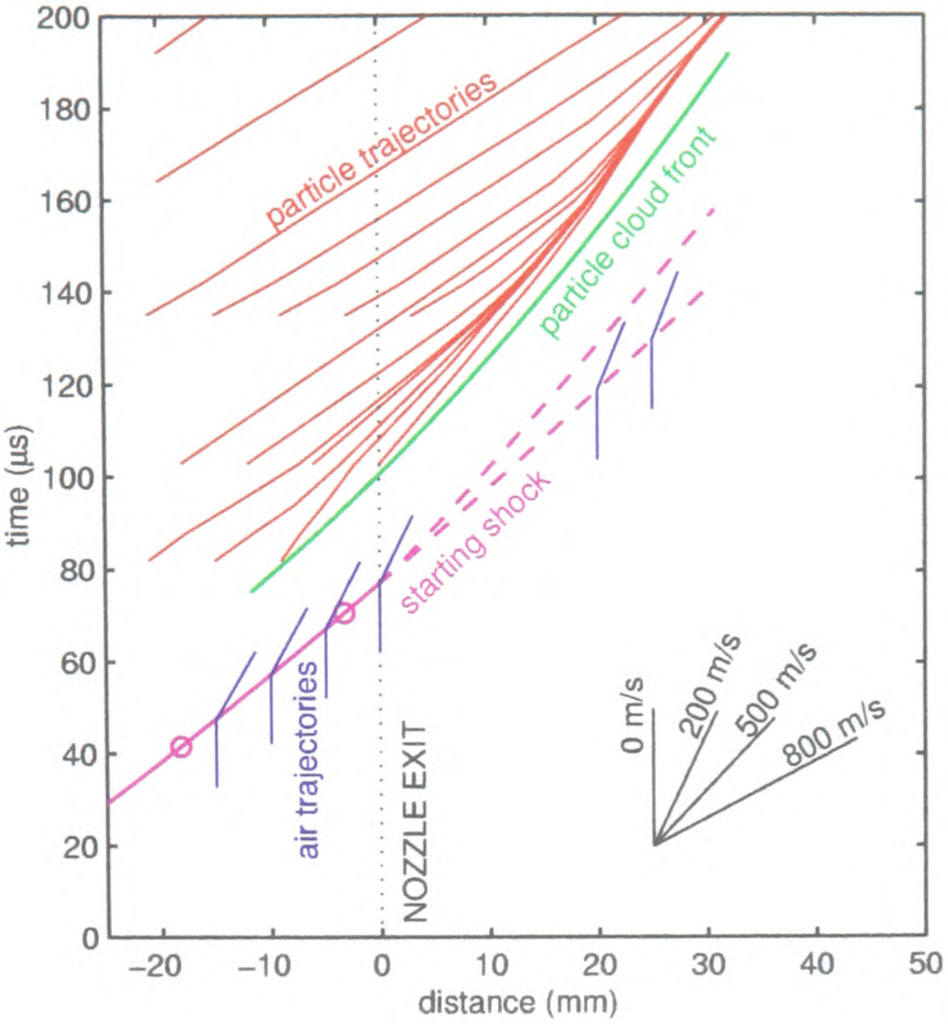


Figure 6.33 A space-time diagram, based on time-resolved DGV and nozzle pressure measurements, summarising experimental data for the first 200 μs of flow.

This relationship between the quasi-steady jet and the leading cloud is illustrated in ***Figure 6.33***, a space-time diagram which is a graphic summary of the time-resolved velocity data and some nozzle pressure data. It includes representative particle trajectories and the trajectory of the particle cloud front, both derived by interpolation of smoothed time-resolved DGV results. It also includes the trajectory of the starting shock, along with the theoretical trajectories followed by air molecules as the starting shock sets them in motion. The shock's speed inside the nozzle is based on nozzle pressure measurements, indicated by circular marks. Two possible shock trajectories downstream of the nozzle exit (shown in dashed lines) were estimated using Chisnell's method [17]. For the first, it was assumed that the shock propagates through an imaginary extension of the nozzle. Alternatively, it was assumed that the shock takes the form of a spherical surface, centred on the apex of the cone formed by the nozzle's divergent section, and bounded by the intersection of the sphere with the plane of the nozzle exit.

This space-time diagram is not a precise representation of the data, as considerable smoothing of the data was involved in its preparation, to ensure smooth particle trajectories. Nevertheless, it graphically illustrates the assertion that particles in the quasi-steady jet constantly pour into the leading cloud. The particle paths from positions far upstream converge onto a common trajectory near the cloud front.

The observation that the speed of the primary shock increases slightly with payload mass, reported in §4.4, falls into place in this understanding of the flow-field. The particles and gas at the front of the high-speed quasi-steady flow must exert a force on the growing low-velocity cloud as they decelerate and merge into it. Their momentum is redistributed in the leading cloud through particle collisions and drag force. Therefore, the leading particle cloud continually accelerates with respect to the gas immediately in front of it. In doing so it compresses that gas, and compression waves propagate forwards at the local speed of sound. They catch up with the starting shock, coalesce with it, and strengthen it.

6.4.4 Uncertainty Estimate

A typical measurement was selected from the series for a 1.0 mg payload of 4.7 μm microspheres described above. Using the method described in §5.3, measurement uncertainty was estimated for this particular case as a representative sample of the data set. The result is shown in ***Figure 6.34***. The estimated uncertainty is quite uniform throughout the field, at 70 – 90 m/s, or 8 – 10% of the maximum velocity in the field. A breakdown of the error sources in ***Figure 6.35*** shows that misalignment and shot noise errors are negligible in comparison with errors due to the iodine cell calibration. These uncertainties are no greater than those estimated for time-integrated measurements near the periphery of the jet, which were dominated by misalignment.

These relatively large error estimates detract slightly from the significance of these measurements as quantitative studies, but the importance of the information they provide, as a unique set of insights into this flow-field, is not diminished. It should be noted that the validity of comparisons within velocity maps and between maps in the same sequence is not at all compromised by the expected uncertainty, since it is largely systematic in nature. Comparisons between velocity maps measured several days apart may be suspect, because of the significance of the iodine cell transmission instability (which is, however, accounted for conservatively in the stated uncertainty estimates).

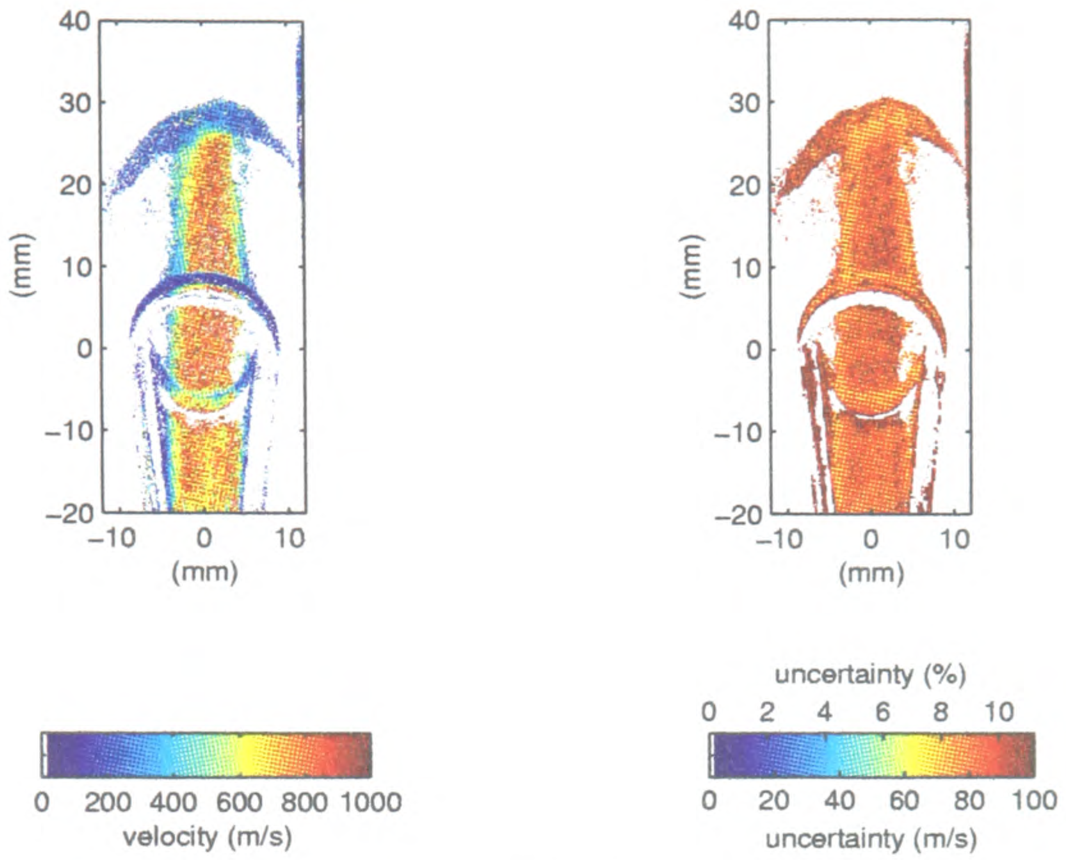


Figure 6.34 (*Left*) an instantaneous velocity field measured 177 μs after membrane rupture in a flow with 1.0 mg of 4.7 μm microspheres, and (*right*) the estimated measurement uncertainty, expressed as absolute velocity and as a percentage of the maximum velocity in the field. Flow is from bottom to top.

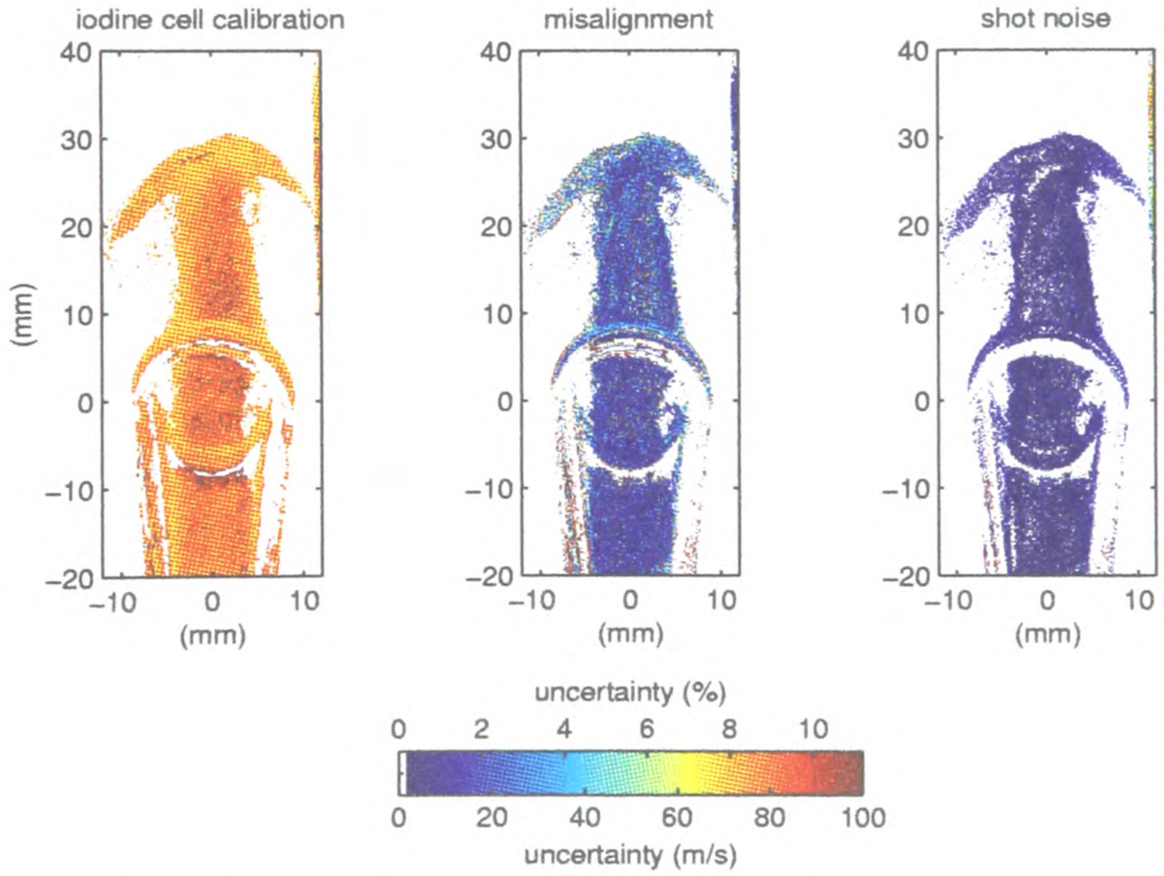


Figure 6.35 Contributions of iodine cell transmission profile inaccuracies, misalignment and shot noise to computed uncertainty in the time-resolved velocity measurement shown in ***Figure 6.34***, expressed in terms of absolute velocity and as a percentage of the maximum velocity in the field.

6.4.5 Velocity Estimates Based on Cloud Front Position

Time-resolved DGV is a new and relatively unproven technique, which has been applied here to measure velocity in a flowfield for which no comparable data are available from any other source.

Any independent velocity measurement would be a valuable as a check on the results, and would build confidence in the DGV results.

Fortunately, a simple measure of the velocity of leading particles is available through the flow-field images acquired in the process of the DGV measurements. By measuring the movement of the front of the cloud from image to image, it is straightforward to calculate the average velocity of the cloud between images. This estimate, called the *front-tracking velocity*, can be compared directly to the velocity reported by DGV for particles at the cloud's leading edge. The travel of the cloud front is most easily gauged in the warped images, since they are undistorted images of the measurement plane.

Before implementing a method of this type, two issues must be approached. First, a consistent and objective means of identifying the position of the cloud front in each image is required. The particle cloud front does not appear as a sharp edge in the image, so for this purpose a cutoff image intensity level is specified. Pixels with intensities below the cutoff are considered background pixels, and rejected. Isolated pixels which exceed the threshold in intensity are also rejected, to filter out spurious bright spots in the background. The pixels which belong to the image of the particle cloud, according to these criteria, are searched to identify the furthest forward pixel in the cloud. A strip along the image of the cloud front, centred on that pixel, is then designated, 21 pixels wide (in the radial direction) and 5 pixels deep (in the upstream direction). The axial location of the cloud front, and the DGV measurement of particle velocity on the cloud front, are taken as averages over all these pixels.

The threshold for identification of the cloud front was set at a normalised image intensity level of 0.125. This was chosen by using a range of cutoff values to identify the cloud front in a set of 7 nominally identical test measurements, and seeking the value around which the derived front location was least sensitive to the cutoff level.

The second complication with the front-tracking velocity calculation arises because the images on which it is based are not successive images of the same evolving flow-field. Each measurement is a glimpse of a unique unsteady flow-field which is not perfectly repeatable. The resulting uncertainties in the front-tracking velocity estimates cannot be eliminated, but if they are quantified, a criterion can be set for the level of agreement demanded between the DGV and front-tracking measures of velocity.

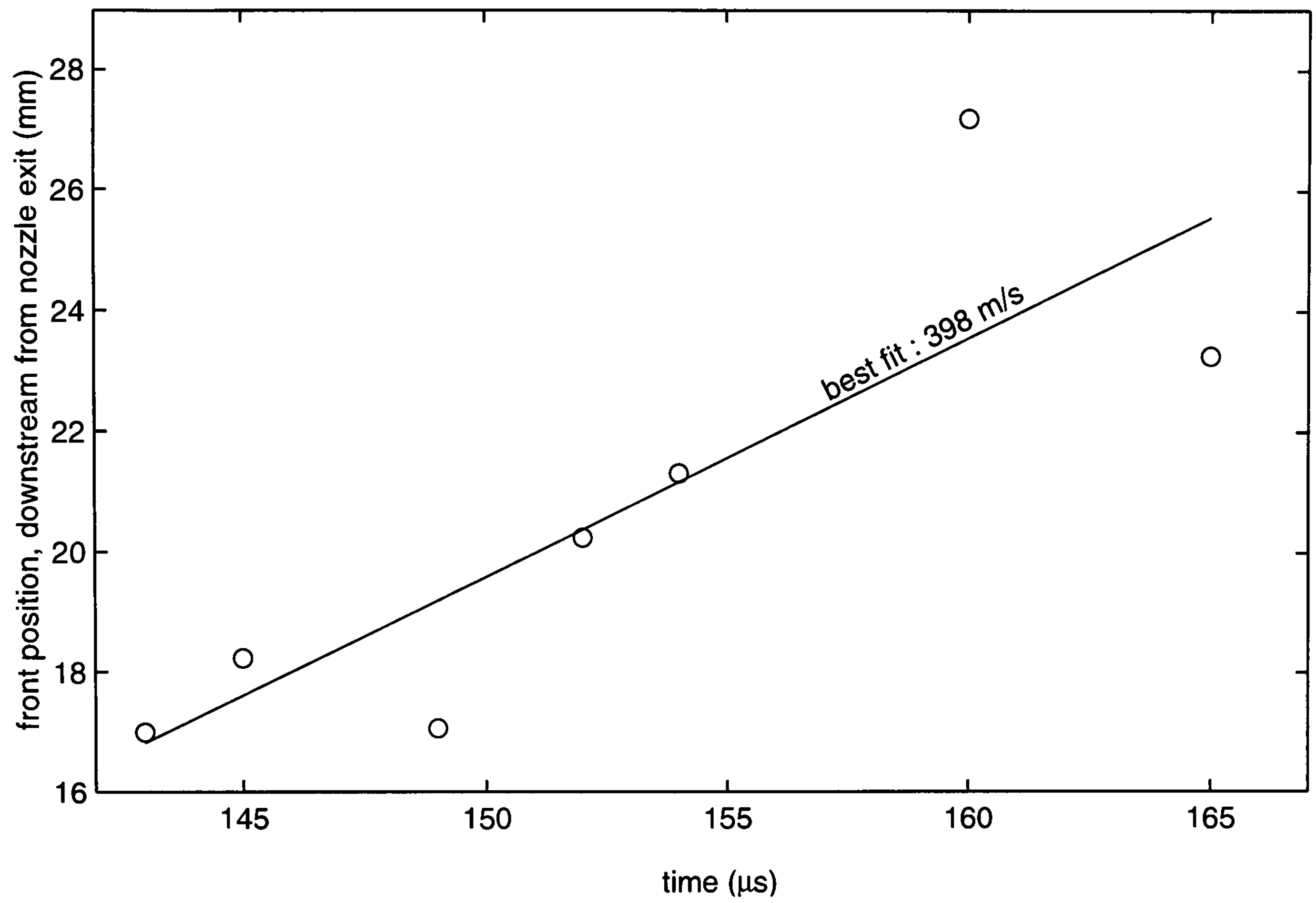


Figure 6.36 Position of the particle cloud front as a function of time from bursting of the downstream membrane, for 7 experiments under nominally identical conditions.

With this in mind, baseline information on shot-to-shot variability was gathered from the 7 nominally identical test measurements mentioned above. The cloud front locations in the 7 test images are plotted in ***Figure 6.36***. The front locations, as well as the estimated time t from the downstream membrane burst to the laser pulse, vary considerably. This variability is probably a combination of errors in identification of the reference time $t = 0$, and physical variations in the timing of the flow-field itself. It is convenient, however, to imagine that time is measured accurately and consistently, and to attach all the variability to the cloud front position.

A linear fit (shown in ***Figure 6.36***) for front location in the 7 test images as a function of time was computed, and the standard deviation from this fit was chosen as a measure of the variability in front position. This deviation is a guess at the variation which would be observed if multiple measurements at an identical time in the flow-field history were available. As an indicator of variability, it is a more conservative choice than the straightforward (and larger) standard deviation of cloud front location, in the sense that it will ultimately yield a lower uncertainty in the front-tracking velocity estimate, setting a tougher criterion for agreement with DGV. In physical terms, this approach represents an assumption that changes in the cloud front position in repeated shots reflect an underlying forward motion of the cloud at constant velocity, on which some random variability is superimposed. It is this variability which is described by the deviation from the linear fit. The standard deviation of cloud front position from a linear fit, computed in this way, is 2.0 mm.

This standard deviation can be regarded as an uncertainty in measurement of the "true" cloud front location, imagined as the mean of many measurements. It can then readily be incorporated in the front-tracking velocity measured between images *i* and *j*, which is defined as:

$$V_{i,j} = \frac{x_j - x_i}{t_j - t_i}, \qquad (6.1)$$

where *x* is the position of the front in each image, and *t* the time at which each image was captured. The error ε_x in front location translates into an error in velocity according to:

$$\varepsilon_{V,i,j} = \frac{\sqrt{\varepsilon_x^2 + \varepsilon_x^2}}{t_j - t_i}. \qquad (6.2)$$

This can easily be calculated for any given pair of cloud front locations, providing an estimate of uncertainty in the front-tracking velocity.

The cloud front is identifiable in 13 of the results included in ***Figures 6.28*** to ***6.30***. They have been used to compute front-tracking velocities, which are shown in ***Figure 6.37***. Each velocity is associated with a spatial location half-way between the front locations in the two images on which it is based. The discrepancies between DGV and particle tracking are far less than the uncertainties in front-tracking velocity, which are in hundreds of m/s. Agreement is particularly good in the region from 10 mm to 23 mm downstream of the nozzle exit.

However, a velocity estimate which is accurate to within hundreds of m/s is not an exacting standard against which to compare DGV. The uncertainties in front-tracking velocity can be reduced to more useful levels by basing the calculations on pairs of images separated by larger time intervals. Therefore, velocities were recalculated by considering sequences of six images within the data set, and basing the velocity on the first and sixth of each sequence (the first and sixth, second and seventh images, and so on, of the whole series of 13). Similarly, velocities were calculated from the first and fourth images in sequences at the beginning and end of the main measurement series, to provide extra data points. The resulting particle front velocities are plotted in ***Figure 6.38*** with the corresponding DGV velocities for comparison.

The results are encouraging. Uncertainties on the front-tracking velocities range from ±50 to ±95 m/s for velocities computed across 6 images, with ±115 m/s on the first point, calculated across only four images. In the region where ample velocity measurements of both types are available — from 15 to 25 mm downstream of the exit — the majority of DGV points lie within the error bars of their neighbouring front tracking points. Upstream of the nozzle exit, the DGV velocities are lower than the one available front-tracking estimate, although all but one of them are within its error bar. The one serious anomaly occurs in the first 10 mm downstream of the exit plane, where DGV estimates are not available. There, particle tracking velocities decrease almost linearly with distance, while the DGV velocities at either end of the region show an increase. Overall, however, agreement between the DGV measurements and the front-tracking velocity estimates is satisfactory.

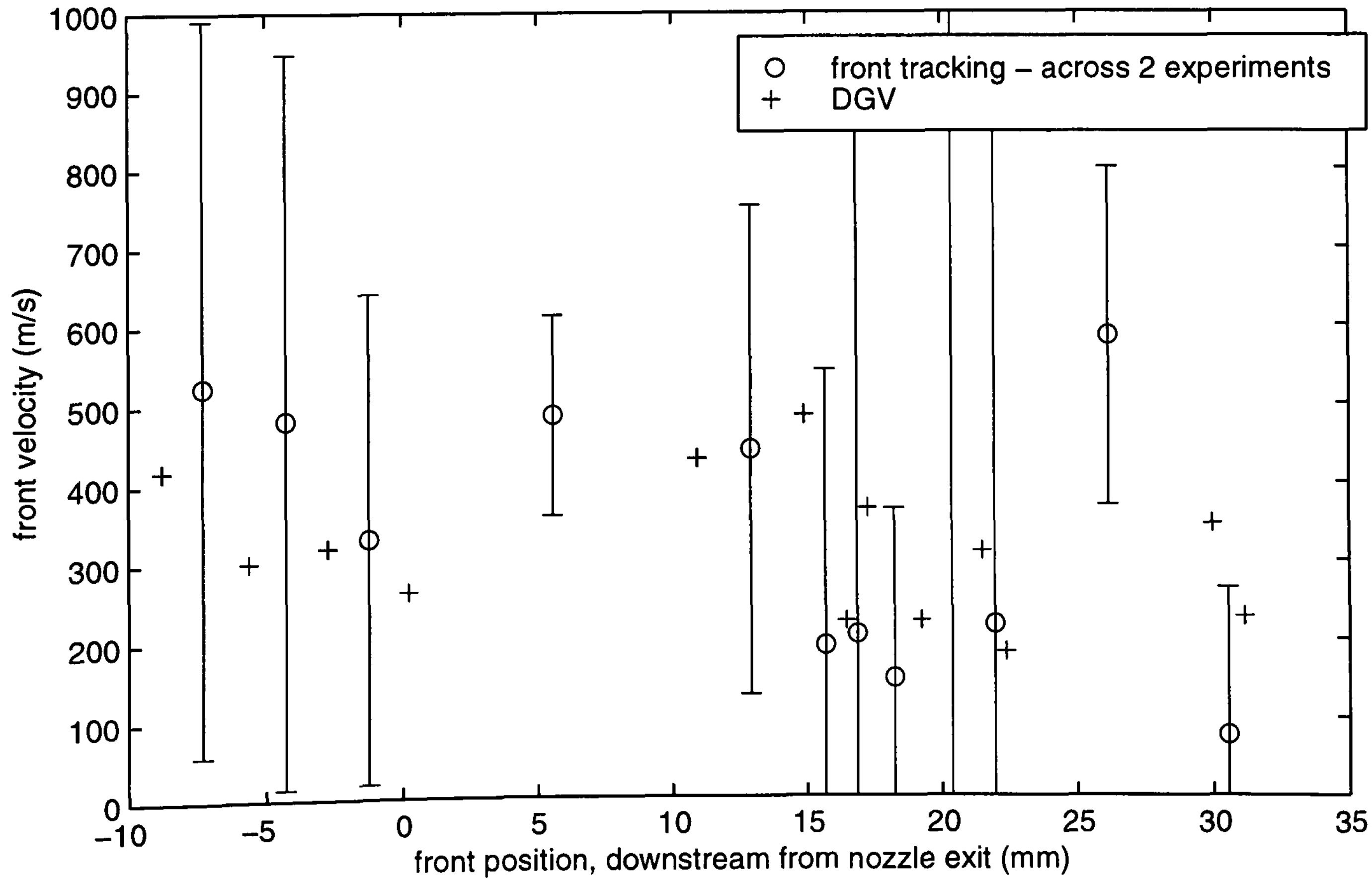


Figure 6.37 A comparison of DGV measurements of velocity at the front of the particle cloud with velocity estimates based on cloud front tracking, computed over intervals between successive images. Error bars show the estimated uncertainty in the front-tracking velocity. One point, with a velocity of 2320 m/s at a position of 20.4 mm, is not visible on this scale.

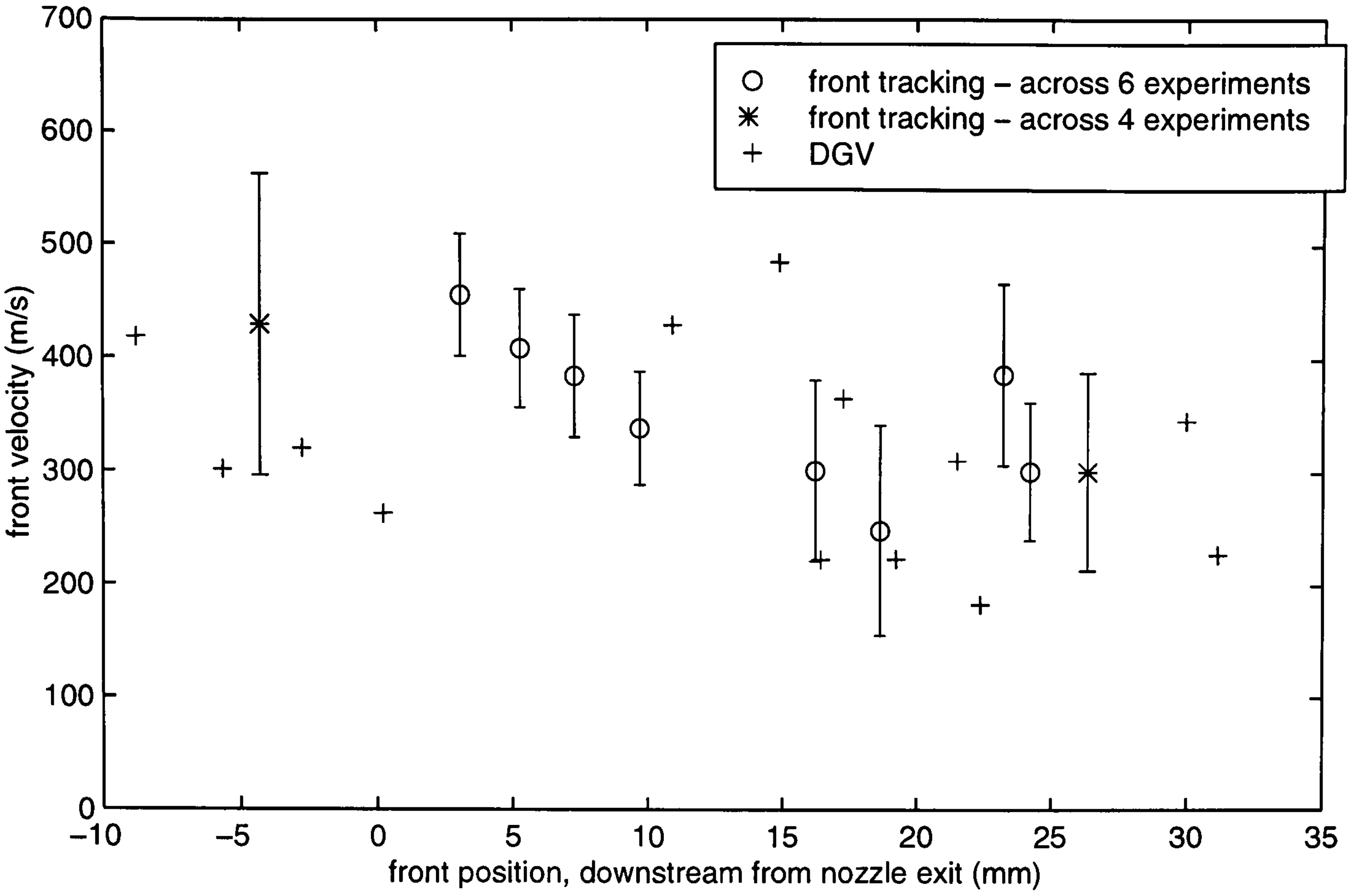


Figure 6.38 A comparison of DGV measurements of velocity at the front of the particle cloud with velocity based on cloud front tracking, computed over intervals spanned by six and four data points. Error bars show estimates of the r.m.s. uncertainty in front-tracking velocity.

6.4.6 Additional Results

To supplement the detailed velocity field history measured for 1.0 mg of 4.7 μm spheres, described in §6.4.2, less extensive measurements have been carried out for various other operating conditions. Further tests with the same payload were carried out with the aim of measuring velocity and visualising the flow as far upstream of the nozzle exit as possible. Measurements were also carried out with lidocaine and silica powder payloads.

In the results shown already, the flow inside the nozzle was measured over the downstream 20 mm of the nozzle. Measurements further upstream were desirable to expand knowledge of flow mechanisms in the nozzle, particularly the suspected separation of the jet from the nozzle wall, and also to provide data for comparison with numerical predictions. To this end, the drug delivery device was repositioned so that the measurement region extended 37 mm upstream of the nozzle exit, a limit imposed by the external geometry of the nozzle. Six of the velocity maps and reference images delivered by this series of experiments are shown in ***Figure 6.39***.

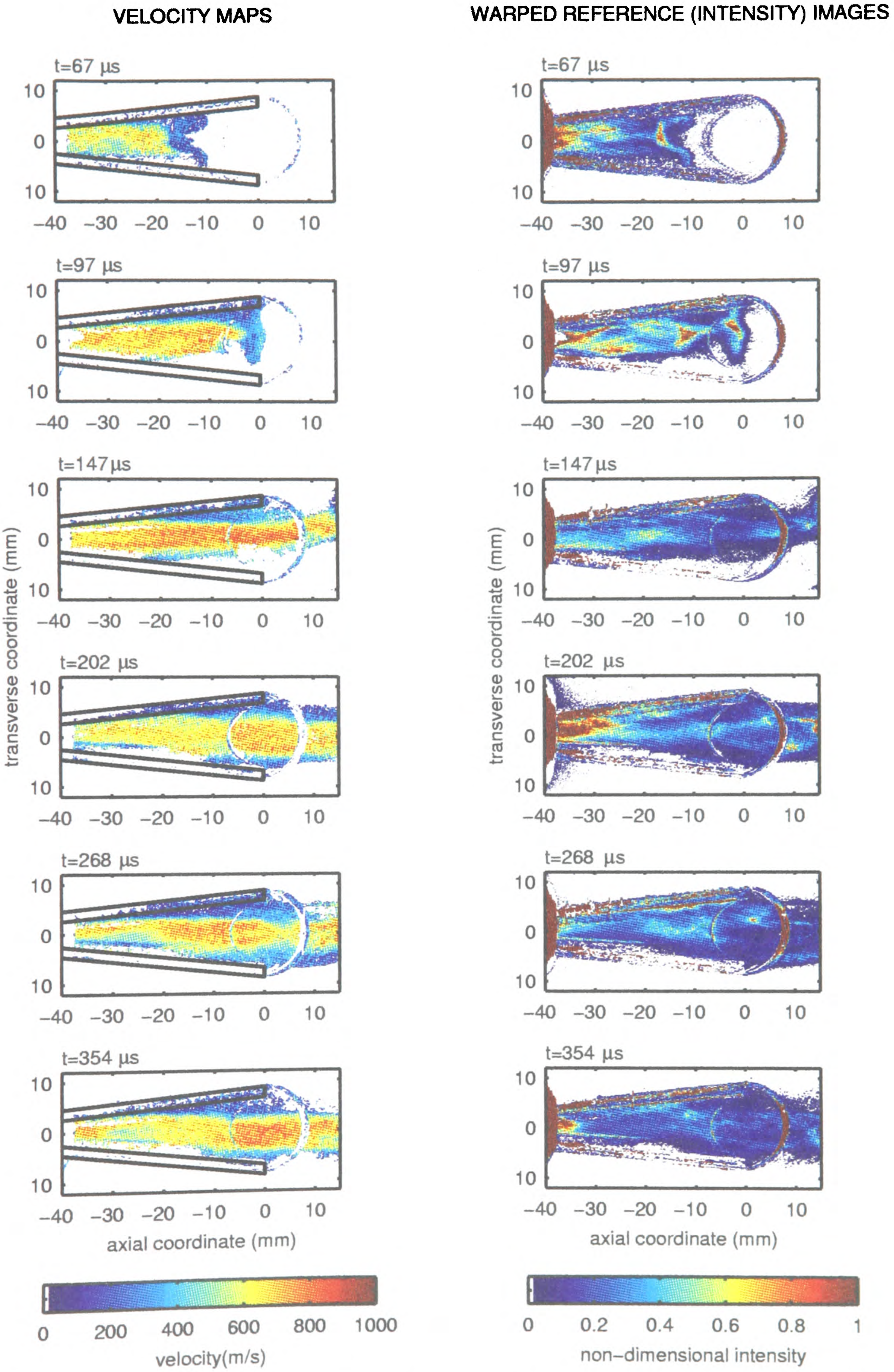


Figure 6.39 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of 4.7 μm microspheres, delivered from a conical nozzle, measured with a wide view of flow in the nozzle.

As noted earlier in §6.4.2, accuracy has not (as yet) been demonstrated for measurements inside the nozzle in regions far upstream of the nozzle exit, where the wall curvature is high. One new potential problem for such measurements is evident in these results, most prominently in the last three measurements shown in ***Figure 6.39***. A thin streak of low measured velocity appears in the lower half of the upstream end of the nozzle, apparently radiating from the apex of the nozzle cone. It is due to a glare on the nozzle, which was visible while setting up the experiment, and appears in the reference images as a bright streak.

These results contain interesting and important information. The front of the particle cloud has a concave shape as it approaches the nozzle exit. In the image for 97 μs, however, the front has a convex profile only 1-2 mm downstream of the exit, demonstrating consistency between this set of measurements and the main series.

Separation of the particle-laden flow from the nozzle wall is clearly discernible in some of the images. Upstream of the separation point, particles at high velocity (over 650 m/s) fill the nozzle. Separation is manifested by the appearance and growth of a wedge of low velocity between the nozzle walls and the high velocity flow. The high velocity core itself becomes detached and effectively continues as a jet. On the upper wall, this occurs approximately 25, 23, 17, 30 and 28 mm upstream of the nozzle exit, in the measurements at $t = 97, 147, 202, 268$ and 254 μs, respectively. The separation point is not discernible on the lower wall of the nozzle, where velocity information is not available because of low image intensity, due to asymmetry of the light sheet with respect to the nozzle centreline. There appears to be no systematic movement of the separation feature over this time range, though this data set may be too small to capture such a trend amid statistical fluctuations. Also, as noted in Chapters 2 and 4, the separation structure may be oscillating.

Velocity fields were measured for payloads of 1.0 mg of silica powder in the expectation that these particles would follow the gas flow closely, so that their velocity would be an indicator of gas velocity. Six reference images and measured velocity maps are shown in ***Figure 6.40***. Though flow structures and the magnitudes of the measured velocities are quite similar to those observed for 4.7 μm polystyrene microspheres, the velocity fields measured for silica are much less smooth. The front of the particle cloud is less sharply defined, and the boundary between the leading cloud and the quasi-steady stream is poorly defined, with a more intricate shape. In two images of the external jet, at 166 and 188 μs, wisps of low-velocity particles appear to trail from the edges of the leading cloud and curl back towards the centreline. They suggest vortices formed by a shearing action between the particle cloud and relatively slow-moving gas around it. This indicates again that the leading particle cloud is moving faster than the gas which surrounds it. The fact that the silica can follow such detailed flow features suggests that it contains particles which are small enough to function as true flow tracers.

Complex, uneven patterns are evident throughout all parts of the measured flow fields. This suggests that the underlying gas flow involves strong unsteady fluctuations, to which the silica particles are small enough to respond, though the microspheres are too large. The shear layer at the edge of the external quasi-steady jet and shock-induced separation in the nozzle are possible sources of these fluctuations.

Measurements of the velocity field in delivery of 1.0 mg of 38-53 μm lidocaine are shown in ***Figure 6.41***. Velocity magnitudes are broadly similar to those measured for the other two particle types. With detailed small-scale structure in the flow, ragged jet boundaries and vortex-like features rolling off the leading cloud, these velocity fields resemble the results for silica more than the results for microspheres. These observations suggest that particles smaller than the 4.7 μm microspheres are present in ample quantities in the lidocaine payloads, despite their nominal particle size range (based on sieve mesh spacing) of 38 – 53 μm. Time-integrated measurements also support this conclusion, as discussed in §6.2.4. In time-integrated velocity maps of flow from the Mach 2.5 contoured nozzle (a device which has been shown to be sensitive to particle size) the spreading of the lidocaine particle jet suggests the presence of very small particles. Time-integrated centreline velocities in the same flows are on a par with those of 26.1 μm microspheres, regardless of the initial nominal particle size. Taken together, all these observations strongly indicate that lidocaine particles are broken up in the flow-field before reaching the nozzle exit.

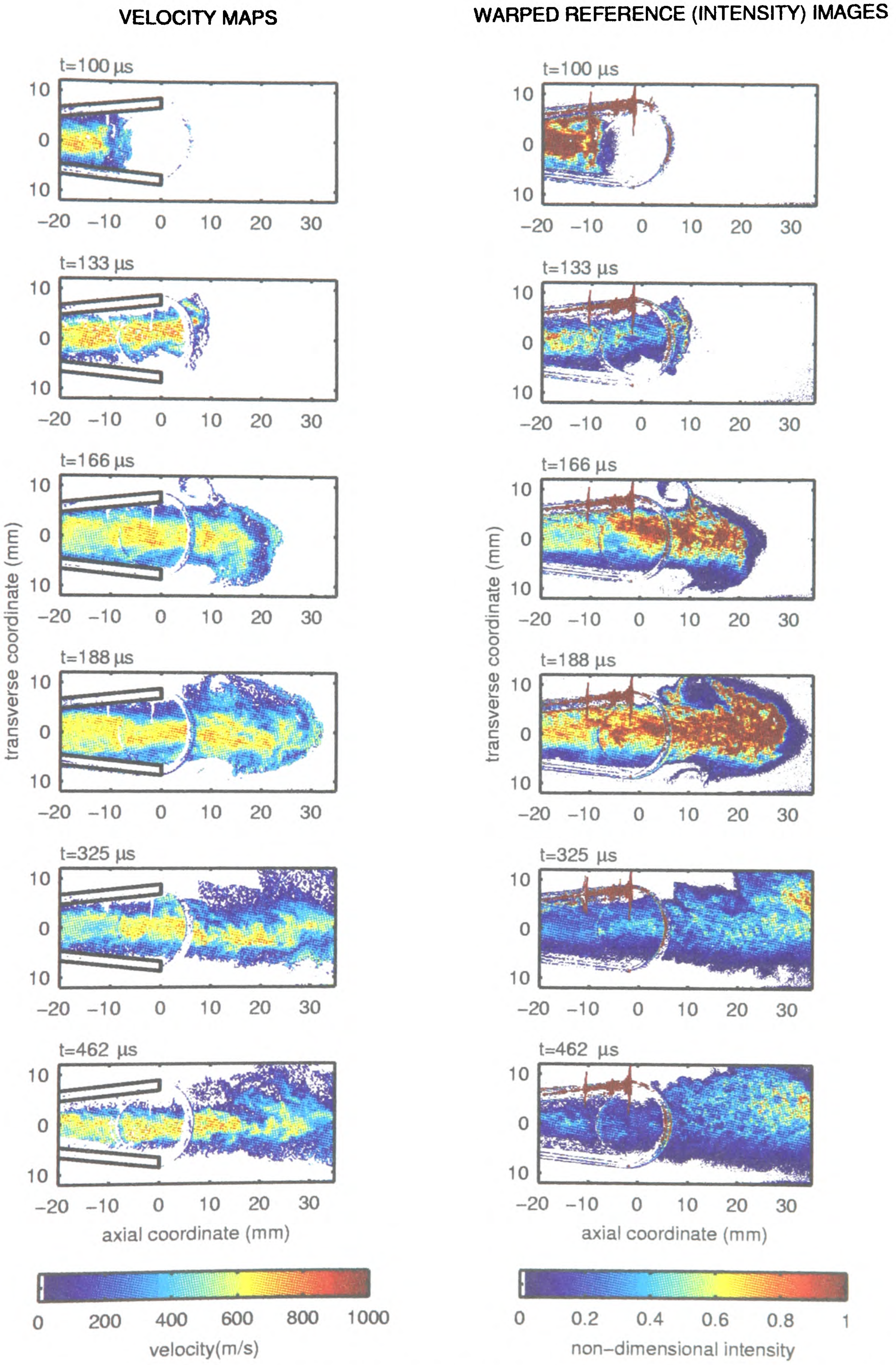
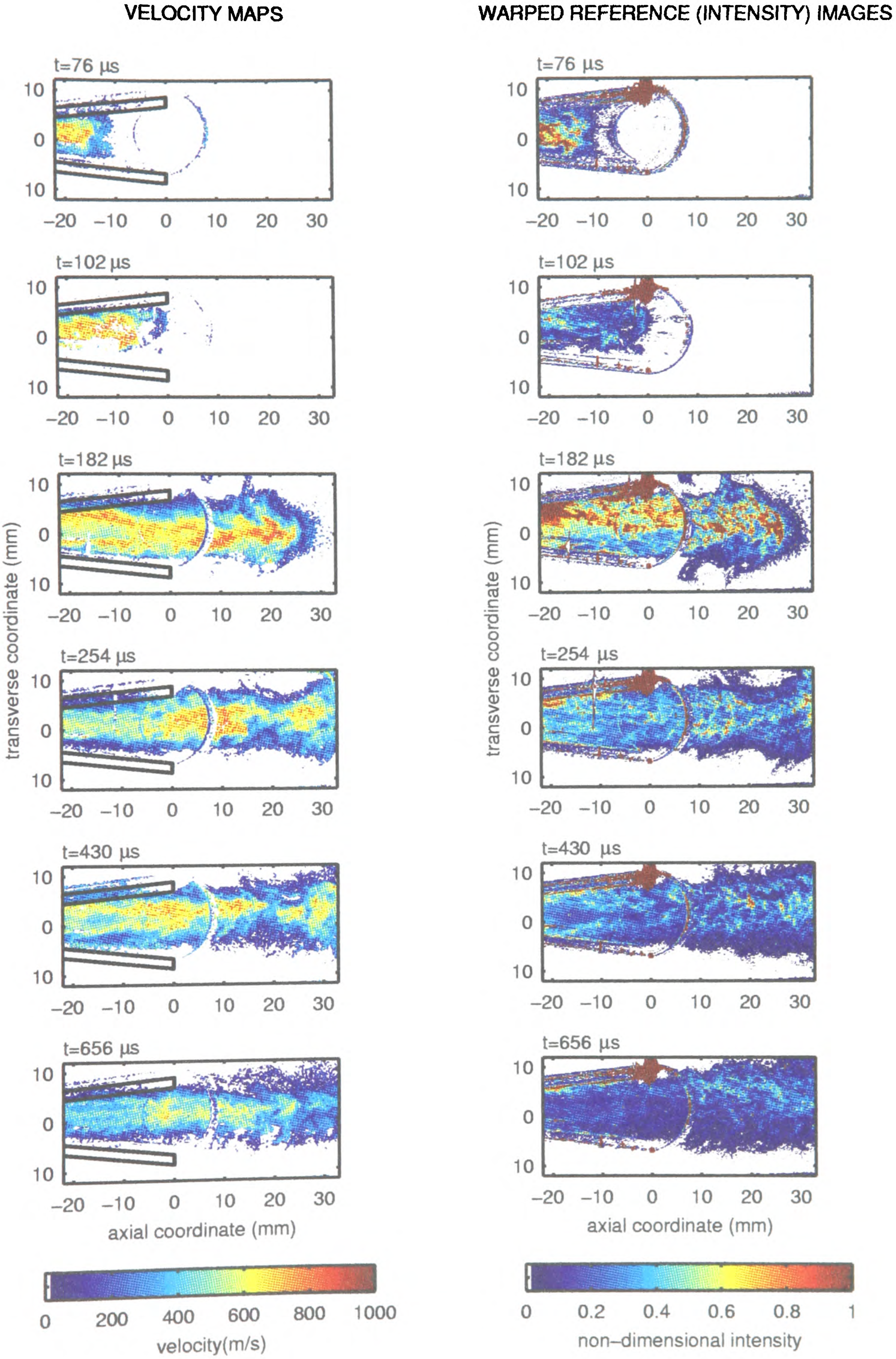


Figure 6.40 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of silica powder, delivered from a conical nozzle.



VELOCITY MAPS

WARPED REFERENCE (INTENSITY) IMAGES

Figure 6.41 Instantaneous measured velocity fields and reference images for a 1.0 mg payload of 38-53 μm lidocaine, delivered from a conical nozzle.

6.5 Comparison of Time-Integrated and Time-Resolved Results

The velocity measurements obtained using time-resolved and time-integrated DGV include a common test point — delivery of a 1 mg payload of 4.7 μm microspheres through a conical nozzle — which is in fact the major test case for time-resolved DGV. This commonality permits a systematic comparison of the results obtained with both methods. There are two motivations for such a comparison. Firstly, this comparison is a test for consistency between the variants of the DGV technique, and, in the event of a positive outcome, a means of bolstering confidence in the experimental results as a whole. The comparison is also an opportunity to add a specific empirical assessment of biasing effects in time-integrated DGV of unsteady flows to the general theoretical discussion of §5.5.4, and assist in physical interpretation of the time-integrated results.

These two goals are quite distinct, and demand manipulation of the time-resolved results in different ways to render them directly comparable with the time-integrated results. In effect, two comparisons have been carried out. They are documented separately in §6.5.1 and §6.5.2.

6.5.1 Consistency

In order to make a meaningful comparison of time-resolved and tme-integrated DGV results, the biases which are inherent in time-integrated DGV must be modelled and applied to the time-resolved data. The precise meaning and significance of *time-integration* in a DGV experiment was discussed in §5.5.4. ***Eq. 5.22*** in that section, repeated below, relates the time histories of particle velocity and light scattering intensity to the reported time-integrated transmission ratio, $\bar{T}$:

$$\bar{T} = \frac{\int_0^\tau T(V(t))I(t)dt}{\int_0^\tau I(t)dt} \quad . \tag{6.3}$$

The bias of time-integrated DGV towards bright and slow particles is implicit in this equation, as explained in §5.5.4. In time-integrated DGV the time-varying data $I(t)$ and $V(t)$ are subsumed into the integrations above, and their details are lost. However, in time-resolved DGV, they are stored. The reference images and velocity maps yielded by the time-resolved experiments can be viewed as samples of the functions $I(t)$ and $V(t)$ at m points in time — that is, $\{I(t_1), I(t_2), \dots I(t_m)\}$ and $\{V(t_1),$

$V(t_2)$, ... $V(t_m)\}$ at each pixel. By constructing a discretised approximation to ***Eq. 6.3***, it is straightforward to process the time-resolved data in a way which mimics the processes of time-integrated DGV, as defined by ***Eq. 6.3***. The most obvious approximation is:

$$\bar{T} \approx \frac{T(t_1)I(t_1)\frac{(t_2-t_1)}{2} + \sum_{j=2}^{m-1} T(t_j)I(t_j)\frac{t_{j+1}-t_{j-1}}{2} + T(t_m)I(t_m)\frac{(t_m-t_{m-1})}{2}}{I(t_1)\frac{(t_2-t_1)}{2} + \sum_{j=2}^{m-1} I(t_j)\frac{t_{j+1}-t_{j-1}}{2} + I(t_m)\frac{(t_m-t_{m-1})}{2}} . \qquad (6.4)$$

In practice, this gave results which were indistinguishable from those of a more accurate equation based on the trapezoidal integration rule. The values $T(t_j)$ are calculated from velocity data, and $\bar{T}$ is converted back to a velocity, by referring to the calibration of the iodine cell in the time-integrated DGV system. This process was applied on a pixel-by-pixel basis to the reference images and velocity maps obtained in 31 time-resolved DGV measurements of the velocity field of 1.0 mg of 4.7 μm particles, delivered through a conical nozzle (the same data presented in §6.4.2).

The final synthetic time-integrated velocity map is shown in ***Figure 6.42*** along with a true time-integrated DGV measurement taken at the same test condition (with an opaque nozzle). Agreement between the two velocity maps is excellent. In other words, the time-resolved data, when artificially subjected to the effects which bias the time-integrated DGV measurements, yield a result which is essentially identical to an actual time-integrated measurement. This very strongly indicates that the time-resolved and time-integrated measurement systems function in precisely the same way on a fundamental level (as they should), and differ only because of the unavoidable biasing effects which are caused by the integration of time-varying images in time-integrated DGV. The two velocity maps shown in ***Figure 6.42*** were obtained using different procedures, different imaging systems and iodine cells, different frequency standards for iodine cell calibration, and radically different lasers. Their consistency must therefore be seen as a strong endorsement of both versions of the DGV technique.

Having established the consistency of the time-integrated and time-resolved results, the contrasts between the time-integrated and time-resolved measurements can be discussed in terms of the observed velocity fields. A typical time-integrated velocity map takes on the appearance of a steady subsonic jet, with gradually varying cross-sectional velocity profiles and gradual axial decay of velocity, accompanied by radial growth of the jet. Instantaneous velocity field histories depict the brief passage of a wide low-velocity leading cloud followed by a quasi-steady jet which possesses striking uniformity in both axial and transverse directions. Later, velocity and particle concentrations in the jet begin to decay with time, and the weakening jet dissipates into the atmosphere.

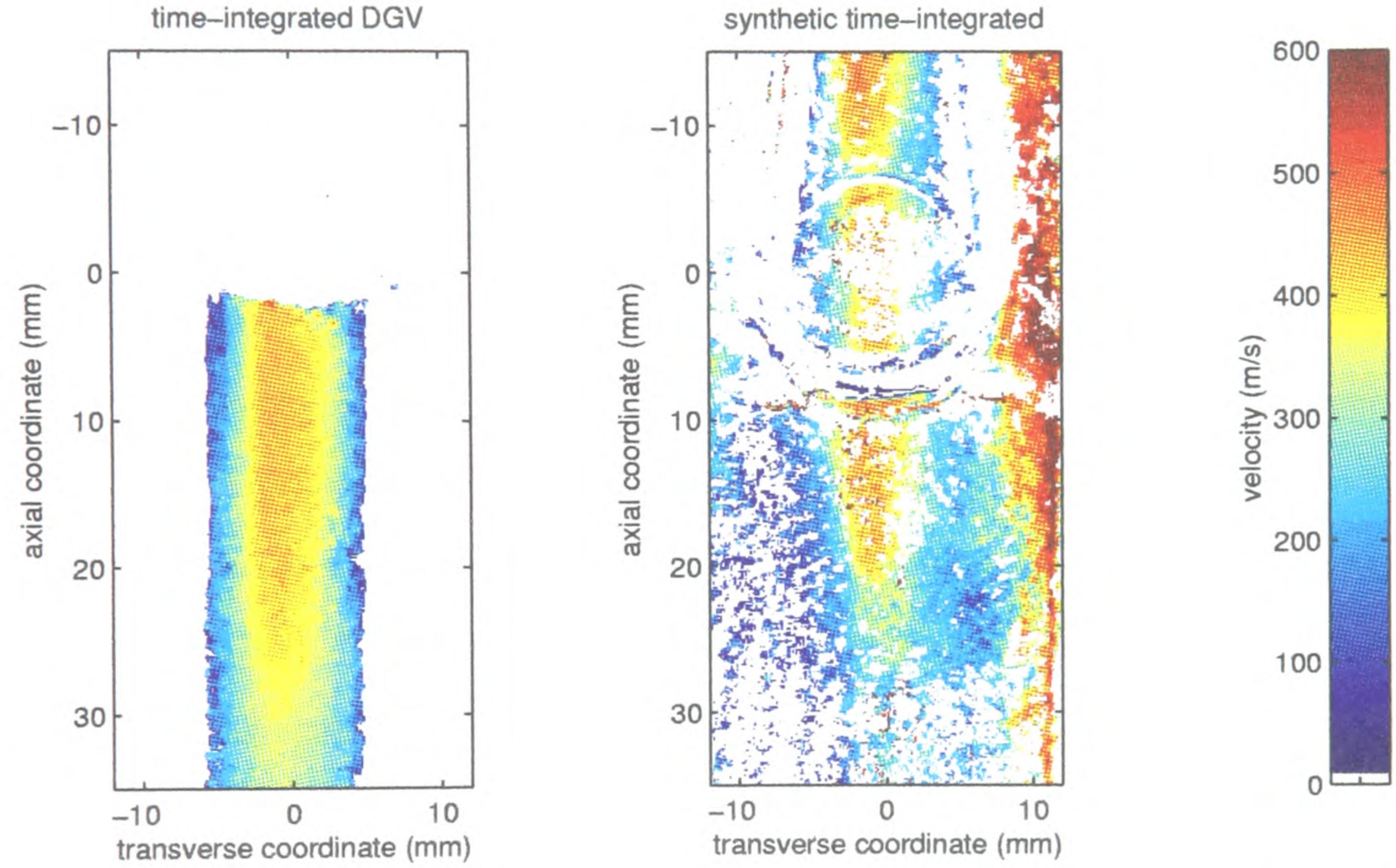


Figure 6.42 (*Left*) a time-integrated DGV measurement for 1.0 mg of 4.7 μm polystyrene microspheres, and (*right*) a synthetic time-integrated velocity field derived from time-resolved results, manipulated in a way which mimics the biasing inherent in time-integrated DGV.

The dissipative character of the time-integrated measured velocity field jet is probably due in part the late stages of the unsteady jet, in which a small flux of low-velocity particles is delivered for a long period. The leading particle cloud may also be an influence on the apparent cross-sectional profile of the time-integrated field, sweeping out a band of low velocity which extends beyond the width of the quasi-steady jet.

6.5.2 Assessment of Biasing in Time-Integrated DGV

It was pointed out in §5.5.4 that time-integrated DGV measurements of velocity in unsteady flows are inherently biased towards the largest and slowest particles in the flow-field, and may be skewed by strong non-linearity in the iodine cell transmission profile. In this section, to empirically examine the significance of these biasing effects, a time-integrated DGV measurement of a particle velocity field is compared with a single velocity field which is synthesised from the time-resolved data set and is representative of the complete flow-field history in a simple and physically meaningful way.

Time-resolved measurements are not affected by the non-linearity of the iodine cell profile, and are free of bias towards slow particles, although they are biased towards large (bright) particles if the particle size distribution is wide. However, in all the time-resolved and time-integrated measurements referred to in this section, 4.7 μm microspheres with a narrow size range (0.6 μm standard deviation on diameter) were used. The time-resolved DGV data should therefore be free of all biases. These reliable data have been used to synthesise a time-integrated velocity field in which the velocity value at each pixel is a simple average of the velocity of all particles which pass through the space corresponding to that pixel over the course of an entire run. This *particle-averaged* velocity field could be considered an ideal time-integrated measurement, since it is probably the most useful and relevant summary of a complete particle velocity field history which can be presented in the form of a single spatial map. The particle-averaged form of the time-resolved measurements will be compared with a time-integrated velocity measurement to gain direct experience of the physical meaning of the time-integrated DGV results.

If, in the course of a run, N particles pass through the small, well-defined, unique measurement volume associated with a particular pixel, then the particle-averaged velocity $\bar{V}_p$ for that pixel is defined by ***Eq. 6.5***.

$$\bar{V}_p = \frac{1}{N}\sum_{i=1}^{N} V_i \tag{6.5}$$

To compute this sum over particles, a means of estimating particle numbers (at least in relative terms) is required. If the length (in the flow direction) of the control volume associated with a particular pixel is s, and $V(t)$ is the velocity at which particles pass through the volume, then there is time for $\frac{V(t)\Delta t}{s}$ particles to traverse the volume in an interval Δt. To obtain the total number of particles which pass through during the interval of interest, this must be multiplied by the average number of particles inside the control volume during that interval. For this purpose, it is assumed that the number of particles in the pixel's control volume at any time is proportional to the instantaneous intensity I of the light received at the pixel (this assumption rests on the conditions that particle size is uniform throughout the flow, and that all the light which arrives at a pixel is scattered directly to the pixel from a particle in the control volume, without obscuration or re-scattering by any other particles). Therefore, the total number of particles which are imaged on a given pixel in a short time is:

$$n(t, \Delta t) = \frac{V(t)\Delta t}{s} kI(t) \ , \tag{6.6}$$

where k is a constant of proportionality, ultimately depending on the physics of light scattering and the optical characteristics of the imaging system.

To exploit the discrete time-resolved DGV data, the time history is carved up into intervals from $t = \frac{t_{j-1} + t_j}{2}$ to $t = \frac{t_j + t_{j+1}}{2}$, and the image intensity field and the velocity field are assumed to remain constant over each such interval, in the state recorded experimentally at $t = t_j$. The duration of the jth interval is $\Delta t_j = \frac{t_{j+1} - t_{j-1}}{2}$. ***Eq. 6.6***, in this framework, becomes ***Eq. 6.7***. ***Eq. 6.5*** can be rewritten as ***Eq. 6.8***, replacing the sum over particles with a sum over the m time intervals.

$$n_j = \frac{V(t_j)\Delta t_j}{s} kI(t_j) \tag{6.7}$$

$$\bar{V}_p \approx \frac{\sum_{j=1}^{m} n_j V(t_j)}{\sum_{j=1}^{m} n_j} \tag{6.8}$$

By substituting ***Eq. 6.7*** into ***Eq. 6.8***, introducing special treatments for the first and last time points, and cancelling the constants k and s, this estimate of the particle-averaged velocity is obtained:

$$\bar{V}_p \approx \frac{V(t_1)^2 I(t_1)\frac{(t_2 - t_1)}{2} + \sum_{j=2}^{m-1} V(t_j)^2 I(t_j)\frac{t_{j+1} - t_{j-1}}{2} + V(t_m)^2 I(t_m)\frac{(t_m - t_{m-1})}{2}}{V(t_1) I(t_1)\frac{(t_2 - t_1)}{2} + \sum_{j=2}^{m-1} V(t_j) I(t_j)\frac{t_{j+1} - t_{j-1}}{2} + V(t_m) I(t_m)\frac{(t_m - t_{m-1})}{2}} \ . \tag{6.9}$$

In this form, $\overline{V}_p$ can readily be estimated for every pixel, given time-resolved warped reference images (distributions of $I(t_j)$) and the corresponding velocity maps (distributions of $V(t_j)$), to yield a map of particle-averaged velocity. This process has been carried out, using the time-resolved data of §6.4.2, to generate the velocity field shown in the right-hand side of ***Figure 6.43***. To recap, the velocity at each pixel in this field is an estimate of the average velocity of all particles which pass through the flow-field region represented by that pixel in the course of a complete flow history. It is a straightforward average, free of any weighting towards particle size or velocity, and is independent of the iodine cell transmission profile.

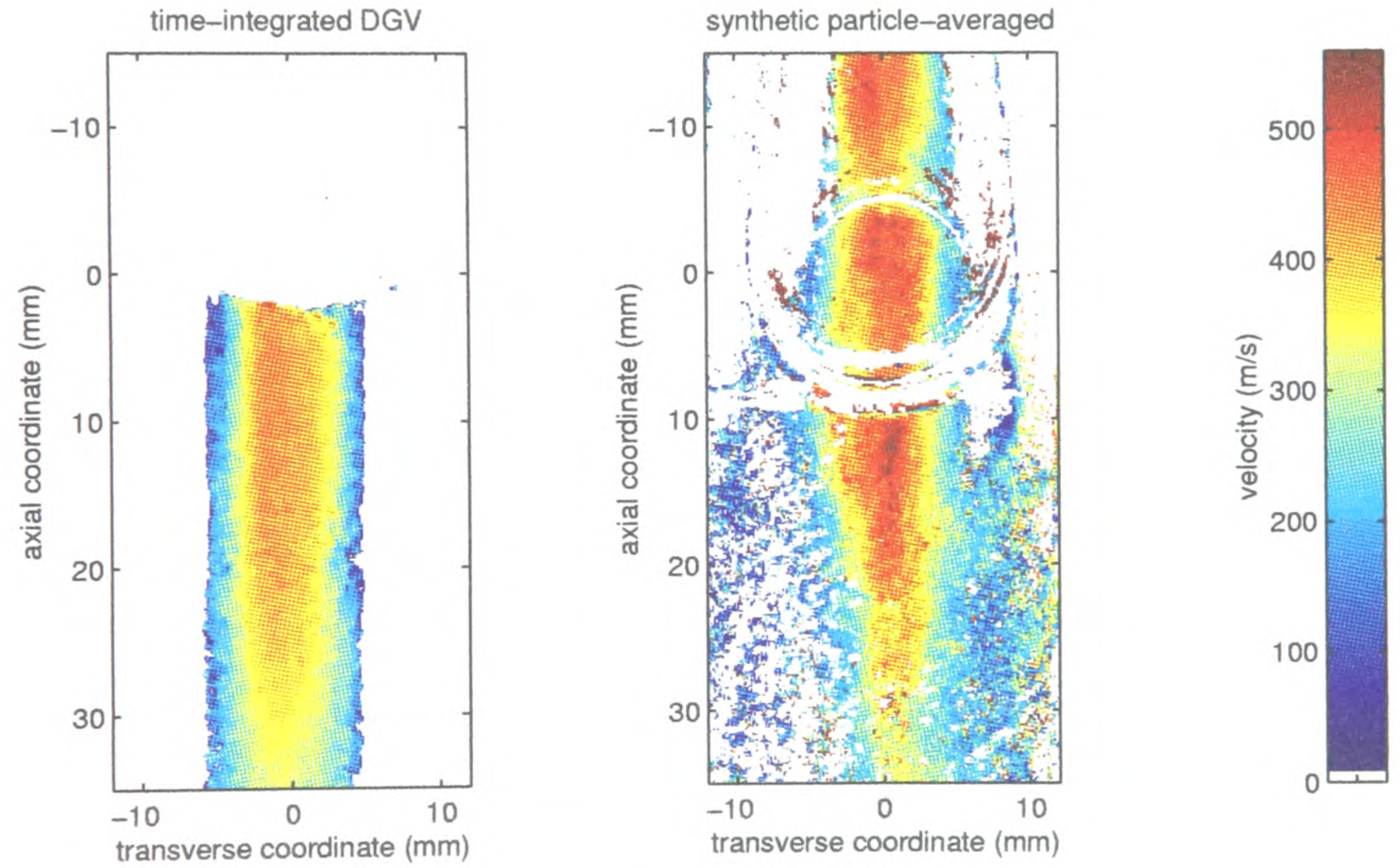


Figure 6.43 (*Left*) a time-integrated DGV velocity measurement for 1.0 mg of 4.7 μm polystyrene spheres. (***Right***) a velocity field generated from time-resolved measurements of a nominally similar flow, equivalent to an unbiased average of velocity over all particles which pass through each pixel.

On the left side of ***Figure 6.43***, the result of an appropriate time-integrated velocity measurement is shown for comparison. The two velocity maps are similar in overall form. The velocities measured in the time-integrated experiment are lower than the computed particle-averaged velocity by up to 100 m/s in the region from 10 to 20 mm downstream of the nozzle exit. Elsewhere, there is good agreement between the velocity maps. This is in line with the prediction that time-integrated measurements are biased towards slow particles, even if particle size is uniform. This result confirms the theoretical understanding of the application of time-integrated DGV to unsteady flows, and suggests that biasing effects in time-integrated DGV can be significant, though they are far from overwhelming in this case.

6.6 Conclusions

Time-integrated DGV measurements have been carried out for payloads of polystyrene microspheres, lidocaine and silica in devices with contoured and conical nozzles. In all the measured velocity fields, velocity falls off gradually with distance from the centreline and distance downstream of the exit. Flow from the contoured nozzle is remarkably uniform across the exit, suggesting a clean flow inside the device, though further downstream, it develops a more rounded profile. The measured exit velocity of microspheres delivered from the contoured nozzle ranges from 750 – 1050 m/s, decreasing with increasing particle size and also with increasing payload mass. These findings indicate that the particles lag the gas at the nozzle exit, and that the payload strips a substantial proportion of the gas's momentum. The velocity of lidocaine particles is lower, and does not vary significantly with payload mass or particle size. The independence of particle size suggests that particles become broken up in the flow. Exit velocities up to 1160 m/s were measured for silica. The magnitude of velocities measured, in general, indicate that quasi-steady supersonic flow is a major mechanism of particle acceleration in this device, although the transient starting process may also have a role.

Much lower time-integrated velocities of 300 – 500 m/s were measured for microspheres and lidocaine particles exiting the conical nozzle, showing no dependence on particle size or payload mass. The velocity values support a theory that the gas exit velocity is low because of a stationary shock or shock system inside the nozzle, and that consequently, all particles are capable of reaching the gas velocity.

Time-resolved measurements provided a revealing description of the flow-field in the delivery of microspheres from the conical nozzle. Particles are delivered in two stages. A wide, highly concentrated cloud of particles emerges first, at 200-400 m/s, followed by a narrow slowly-varying jet, which pours into the back of the leading cloud at 650-800 m/s. The leading cloud is thought to be associated with the gas starting process in the gas flow, while the jet is consistent with a shocked, separated quasi-steady nozzle flow. The leading cloud is fed, and to some extent driven forwards, by the quasi-steady stream. In measurements far upstream in the nozzle, the separation of the jet from the wall is clearly visualised. The instantaneous velocity fields of lidocaine and silica payloads are essentially similar to those measured for microspheres, though they suggest widespread local non-uniformities in the flow.

The time-resolved measurements of velocity at the front of the particle cloud agree well with the velocity estimated by tracking the cloud front from shot to shot, within the limits of the estimate's accuracy. It was also shown that the time-resolved results are in close agreement with the time-integrated measurements, when the inevitable biasing in time-integrated DGV is accounted for by an analytical model.

The time-resolved data were also used to synthesise an estimate of the velocity field averaged without bias over all particles. A comparison of the resulting particle-averaged velocity field with a time-integrated result suggested that in these measurements, time-integrated DGV underestimates average particle velocity by up to 100 m/s (approximately 20%), confirming that caution is required in the use of the time-integrated technique with unsteady flow. This discrepancy represents the effects of biasing in time-integrated DGV.

Chapter 7

Numerical Modelling

Mathematical models of the flow of gas and particles in the drug delivery device have been developed as part of this research, with the aim of producing a software code capable of predicting device performance. There have been several motivations for this activity. Perhaps most obviously, a predictive tool is attractive as an aid to future device design. The task of developing a numerical model from first principles is also useful as a structure for formalising, testing and refining the understanding of fluid dynamics in the drug delivery devices. This understanding is the main goal of this thesis. Computed predictions can also provide new insights into certain aspects of device operation. This is where they can potentially be most valuable, by providing a complete description of the flow-field, where even the most powerful of experiments can yield information about a limited number of variables in a limited region. Use of computations in this way must be backed up with evidence that they are representative of the important physical phenomena, even if their numerical accuracy is not excellent.

Constraints on the time and effort available for this modelling work influenced the objectives of the project. The capability to correctly predict trends in device operation, while predicting parameters to moderate accuracy, is a realistic target. Ideally, the software implementation of the model should be easy to use and computationally inexpensive. It should require as few inputs as possible, other than the values of physical parameters which describe the geometry and operating conditions of a real device. In short, the goal is a set of simple, robust, workable engineering approximations to the flow in transdermal drug delivery devices.

These objectives rule out Computational Fluid Dynamics (CFD) in the usual sense of the term, meaning a numerical solution of the full Navier-Stokes or Euler equations in two or three dimensions. Instead, the models used here take advantage of analytical theories and empirical correlations which describe simplified flows similar to the real flows of interest. The development of the model has been a process of successively refining the set of assumptions regarding the flow physics to achieve better agreement with experiments.

With one (arguably justifiable) exception, numerical parameters in the models have not been tuned to force agreement with experimental results. Excessive reliance on such tuning (effectively, calibration) would result in a model of limited applicability, which might become little more than a curve fit to experimental results.

The drug delivery device is divided into two subsystems for modelling purposes. The reservoir system, consisting of the cylinder, valve and rupture chamber, is dealt with in §7.1. Modelling of the second system, which consists of the gas and particle flows in the nozzle, is described in §7.2. In each section, the mathematical model and its numerical implementation is described, and its predictions are presented and compared with experimental results.

The division into two systems is permitted by the assumption that the rupture chamber contains gas which is essentially stagnant, flowing slowly and isentropically into the nozzle. The static pressure and temperature in the rupture chamber can then be taken as the total pressure and temperature for flow in the nozzle. The task of the reservoir model, then, is to determine the time history of total pressure and temperature for the nozzle flow by analysis of the cylinder and rupture chamber flows. The task of the nozzle flow model is to compute the flow through the nozzle, and the behaviour of particles in that flow, given the total conditions prescribed by the reservoir model. In both areas, the assumption of quasi-steady flow underlies all modelling. The complexity of the strongly unsteady early flow is beyond the scope of the present work.

7.1 Gas Dynamics of the Reservoir System

7.1.1 Mathematical Model

A schematic diagram of a generalised reservoir system is shown in ***Figure 7.1***. The complex geometries of the real system are idealised as a pair of chambers, labelled c (cylinder) and r (rupture chamber), which are linked by an opening labelled v for the valve. The rupture chamber has a second opening, t, the nozzle throat. The gas within each chamber is assumed to be uniform and stationary, implying that total temperature and pressure are equal to static temperature and pressure, respectively. This assumption has been checked for the rupture chamber by treating it as an upstream extension of the nozzle and assuming quasi-one-dimensional isentropic flow. With these assumptions, the highest possible Mach number in the rupture chamber occurs with the conical nozzle (which has a larger throat than the contoured nozzles) when the nozzle throat is choked. This maximum Mach number in the rupture chamber is 0.143. The corresponding error in equating total properties to static properties is then 1.7% for pressure, and less for temperature.

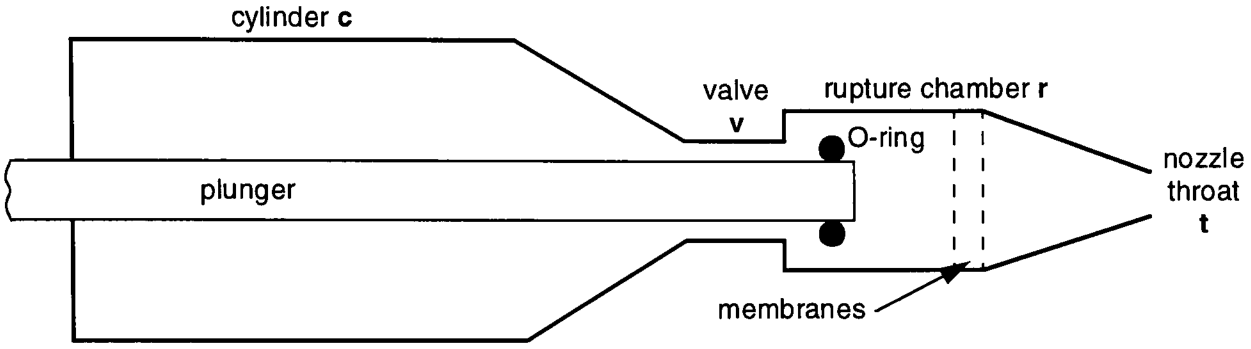


Figure 7.1 A schematic diagram of the reservoir system, showing the valve in the open position.

The overall modelling strategy is to derive equations relating the mass in each chamber to that chamber's pressure, and equations which relate the mass flow rate through each opening to the pressure in one or both chambers. The system of equations is closed by trivial differential relationships between mass and mass flow rate, the ideal gas law, and thermodynamic relations. The details of this model are presented in the paragraphs which follow. Equations which are given below without proof can readily be derived from the area-Mach number relation for quasi-one-dimensional isentropic flow (***Eq. 2.6***), the isentropic relationships between total and static thermodynamic properties (***Eqs. 2.4*** and ***2.5***), and the ideal gas law, $p = \rho RT$.

The mass of gas in each chamber is a simple function of temperature, pressure, and the chamber volume, as expressed in ***Eq. 7.1*** and ***Eq. 7.2***.

$$p_c = \frac{V_c}{RT_c m_c} \tag{7.1}$$

$$p_r = \frac{V_r}{RT_r m_r} \tag{7.2}$$

Assuming isentropic flow through the cylinder outlet valve and the nozzle throat, the mass flow rate through either orifice can be expressed as a function of Mach number, open area A and the total conditions, as in ***Eq. 7.3***:

$$\dot{m} = \rho A u = p_0 A \sqrt{\frac{\gamma}{R T_0}} M \left(1 + \frac{\gamma - 1}{2} M^2\right)^{-\frac{(\gamma+1)}{2(\gamma-1)}} . \tag{7.3}$$

It is convenient to rewrite the Mach number in this equation in terms of the ratio of total pressure to static pressure, giving the following expression:

$$\dot{m} = A \sqrt{\frac{2\gamma}{R(\gamma-1)}} \left(\frac{p_0}{p}\right)^{\frac{(\gamma+1)}{2\gamma}} \sqrt{\frac{p_0}{p} - 1} \frac{p_0}{\sqrt{T_0}} . \tag{7.4}$$

In the case of the cylinder outlet valve, flow is assumed to emerge into the rupture chamber as a free jet. If the valve flow is subsonic, static pressure in the valve matches the pressure in the rupture chamber, p_r. It can be shown that isentropic flow through the valve (or any orifice) is sonic when the local total-to-static pressure ratio exceeds a critical level defined by $p_0/p \geq \left(\frac{\gamma+1}{2}\right)^{\frac{\gamma}{\gamma-1}}$. For helium ($\gamma = 5/3$), this is 2.05. If the pressure ratio is raised above the critical value, the valve flow becomes independent of the downstream pressure, and the Mach number remains at 1. Two relationships are therefore required for valve mass flow, corresponding to subsonic and sonic (or choked) flow. They are derived from ***Eq. 7.3*** and ***Eq. 7.4***, and both are given below in ***Eq. 7.5***.

$$\dot{m}_v = \begin{cases} A_v \sqrt{\frac{2\gamma}{R(\gamma-1)}} \left(\frac{p_c}{p_r}\right)^{\frac{(\gamma+1)}{2\gamma}} \sqrt{\frac{p_c}{p_r} - 1} \frac{p_c}{\sqrt{T_c}} & \text{for } \frac{p_c}{p_r} \leq \left(\frac{\gamma+1}{2}\right)^{\frac{\gamma}{\gamma-1}} \quad \text{(subsonic)} \\ A_v \sqrt{\frac{\gamma}{R}\left(\frac{\gamma+1}{2}\right)^{-\frac{\gamma+1}{\gamma-1}}} \frac{p_c}{\sqrt{T_c}} & \text{for } \frac{p_c}{p_r} \geq \left(\frac{\gamma+1}{2}\right)^{\frac{\gamma}{\gamma-1}} \quad \text{(choked)} \end{cases} \tag{7.5}$$

Cylinder pressure p_c is the total pressure in the valve flow, and rupture chamber pressure p_r is the exit static pressure for subsonic flow. In practice, the value of p_c used to compute the valve mass flow is reduced by an empirical factor which represents friction and heat transfer effects in the valve. This is explained in detail in the next section.

The analysis of flow through the throat differs slightly. For subsonic throat Mach numbers, it is convenient to describe downstream conditions in terms of the flow at the nozzle exit, since the nozzle exit pressure p_e is always known — it is simply the atmospheric pressure, p_a. Mass flow is then calculated through the nozzle exit, rather than the throat itself. This is a valid approach because in quasi-steady flow, mass flow rate is the same at all axial locations in the nozzle.

At the critical condition for choking of the throat, flow expands isentropically and subsonically from Mach number 1 at the throat to atmospheric pressure p_e and Mach number M_{e-} at the exit. As shown in §2.1.2 (***Eq.** 2.7*), the critical value for the pressure ratio p_r/p_e can be defined concisely in terms of M_{e-}. The expression for choked flow is similar to that for cylinder outlet valve. The two expressions for mass flow through the throat are as follows:

$$\dot{m}_t = \begin{cases} A_e\sqrt{\dfrac{2\gamma}{R(\gamma-1)}}\left(\dfrac{p_r}{p_a}\right)^{\frac{(\gamma+1)}{2\gamma}}\sqrt{\dfrac{p_r}{p_a}-1}\;\dfrac{p_r}{\sqrt{T_r}} & \text{for}\quad \dfrac{p_r}{p_a}\le\left(1+\dfrac{\gamma-1}{2}M_{e-}^2\right)^{\frac{\gamma}{\gamma-1}}\text{(subsonic)} \\ A_t\sqrt{\dfrac{\gamma}{R}\left(\dfrac{\gamma+1}{2}\right)^{-\frac{\gamma+1}{\gamma-1}}}\;\dfrac{p_r}{\sqrt{T_r}} & \text{for}\quad \dfrac{p_r}{p_a}\ge\left(1+\dfrac{\gamma-1}{2}M_{e-}^2\right)^{\frac{\gamma}{\gamma-1}}\text{(choked)} \end{cases} \tag{7.6}$$

The mass flow rates defined above by ***Eqs.*** *7.5* and *7.6* are related to the mass of gas in each chamber by the following differential equations:

$$\frac{dm_c}{dt} = -\dot{m}_v \tag{7.7}$$

$$\frac{dm_r}{dt} = \dot{m}_v - \dot{m}_t\ . \tag{7.8}$$

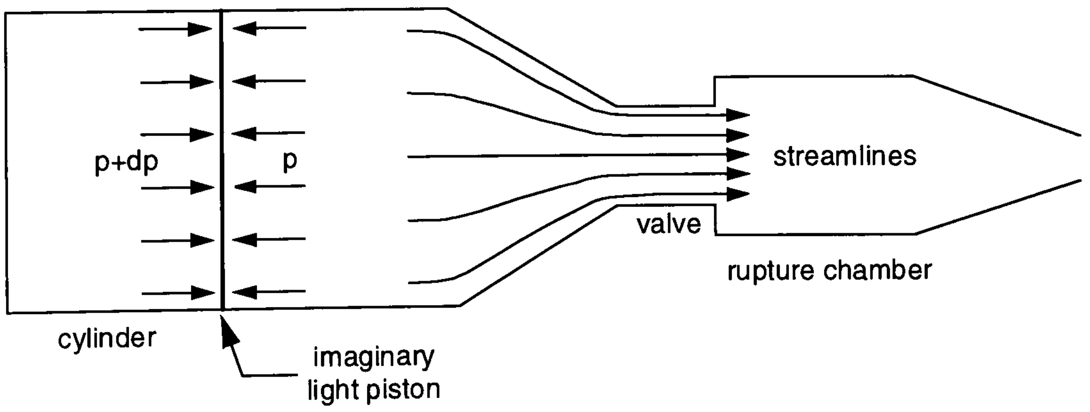


***Figure** 7.2* Simplified diagram of flow in the cylinder.

It remains to model the temperature variations in each chamber. The situation in the cylinder is represented in ***Figure** 7.2*, which shows an imaginary light piston dividing the cylinder. The gas far downstream of this arbitrary divide is beginning to accelerate into the nozzle. The gas upstream of the piston is essentially static, and destined to stay in the cylinder for a long time. The upstream gas does work on the piston (in reality, an arbitrary boundary surface in the gas) to accelerate the downstream gas. The pressure difference across the piston is infinitesimal, so this work is a reversible thermodynamic process. Neglecting heat transfer across the piston, it is also adiabatic. It is therefore isentropic, and temperature in the stationary gas (which, in reality, almost fills the cylinder) is related to pressure by the isentropic relation $Tp^{\frac{-\gamma}{\gamma-1}} = \text{const}$, which can be differentiated to give ***Eq. 7.9*** below.

$$dT_c = \frac{(\gamma-1)}{\gamma}T_c\frac{dp_c}{p_c} \tag{7.9}$$

Similar arguments hold for stationary gas in the rupture chamber, though as well as doing work (on gas flowing into the nozzle), it is worked upon (by gas flowing in from the cylinder, tending to compress it). However, temperature in the rupture chamber is also influenced by the temperature of incoming gas from the cylinder, which mixes with it. Heat transfer in the valve is invoked to estimate the temperature of that gas. The valve consists of a slender annular duct, the walls of which begin the run at uniform temperature. It has a large surface exposed area to the gas, which flows at Mach 1 for much of the run. The heat transfer coefficient between the valve and the gas should therefore be high. This suggests the approximation that gas emerges from the valve at the same temperature as the steel in the valve structure. The most straightforward assumption for the temperature of the metal is that it remains at ambient temperature. The inflowing gas is then assumed to mix fully and instantaneously with the gas in the rupture chamber. In practice, this set of assumptions has given better results than the other possible simplistic assumption, namely that no heat transfer takes place.

Since it is a basic feature of this model that gas flowing from the valve comes to rest in the rupture chamber, it would be natural to assume that the gas mixes out and stagnates at a temperature equal to the total temperature at the valve exit. However, it is assumed in the model that the gas mixes and comes to rest at its inflow static temperature (that is, ambient temperature). This assumption is not well grounded in theory, but appears to give results that are more consistent with the experimentally observed behaviour of the system. This claim is supported by some additional analysis in §7.1.2.

The two influences on rupture chamber temperature — isentropic work and mixing out of the inflowing gas — are expressed in ***Eq. 7.10***, where T_a denotes ambient temperature.

$$dT_r = \frac{\dot{m}_v}{m_r}(T_a - T_r)dt + \left(\frac{\gamma-1}{\gamma}\right)T_r\frac{dp_r}{p_r} \tag{7.10}$$

Eqs. 7.1, ***7.2*** and ***7.6*** to ***7.10*** constitute a system of 8 equations in 8 unknowns, which fully describe the quasi-steady flow from the cylinder, through the rupture chamber and into the nozzle. Additional special treatments are required to simulate the early stages of flow, which set up initial conditions for the quasi-steady history. Initially, the unbroken membranes seal off the nozzle inlet, and mass flow through the nozzle throat must be held at zero. As gas flows in through the cylinder valve, pressure builds up in the rupture chamber until it exceeds membrane burst pressure. After this point, in reality, the rupture chamber and the drug cassette behave as a short shock tube. The assumed end result of the ensuing processes is that gas in the rupture chamber expands into the cassette volume, doing no work and experiencing no heat transfer (its interaction with air in the cassette is neglected). Therefore, the internal energy (and temperature) of this gas remain constant. Pressure drops according to ***Eq. 7.11***, in which the subscripts 1 and 2 signify states before and after burst, respectively.

$$\frac{p_{r,2}}{p_{r,1}} = \frac{\rho_{r,2}RT_{r,2}}{\rho_{r,1}RT_{r,1}} = \frac{V_{r,1}}{V_{r,2}} \qquad (7.11)$$

When rupture chamber pressure again rises above burst pressure, the downstream membrane bursts and the rupture chamber becomes a shock tube for the second time. In this case, with no obstacle immediately downstream of the membrane, significant mass flow from the rupture chamber can occur. This flow is short-lived and is soon replaced by quasi-steady mass flow driven by the pressure ratios across the valve and the nozzle. For modelling purposes, the duration of the shock tube flow is estimated as the time taken for an expansion wave to travel (at the local speed of sound) to the back wall of the rupture chamber, and after reflection, back to the membrane location. During this time, gas is assumed to flow through the rupture chamber exit plane with the density and contact surface velocity predicted by standard shock tube theory (as derived by Anderson [10], for example). The brief shock tube flow is modelled as a step change. Mass in the rupture chamber is decremented by the total outflow which occurs in this period, and density is adjusted accordingly. Isentropic relations are used to recalculate pressure and temperature, since an expansion wave is an isentropic process.

Throughout the early stages of flow, the composition of gas in the rupture chamber must be monitored, since it consists initially of air, but is progressively diluted in Helium. The ideal gas constant R and the constant volume specific heat c_v of the mixture are given by mass averages of the values for air and helium. This approach is consistent with a partial pressures model of the mixture. The specific heat ratio for the mixture is then calculated from the relationship $\gamma = 1 + \frac{R}{c_v}$.

7.1.2 Empirical Loss Coefficient for Flow in the Cylinder Outlet Valve

Mr. Mark Roberts [33], who carried out some simulations with an early version of the numerical model described here, found that agreement between computed and experimental results could be greatly improved, across the full range of test conditions, if the valve cross-sectional area used in the model was scaled by a factor of 0.645. This factor is equivalent to a discharge coefficient for the flow through the valve. The correction is retained in the current work, and is the one empirically calibrated feature of the model. However, new work is described in this section in which an additional numerical analysis is used to explore the physical significance of the discharge coefficient as a description of friction and heat transfer effects in the valve. The objective of this work is to present the discharge coefficient as a well-understood flow parameter which requires measurement for particular device configurations, rather than as a superficial description of unknown effects. To some extent, this work refutes objections to the use of an empirical input to the numerical model.

While the valve is choked, mass flow rate in isentropic flow is proportional to $A_v p_c / \sqrt{T_c}$. The application of a discharge coefficient, defined as a scale factor on the mass flow rate, is therefore exactly equivalent to a reduction in the value of the parameter $p_c / \sqrt{T_c}$, the ratio of cylinder pressure to the square root of cylinder temperature. These quantities appear in this context as the total pressure and total temperature in the choked valve flow, and it can be expected that friction and heat transfer act in the valve flow to reduce total pressure and increase total temperature. The discharge coefficient introduced above can therefore be interpreted in a physically meaningful way as the ratio of $p_0 / \sqrt{T_0}$ at the valve outlet to $p_c / \sqrt{T_c}$ (which is the value of $p_0 / \sqrt{T_0}$ at the valve inlet).

The flow through the valve has been approximately modelled at a range of conditions to investigate changes in total pressure and total temperature. This project is distinct from the main modelling of mass flows in the cylinder and rupture chamber. Its aim is to assess whether the combined influence of friction and heat transfer is similar in magnitude to the influence of the empirically evaluated discharge coefficient, and whether the effect is constant over a range of operating conditions, as the empirical coefficient appears to be. This exercise is an investigation, rather than a more ambitious attempt to predict useful values of the discharge coefficient.

The open valve is essentially an annular duct, 11 mm long, with inner and outer diameters of 4.7 and 5.0 mm, respectively. Its hydraulic diameter, D_h, is therefore 0.3 mm. Differential relations for one-dimensional compressible flow through a constant-area duct with friction and heat transfer are shown below as ***Eqs. 7.12*** and *7.13*. They are adapted from Shapiro's text [11].

$$\frac{dM^2}{M^2} = \frac{\left(1 + \frac{\gamma - 1}{2}M^2\right)}{1 - M^2}\left[4\gamma M^2 f \frac{dx}{D_h} + (1 + \gamma M^2)\frac{dT_0}{T_0}\right] \qquad (7.12)$$

$$\frac{dp_0}{p_0} = -\frac{\gamma M^2}{2}\left(4f\frac{dx}{D_h} + \frac{dT_0}{T_0}\right) \qquad (7.13)$$

These equations specify the changes in Mach number and total pressure over an infinitesimal duct length, *dx,* as functions of the change in total temperature T_0, due to heat transfer, and the local friction factor $f = 2\tau_w/(\rho u^2)$ (where τ_w is the wall shear stress). These equations have been implemented numerically by marching in the flow direction (positive x) from the valve inlet, replacing differentials dM^2, dx, etc. with step differences $\Delta(M^2)$, Δx, etc. At each step, the friction factor is calculated from this correlation for fully developed turbulent flow, given by White [60] and originally due to Haaland:

$$\frac{1}{f^{1/2}} = -1.8\,\log_{10}\left(\frac{6.9}{Re_D}\right) . \qquad (7.14)$$

The Reynolds number Re_D is based on the hydraulic diameter, with an adjustment for the geometry of the annulus, also given by White.

The change in total temperature over a given length is related to heat transfer according to ***Eq. 7.15***, where c_p is the constant pressure specific heat capacity of the gas, dQ is the infinitesimal heat transfer to the fluid across a wall area dA_{ht}, h is the heat transfer coefficient, and T_s is the wall temperature.

$$\dot{m}c_p dT_0 = dQ = h(T_s - T)dA_{ht} \qquad (7.15)$$

The heat transfer coefficient, in turn, comes from a correlation due to Sieder and Tate for fully developed turbulent flow, given by Incropera and DeWitt [61], and given below as ***Eq. 7.16***. Nu_D represents the Nusselt number, k is the thermal conductivity of helium, Pr is the Prandtl number for helium, μ is its viscosity at flow conditions, and μ_s is the viscosity at wall temperature. Sutherland's law [11] is used to calculate viscosity as a function of temperature.

$$Nu_D = \frac{hD_h}{k} = (0.027)Re_D^{4/5}Pr^{1/3}\left(\frac{\mu}{\mu_s}\right)^{0.14} \tag{7.16}$$

To complete the model for heat transfer, it is assumed that the wall temperature remains equal to ambient temperature. This assumption is crude, but is probably the best approximation which avoids the further complication of conduction within the metal.

The equations above have been combined into a procedure for calculating the variation of flow along the length of the valve duct, given the total temperature, total pressure, and inlet Mach number. With a subsonic inlet flow, Mach increases along the duct. In choked flow, M rises to 1 precisely at the duct exit. The inlet Mach number must be guessed and iteratively adjusted to achieve this condition.

Results of the valve flow model are presented in ***Figure 7.3***, in the form of axial profiles of flow through the valve at various inlet conditions. The chosen inlet total pressures and total temperatures represent states of the gas in the cylinder at various instants in a typical run, as predicted using the main numerical model of reservoir system flow. It is immediately apparent that the effects of friction and heat transfer are important — total pressure drops by 25 to 30% across the valve, and the inlet Mach number is less than 0.5, although the exit flow is sonic. A detailed breakdown of the results shows that heat transfer is responsible for approximately 5% of the overall total pressure loss in cases of high inlet temperature (290 K), or as much as 45% in the case of the lowest inlet temperatures tested (134 K).

Total temperature rises monotonically through the valve because of positive heat transfer to the cool gas. Static temperature may rise or fall in the inlet region, depending on the inlet state, though it always falls off sharply near the valve exit. This reflects the relative importance of friction (which, in subsonic flow, always causes temperature to decrease) and (positive) heat transfer, which causes temperature to increase if Mach number is less than $\gamma^{-1/2}$, or to decrease at higher Mach numbers, as pointed out by Anderson [10]. For helium ($\gamma = 5/3$), this threshold Mach number is 0.775.

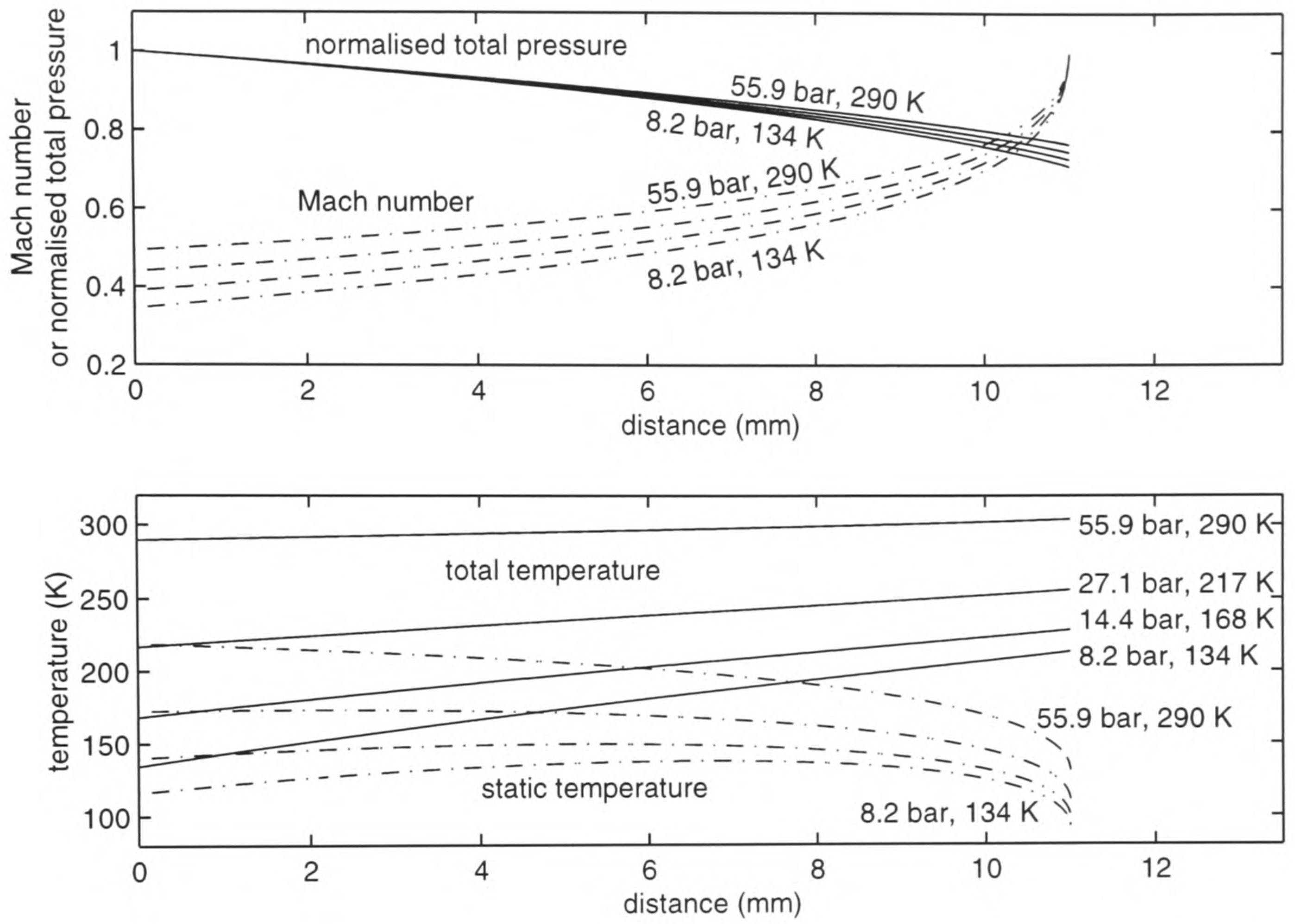


Figure 7.3 Computed axial profiles, at various times, of (*top*) total pressure (normalised to its inlet value) and Mach number, and (*bottom*) total and static temperature in choked flow through the cylinder outlet valve. Annotations on each profile indicate the instantaneous values of upstream (cylinder) total pressure and total temperature for which the profile is calculated.

These temperature profiles provide some support for the approximation (mentioned in §7.1.1) that the gas flowing from the valve comes to rest in the rupture chamber at ambient temperature. Effectively, this implies that the total temperature at the valve exit is equal to ambient temperature. In the valve flow profiles of ***Figure 7.3***, computed for an ambient temperature of 298 K, the total temperature rises monotonically to 200 – 300 K at the valve exit. The more obvious approximation that the static temperature approaches ambient temperature would be much less realistic, in view of the complex nature of static temperature variations in the valve. It should be emphasised, however, that the assumption regarding total temperature was adopted primarily on the basis that it yielded accurate predictions of rupture chamber pressure. It is not a meaningful description of the flow physics, although fortuitously, it does approximate the situation predicted by these computations.

To complete this analysis of valve flow, computations such as those shown in ***Figure 7.3*** can be used to calculate the discharge coefficient — the ratio of outlet to inlet values of $p_0/\sqrt{T_0}$ — which serves as a lumped parameter to describe friction and heat transfer. The time-varying value of this ratio has been calculated for a device with a 5 ml cylinder, a 0.2 ml rupture chamber, and a conical nozzle, at two fill pressures of 30 bar and 60 bar. The results of both sets of calculations are shown in ***Figure 7.4*** as a function of the unsteady cylinder pressure. The computed values range from 0.52 (for the lowest pressures) to 0.75 (at the highest pressures and earliest times).

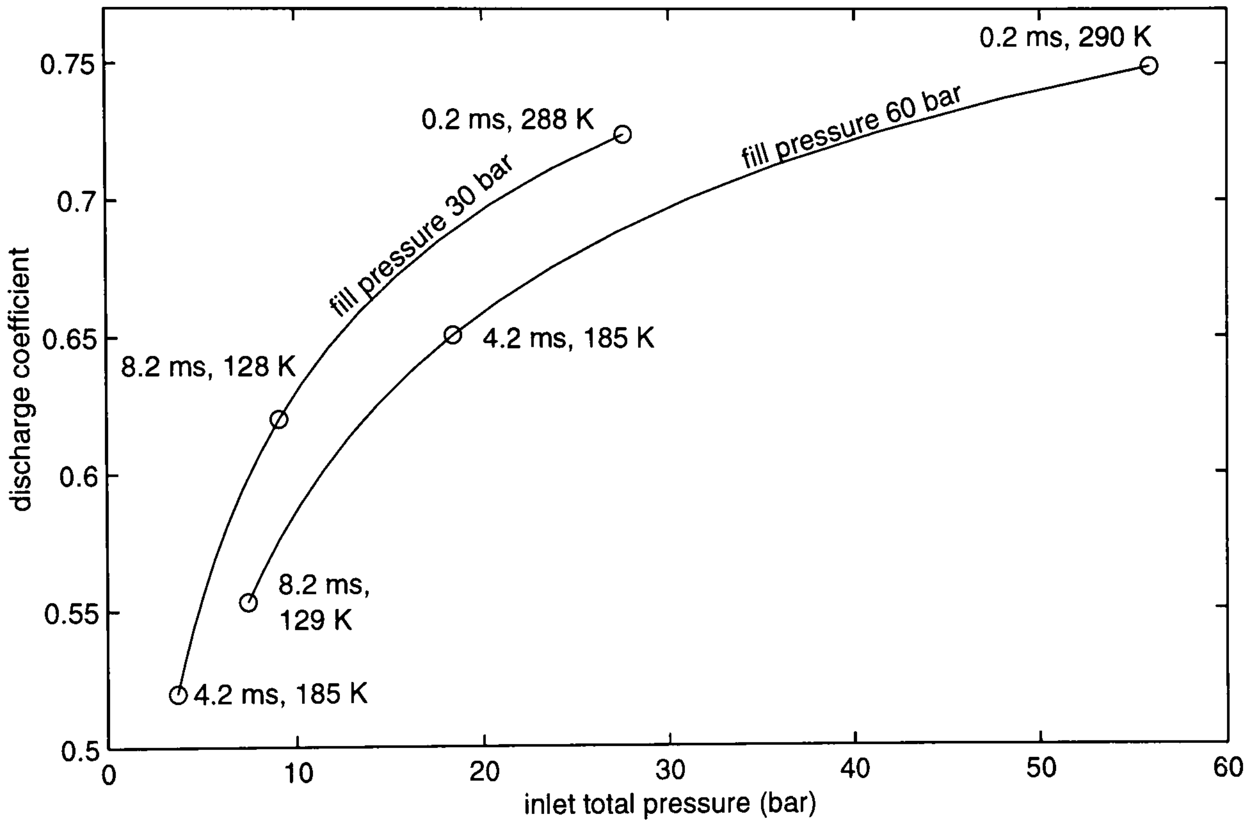


Figure 7.4 Computed discharge coefficients across the cylinder outlet valve, as a function of time-varying cylinder pressure, for fill pressures of 30 and 60 bar. Annotations show the time (from downstream membrane bursting) and cylinder temperature at various points.

To return to the stated objective of this study of the valve flow, the numerical predictions of valve flow *are* broadly consistent with experimental observations of a total pressure loss across the valve. The computed ratios of $p_0/\sqrt{T_0}$ across the valve, ranging from 0.52 to 0.75, are in rough agreement with the empirically evaluated (roughly constant) value of 0.645. This finding allows the empirical discharge coefficient (which is used routinely in the final version of the reservoir flow modelling code) to be viewed as a well-understood parameter which must be measured for a given device type, rather than as a crude description of unknown mechanisms.

7.1.3 Numerical Solution Method

The model of flow in the reservoir system, described in §7.1.1, has been implemented in a programme written in C. The program computes the gas mass, pressure and temperature for each chamber at successive times separated by a fixed step Δt. Of the eight governing equations, four are implemented directly without any approximation. The remaining four, ***Eqs. 7.7 – 7.10***, are differential equations which must be discretised.

An explicit time-marching scheme is used: a differential in any variable, say dp, required at time t_n, the beginning of the nth time step, is replaced with the difference $(p_{\mathrm{n+1}} - p_\mathrm{n})$. In each time step, the cylinder variables are updated first, followed by the rupture chamber variables. The code is computationally inexpensive and runs in a matter of seconds, so that very small time steps can be used to reduce discretisation error. The algorithm followed for the nth time step is outlined below:

- compute valve mass flow using ***Eq. 7.5***, with discharge coefficient
- update mass in cylinder using $m_{c,n+1} = m_{c,n} - \dot{m}_v \Delta t$ (discretised ***Eq. 7.7***)
- update cylinder pressure using ***Eq. 7.1***
- update cylinder temperature using $T_{c,n+1} = T_{c,n} + \left(\frac{\gamma - 1}{\gamma}\right) T_{c,n} \frac{p_{c,n+1} - p_{c,n}}{\frac{1}{2}(p_{c,n+1} + p_{c,n})} \Delta t$

  (discretised ***Eq. 7.9***)
- compute throat mass flow using ***Eq. 7.6***
- update mass in rupture chamber using $m_{r,n+1} = m_{r,n} + (\dot{m}_v - \dot{m}_{vt}) \Delta t$ (discretised ***Eq. 7.8***)
- update rupture chamber pressure $p_{\mathrm{n+1}}$ using ***Eq. 7.2***
- update rupture chamber temperature $T_{\mathrm{n+1}}$ using

$$T_{r,n+1} = T_{r,n} + \left[\frac{\dot{m}_v (T_a - T_{r,n})}{\frac{1}{2}(m_{r,n+1} + m_{r,n})} + \frac{\gamma - 1}{\gamma} T_{r,n} \frac{p_{r,n+1} - p_{r,n}}{\frac{1}{2}(p_{r,n+1} + p_{r,n})} \right] \Delta t \quad \text{(discretised } \textbf{\textit{Eq. 7.10}}\text{)}$$

- increment n and return to top of loop.

7.1.4 Comparison of Computed and Experimental Results

In this section, some computations of reservoir flow, obtained with the algorithm outlined above, are compared with measurements of rupture chamber pressure in a device fitted with a conical nozzle. These are the only available data which are appropriate for a direct comparison. In all three of the test cases examined here, the 0.2 ml rupture chamber and the conical nozzle were used with 20 μm polycarbonate burst membranes. There were three different combinations of fill pressure and cylinder volume: 60 bar with a 5 ml cylinder, the operating condition used in DGV experiments; 30 bar with a 5 ml cylinder; and 60 bar with a 20 ml cylinder. The latter two conditions represent extreme configurations of the device. The measured and predicted histories of rupture chamber pressure for these test cases are shown on two time scales in ***Figure 7.5***.

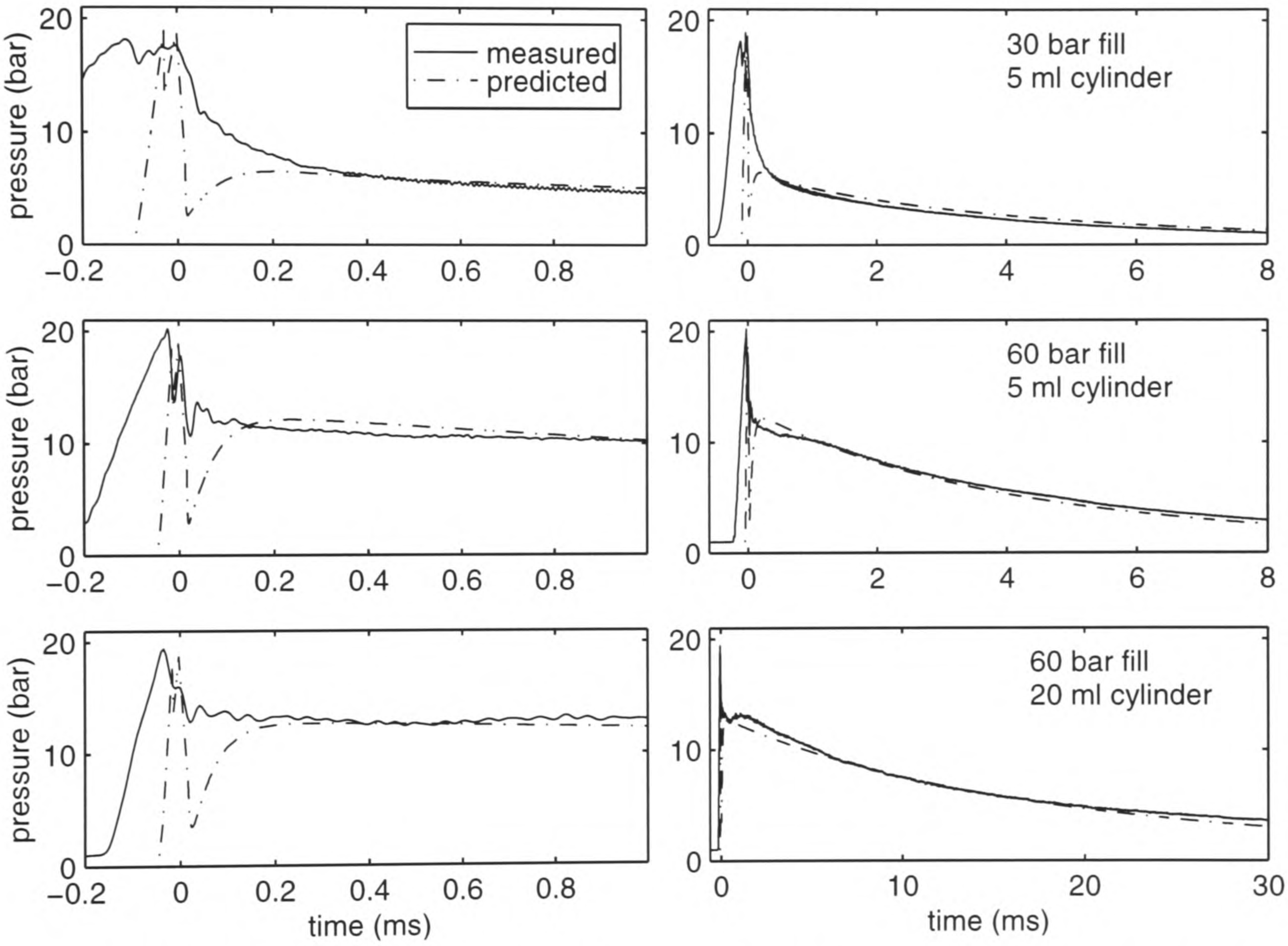


Figure 7.5 Comparisons of predicted and measured rupture chamber pressure in three test cases. Each case is represented on two time scales in side-by-side plots.

On the whole, agreement is excellent at all times more than 0.2 ms after bursting of the downstream membrane. The model severely overpredicts the rate of pressure rise as the rupture chamber fills before bursting, probably because there has been no attempt to assess or model the finite opening time of the valve. However, the filling period has no practical significance. The pressure variations in the first 0.2 ms after membrane bursting are poorly predicted, since the complex processes of wave reflection and interaction have been greatly simplified in the model. However, the scope of this modelling work was restricted at the outset to the quasi-steady regime, and in the period after the first 0.2 ms of flow, the accuracy of the predictions is very satisfactory.

7.1.5 Further Predictions

Detailed predictions of the reservoir blowdown are shown in ***Figures 7.6*** and ***7.7*** for devices with contoured and conical nozzles, respectively. These simulations represented DGV test conditions, with a fill pressure of 60 bar, and 20 μm polycarbonate membranes bursting at 19 bar. In all predictions of rupture chamber pressure (including those shown in ***Figure 7.5***), pressure drops to a minimum in the crudely modelled shock tube stage of flow, just after downstream membrane bursting. It then begins to rise again as inflow from the cylinder takes over, driven by an upstream pressure only slightly lower than the fill pressure. As pressure builds up in the rupture chamber, mass flow through the nozzle rises. At some stage, it exceeds the inflow, and rupture chamber pressure begins to drop.

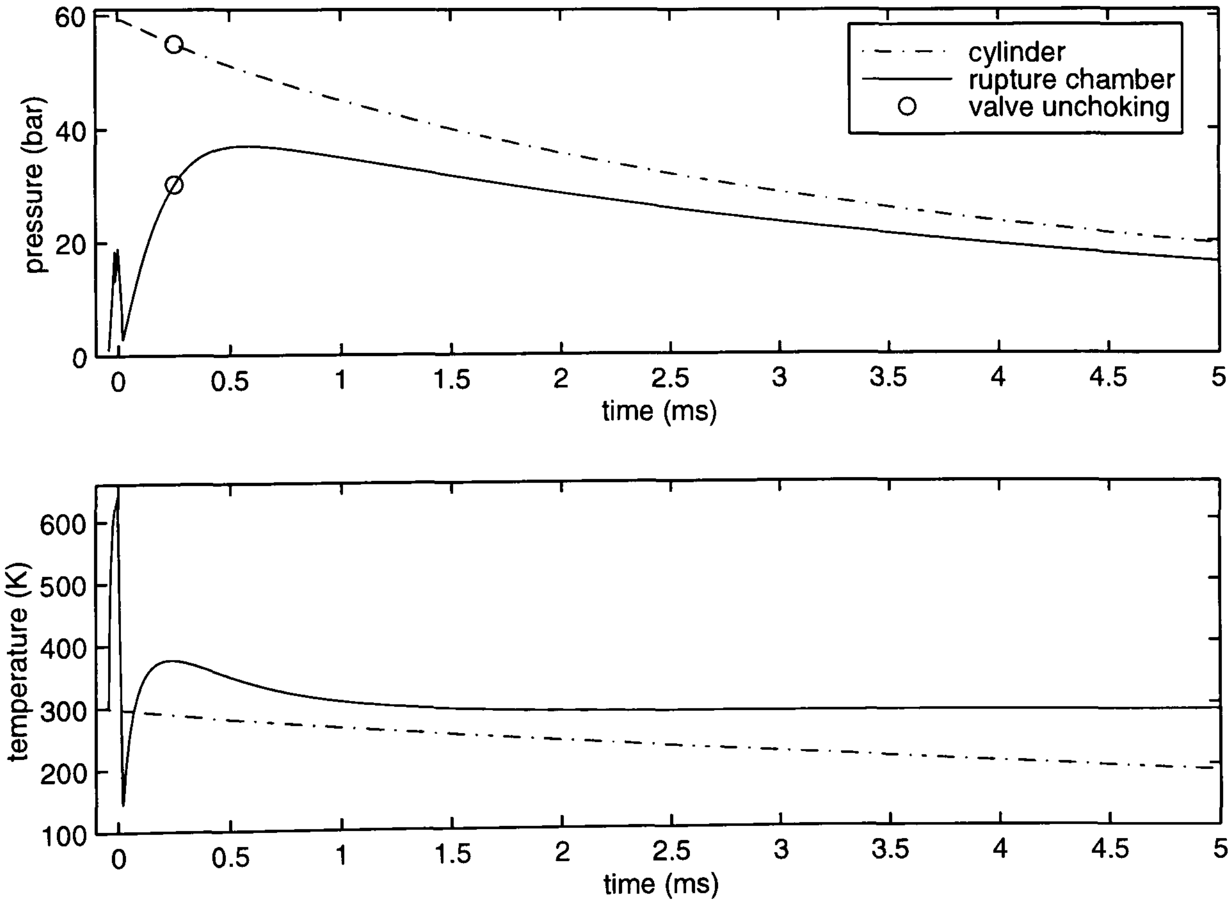


Figure 7.6 Predicted reservoir temperatures and pressures in a device with a 5 ml cylinder, 0.2 ml rupture chamber, 60 bar fill pressure, 20 μm polycarbonate membranes and a contoured nozzle.

In both configurations, temperature and pressure in the cylinder decay monotonically from their initial values, in accordance with the isentropic expansion there. Predicted temperature drops to below 200 K within 5 ms. Temperature in the rupture chamber swings from ambient, 298 K, to 641 K when compressed by inflow before the first membrane rupture. It then falls to 143 K in the expansion which follows bursting of the second membrane. In reality, it may well be that such large temperature swings are moderated to some extent by heat transfer from the surroundings. The large throat of the conical nozzle allows relatively large mass flow from the rupture chamber, keeping rupture chamber pressure much lower than in the contoured nozzle device.

With the contoured nozzle (***Figure 7.6***), which has a smaller throat area than the conical nozzle by a factor of 4, rupture chamber pressure initially rises very rapidly, to the point where the pressure ratio across the valve becomes insufficient to maintain choked flow through the valve. Predictions for the contoured nozzle beyond this point are questionable, since the discharge coefficient of 0.645 was supported by a consideration of choked valve flow only.

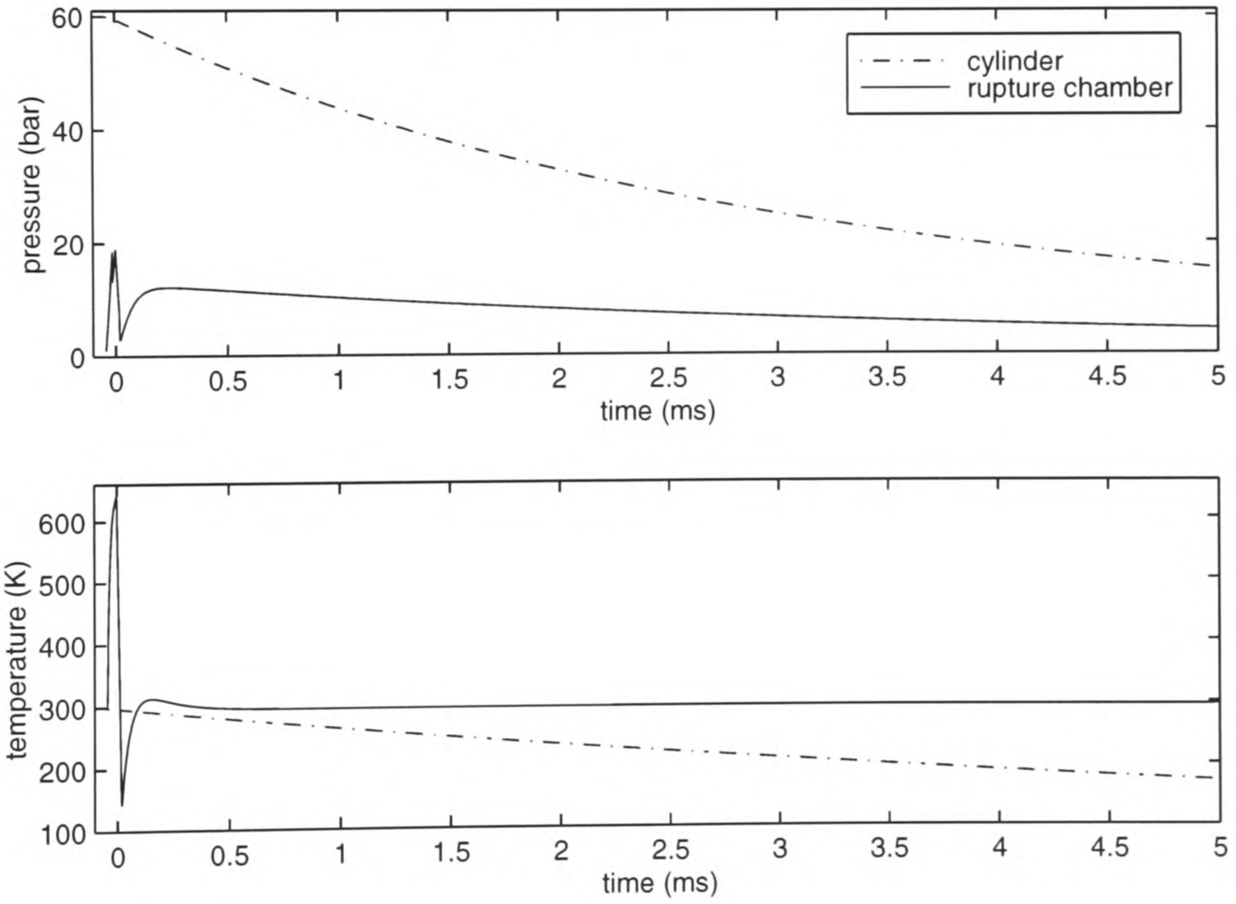


Figure 7.7 Predicted reservoir temperatures and pressures in a device with a 5 ml cylinder, 0.2 ml rupture chamber, 60 bar fill pressure, 20 μm polycarbonate membranes and a conical nozzle.

With the large throat of the conical nozzle draining the rupture chamber, as in the result shown in ***Figure 7.7***, cylinder pressure remains more than 3.5 times as high as rupture chamber pressure over the first 5 ms of flow. Mechanisms of pressure drop other than friction and heat transfer in the valve must be at work. It is also interesting to note that the pressure drop across the valve is much greater with the conical nozzle than with the contoured (small throat) nozzle. These observations prompt a reconsideration of the importance of the valve and the rupture chamber in quasi-steady flow.

Rupture chamber pressure depends on the mass of gas in the chamber, which is determined by the history of mass flow in and out of the rupture chamber. While the valve is choked, flow through the valve depends on upstream (cylinder) pressure and temperature only, and conditions in the rupture chamber have no influence on the flow in the valve. In general, therefore, the static pressure of flow exiting the valve and the pressure in the rupture chamber are mismatched. In sonic valve flow, the static pressure of flow in the valve exit can never be less than the rupture chamber pressure, since unchoking would result. Therefore, flow mechanisms must exist which resolve the mismatch by driving the static pressure of the incoming gas *down* to the pressure of the rupture chamber.

Not only that, but in quasi-steady mode, the bulk flow through the rupture chamber (as an upstream extension of the nozzle) must satisfy the area-Mach number relation, ***Eq. 2.6***. This means that the sonic inflow from the valve must be decelerated to a relatively low Mach number. Since there is a need to reduce the static pressure of the inflowing gas, this reduction of Mach number must be accompanied by a reduction in total pressure. Essentially, this total pressure loss is demanded because no information can propagate upstream through the choked valve to enable inflowing gas to adjust to rupture chamber conditions before arriving in the rupture chamber. It is distinct from any losses which occur within the valve duct itself.

On the fundamental level, it is possible only to speculate about the nature of flow in the rupture chamber and the mechanisms which effect the necessary total pressure loss. The inflow from the valve is an annular jet, with static pressure higher than the “ambient” pressure in the rupture chamber, and a Mach number of 1 (in some average sense). Since the valve flow is known to be influenced by friction, the inflow is certainly non-uniform, with regions of subsonic and supersonic flow. The resulting flow is therefore likely to be a marginally supersonic, underexpanded annular jet, containing expansion waves, shocks, and viscous features. The required decrease in total pressure loss can therefore be brought about through shocks and viscous effects.

To recap, large pressure drops between the cylinder and rupture chamber have been observed in computations which agree well with experiment in their predictions of rupture chamber pressure. Because the valve flow is choked, it is independent of the rupture chamber conditions. Gas arriving in the rupture chamber must somehow be matched to local conditions on entry, and this matching entails a net expansion and deceleration. This combination demands processes which reduce total pressure (or equivalently, raise entropy). Ultimately, these processes have a negative impact on the effectiveness of the device, since total pressure is closely related to the nozzle flow's capability to accelerate particles. Total pressure determines both the density in the gas and the shock structure (if any) in the nozzle. The key to solving this problem lies in manipulation of the mass flows in and out of the rupture chamber to maintain high pressure there.

7.2 Gas and Particle Dynamics in the Nozzle

The goal of the work described in this section is a capability to make approximate predictions of the velocity of gas and particles through the nozzle and at the nozzle exit, given the upstream conditions. This is a complex problem, which has been restricted by assuming quasi-steady flow in the nozzle. The same assumption was described in §7.1 for flow in the reservoir system; it implies that temporal variations in the flow are slow enough that the flow at any instant is identical to a steady flow. This simplification places the early stages of particle delivery beyond the scope of the analysis. The mathematical models of gas and particle flow are described in §7.2.1 and §7.2.2 respectively, and their numerical implementation is outlined in §7.2.3. The predictions of the model are compared with measurements in §7.2.4, and some further predictions are presented and discussed in §7.2.5.

It is assumed that the particles have no significant effect on the gas flow-field, and do not interact with each other. The flow of each phase can then be analysed separately. The gas flow is studied as if no particles were present, and the flow of particles is represented by the motion of a single particle through a known gas flow-field.

7.2.1 Gas Dynamics Model

Flow of gas through the nozzle is modelled using isentropic quasi-one-dimensional flow theory, along with normal shock theory when required. Attention is restricted to conditions at which the flow in the nozzle throat is sonic. As pointed out in §2.3.1, this condition requires a total pressure of just over 1.00 atmospheres for the conical nozzle, or 1.06 atmospheres for the Mach 2.5 nozzle, if isentropic flow is assumed. In reality, when the nozzle flow is marginally choked, it is possible that subsonic flow downstream of the throat separates under the influence of an adverse pressure gradient. In the extreme case, separation might occur at the throat without downstream re-attachment, in which case the flow downstream of the throat would take the form of a sonic jet at atmospheric pressure. The total pressure required to drive this flow can be deduced from ***Eq. 2.4*** as $\left(\frac{\gamma+1}{2}\right)^{\frac{\gamma}{\gamma-1}}$ times atmospheric pressure, or (for helium, $\gamma = 5/3$) 2.05 atmospheres. All the available evidence indicates that at operating conditions of interest, rupture chamber pressure remains higher than this level for a period longer than the time taken for complete payload delivery. For example, with a conical nozzle and a 5 ml cylinder filled to 60 bar, rupture chamber pressure drops to approximately 3 bar after 8 ms of flow, according to both the measured and computed results presented in ***Figure 7.5***. The longest time taken for 90% payload delivery was observed as less than 4 ms, as reported in §4.3.2. It is therefore valid to neglect late stages of flow in which the throat Mach number is less than 1.

Therefore, given that the throat flow is choked for all conditions of practical interest, the critical area A^* is equal to the physical throat area A_t. The Mach number at any location in the nozzle is then a function of the local cross-sectional area, and can be found by an iterative solution of ***Eq. 2.6***. Once the Mach number at a location has been found in this way, and the total pressure and temperature have been specified (typically, by the result of modelling the reservoir system), it is straightforward to calculate the local static conditions, using ***Eqs. 2.4*** and ***2.5***. At any point in time, then, the state of the flow in the nozzle is fully specified by the nozzle geometry and the total conditions.

In the conical nozzle, however, as hypothesised in Chapter 2 and confirmed experimentally in Chapter 4, flow departs from the isentropic idealisation. In the typical operating regime, the ratio of total pressure to exit static pressure is too low to drive supersonic flow from the throat to the exit, and so a shock system forms somewhere in the divergent part of the nozzle. Some computation of the position and strength of this shock or system of shocks is essential. The most likely configuration is a system of oblique shocks, as explained in Chapter 2. However, the only shock structure which lends itself to simple analytical modelling is a single normal shock. It is therefore adopted in the current model as an approximation to the physical situation.

The position and strength of the shock are fixed by the requirement that it must effect a jump from the upstream supersonic flow to a downstream subsonic flow, which must exit the nozzle at atmospheric pressure. The jump must satisfy the normal shock relations, given in many textbooks. In particular, there is a change in total pressure, given by ***Eq. 7.17***, and a change in Mach number, given by ***Eq. 7.18***, where the subscripts 1 and 2 denote conditions immediately upstream and downstream of the shock, respectively.

$$\frac{p_{0,1}}{p_{0,2}} = \left[1 + \frac{2\gamma}{\gamma + 1}(M_1^2 - 1)\right]^{\frac{-1}{\gamma - 1}} \left[\frac{(\gamma + 1)M_1^2}{2 + (\gamma - 1)M_1^2}\right]^{\frac{\gamma}{\gamma - 1}} \tag{7.17}$$

$$M_2^2 = \frac{1 + \frac{\gamma - 1}{2}M_1^2}{\gamma M_1^2 - \frac{\gamma - 1}{2}} \tag{7.18}$$

To determine the location of a normal shock, these equations must be solved simultaneously, along with ***Eqs. 2.4*** and ***2.5*** for isentropic flow and the area-Mach number relation ***Eq. 2.6***, which govern flow within the isentropic regions to either side of the shock. Liou [62] gives a particularly convenient solution procedure, which yields the total pressure ratio across the shock, and has been adopted in the current work. Time-resolved DGV images shown in Chapter 6.4.2 suggest that flow separation is associated with the shock wave or system, leading to a uniform parallel jet downstream of the shock. To account (approximately) for this fact, it is assumed that flow downstream of the shock is entirely uniform.

7.2.2 Particle Dynamics Model

The model described above in §7.2.1 yields axial profiles of gas flow variables for specified total pressure and temperature at the nozzle inlet. In order to calculate the motion of a particle through this field, it is assumed that the flow-field varies negligibly in the course of a particle's transit through the nozzle. It is assumed that particles do not interact with each other or perturb the gas flow, so that the flow of particles through the device at any time in the history of the flow-field can be analysed by considering an isolated test particle at that time. The problem of the gas-particle interaction then reduces to the problem of determining the force on a particle as a function of particle properties and local flow variables.

Drag is defined as the force in the direction of flow on a body immersed in a flow, and is conventionally expressed non-dimensionally as the drag coefficient C_D, as in ***Eq. 7.19***.

$$C_D = \frac{D}{\frac{1}{2}\rho(u - u_p)^2 A} \tag{7.19}$$

D is the drag force, ρ is the local fluid density, u is the free-stream velocity of the fluid, u_p is the velocity of the body, and A is the frontal area presented to the flow by the body. (One-dimensional flow is assumed throughout this analysis, and velocity is treated as a scalar). The relative velocity $u - u_p$ is called the *slip velocity* in the context of two-phase flow. Extensive data for the drag coefficient of particles are available in the literature. In the current work, attention is restricted to spherical particles. A complete treatment of non-spherical particles would be complex, and might not give significant improvements in overall accuracy.

Henderson [63] drew on various sources of experimental data (notably those of Bailey and Hiatt [64], which cover the Mach number regimes of interest in this work) to produce a set of three equations which constitute a correlation for sphere drag coefficient as a function of Reynolds number (based on particle diameter and slip velocity) and Mach number (based on slip velocity). Henderson's correlation is claimed to be the most accurate algebraic description of sphere drag over wide-ranging conditions, with a maximum deviation of 16% from the empirical data on which it is based.

An alternative class of correlations has been developed by researchers who measured the acceleration of particles from rest in shock tubes, with the aim of characterising the effects of unsteady flow on particle drag. Correlations for drag coefficient of spheres in shock tubes, as a function of Reynolds number, have been published by Rudinger [26, pp. 119-125], Kurian and Das [65], Igra and Takayama [66] and Rogue *et al.* [29], among others.

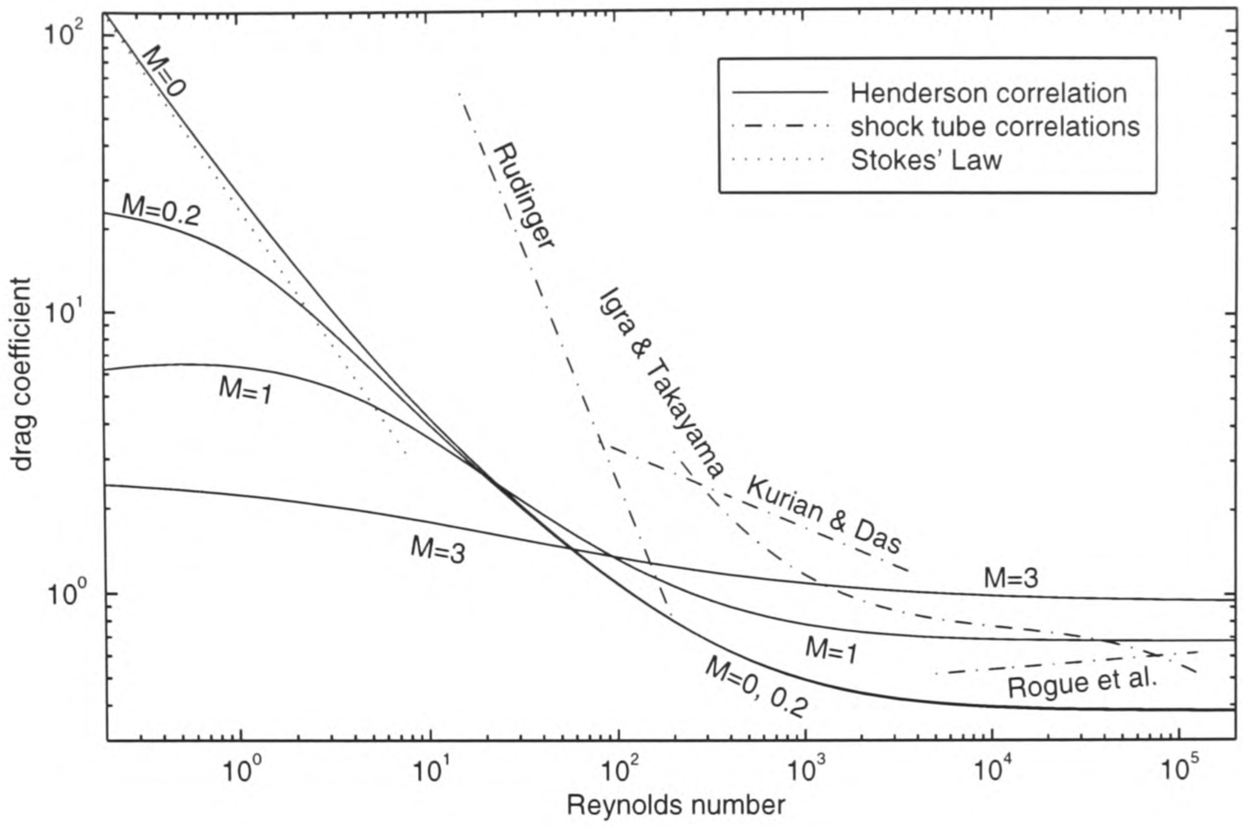


Figure 7.8 Drag coefficient of a sphere as a function of Reynolds number, according to Henderson's correlation for steady flow at selected Mach numbers, various correlations based on shock tube measurements, and Stokes' law for low Reynolds number.

These shock tube correlations, and samples of Henderson's steady drag correlations, are plotted in ***Figure 7.8***, along with Stokes' Law, which is valid at Reynolds numbers less than 1. The curves based on shock tube data differ significantly from Henderson's curves for steady flow, bearing out the assertion that unsteady effects are important. However, they are unattractive as practical tools for modelling. They are not consistent with each other, and they do not incorporate the influence of Mach number, or of the specific character of the unsteadiness. Their scope appears to be limited to flows very similar to those on which they are based. For these reasons, the correlations based on shock tube flow are rejected in favour of Henderson's correlation in the current application.

7.2.3 Numerical Solution Method

The numerical implementation of the particle dynamics model is straightforward. Before the main computation begins, the total pressure ratio across the normal shock (if any) required to match nozzle operating conditions is calculated using Liou's method. Beginning at the nozzle inlet plane, the solution algorithm marches forward in time and downstream in x, computing the drag on the test particle, and using this to update the particle's acceleration, velocity and position. As in the model of the reservoir flow system, time derivatives are approximated by first-order backward differences. At the beginning of the nth time step, the particle's position x_n and velocity $u_{p,n}$ are provided by calculations made on the previous step. The time-marching algorithm followed is outlined below:

- evaluate nozzle area at $x = x_n$, using nozzle geometry specification
- evaluate Mach number at $x = x_n$, using area-Mach number relation
- if shock is not already detected at $x < x_n$, check if a shock at $x = x_n$ provides the required total pressure drop — if so, calculate properties downstream of the shock for subsequent (downstream) calculations
- calculate gas density ρ, temperature T and velocity u at $x = x_n$ using isentropic flow relations (if not downstream of a shock) *or* use shocked gas properties, as appropriate
- evaluate viscosity μ as function of T using Sutherland's Law
- calculate slip Reynolds number $Re_d = \frac{\rho|u - u_p|d_p}{\mu}$ and slip Mach number $M = \frac{|u - u_p|}{\sqrt{\gamma RT}}$
- use Henderson's correlation to give $C_D(M, Re_d)$
- compute new particle acceleration $\frac{du_p}{dt} = \frac{3}{4d_p}\frac{\rho}{\rho_p}(u - u_p)|u - u_p|C_D$
- compute new particle velocity $u_{p,n+1} = u_{p,n} + \frac{du_p}{dt}\Delta t$
- compute new particle position $x_{p,n+1} = x_{p,n} + u_{p,n+1}\Delta t$
- increment n and advance to next time step.

7.2.4 Comparison of Computed and Experimental Results

The time-resolved DGV data for the conical nozzle, including measurements inside the nozzle, provide the most complete basis for a test of particle velocity computations against measurement. Predictions of the velocity of a 4.7 μm sphere in a conical nozzle at times 0.2 ms, 0.7 ms and 1.2 ms after the start of nozzle flow are shown in ***Figure 7.9***, along with smoothed centreline velocity profiles taken from measurements made at approximately corresponding times. The strong dips in the measured velocity, at approximately 7 – 8 mm upstream and downstream of the exit, are artifacts associated with images of the nozzle rim.

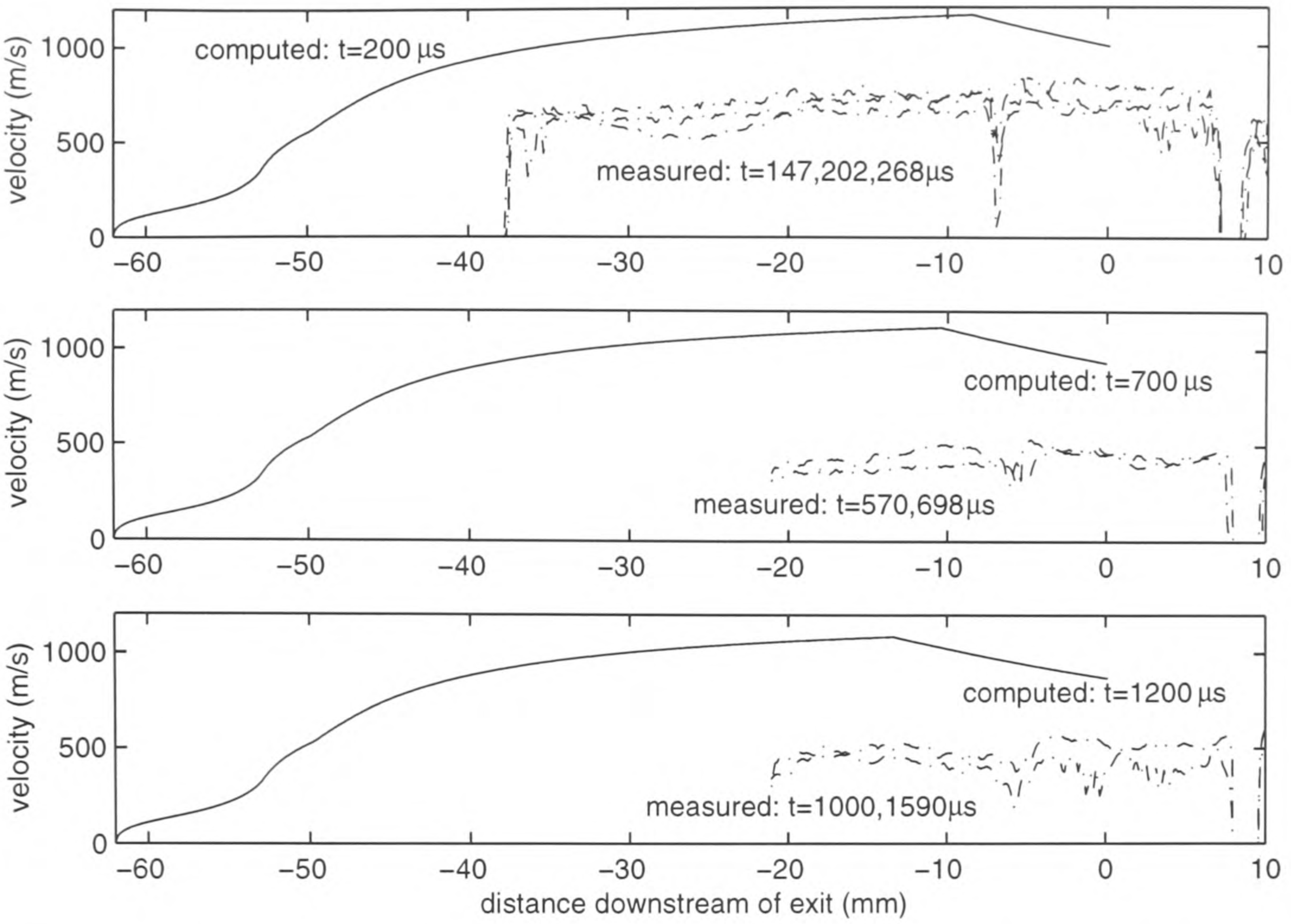


Figure 7.9 Comparisons of measured and computed particle velocity at various times in the conical nozzle. The particle is a 4.7 μm polystyrene sphere, and the fill pressure is 60 bar.

On the whole, computations overestimate the measured velocity by a large margin. The main reason for this is probably the inadequacy of theoretical modelling of the shock structure and separation in the nozzle. While the simple shock model used in these computations places a solitary normal shock less than 10 mm upstream of the nozzle exit, it is clear from the time-resolved velocity field measurements shown in §6.4 (particularly ***Figure 6.39***) that flow separates well upstream of that point. Other physical mechanisms which are not represented in the model, and which probably contribute to the discrepancies, include substantial momentum transfer from the gas (to the particles) with consequent loss of velocity in both phases, and boundary layer effects upstream of the shock structure.

In ***Figure 7.10***, the computed time history of the exit velocity of 4.7 μm microspheres from the conical nozzle is plotted with the measured time-integrated exit velocity values (averaged over all values of payload mass) and a time history of measured exit velocity. Again, the computations overpredict the measured velocity. In the early stages of the flow, the computed 4.7 μm particle exit velocities are approximately 50% higher than measurements. In the period from 0.5 ms to 3 ms after the beginning of flow, this margin rises to 100%. The computed exit velocities show some sensitivity to particle size, with the 4.7 μm particles exiting the nozzle 200 to 300 m/s faster than the 26.1 μm particles. The average measured time-integrated exit velocities, on the other hand, are effectively independent of particle size (the complete exit velocity data, shown in ***Figure 6.26***, indicate that their variation over a range of approximately 80 m/s is not correlated to particle size).

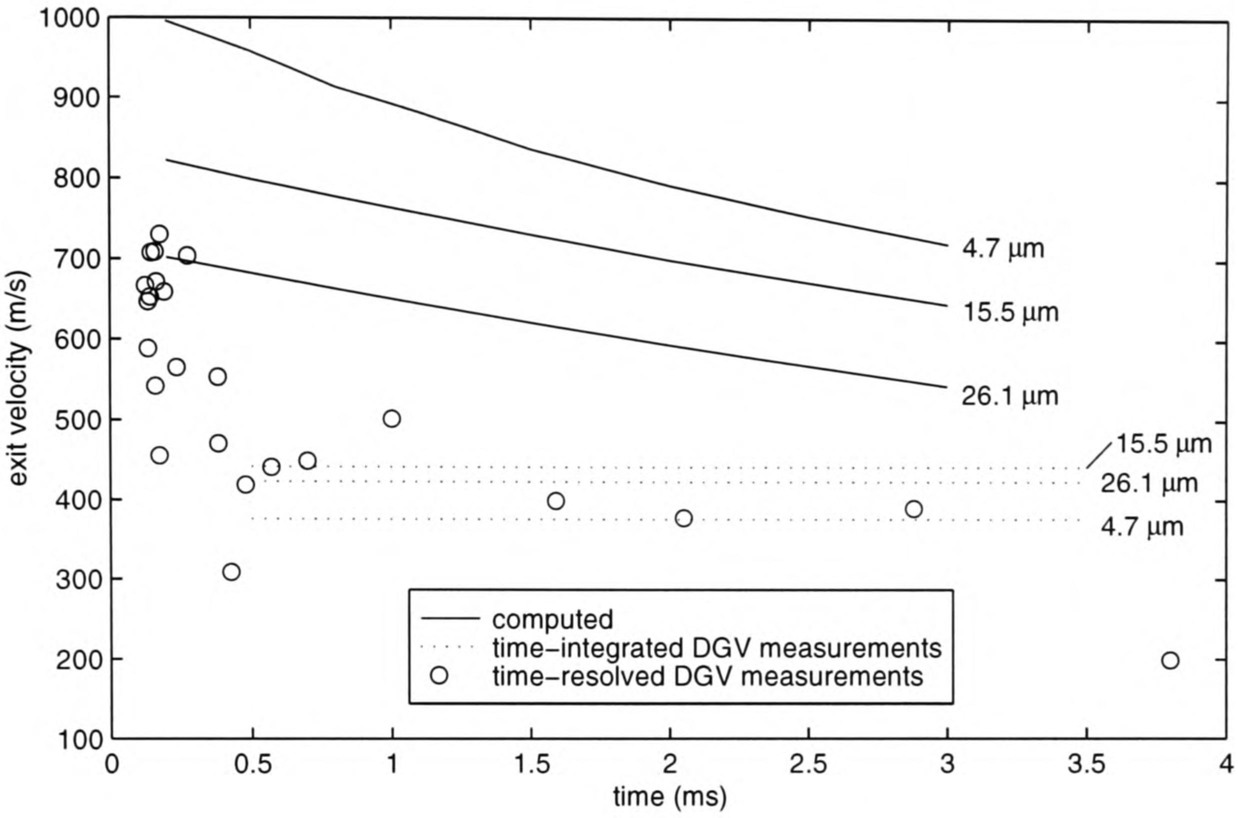


Figure 7.10 Variation of particle exit velocity with time, computed for polystyrene microspheres of three sizes and measured using time-resolved DGV for 4.7 μm microspheres. Time-integrated measurements for each type of microsphere are also shown.

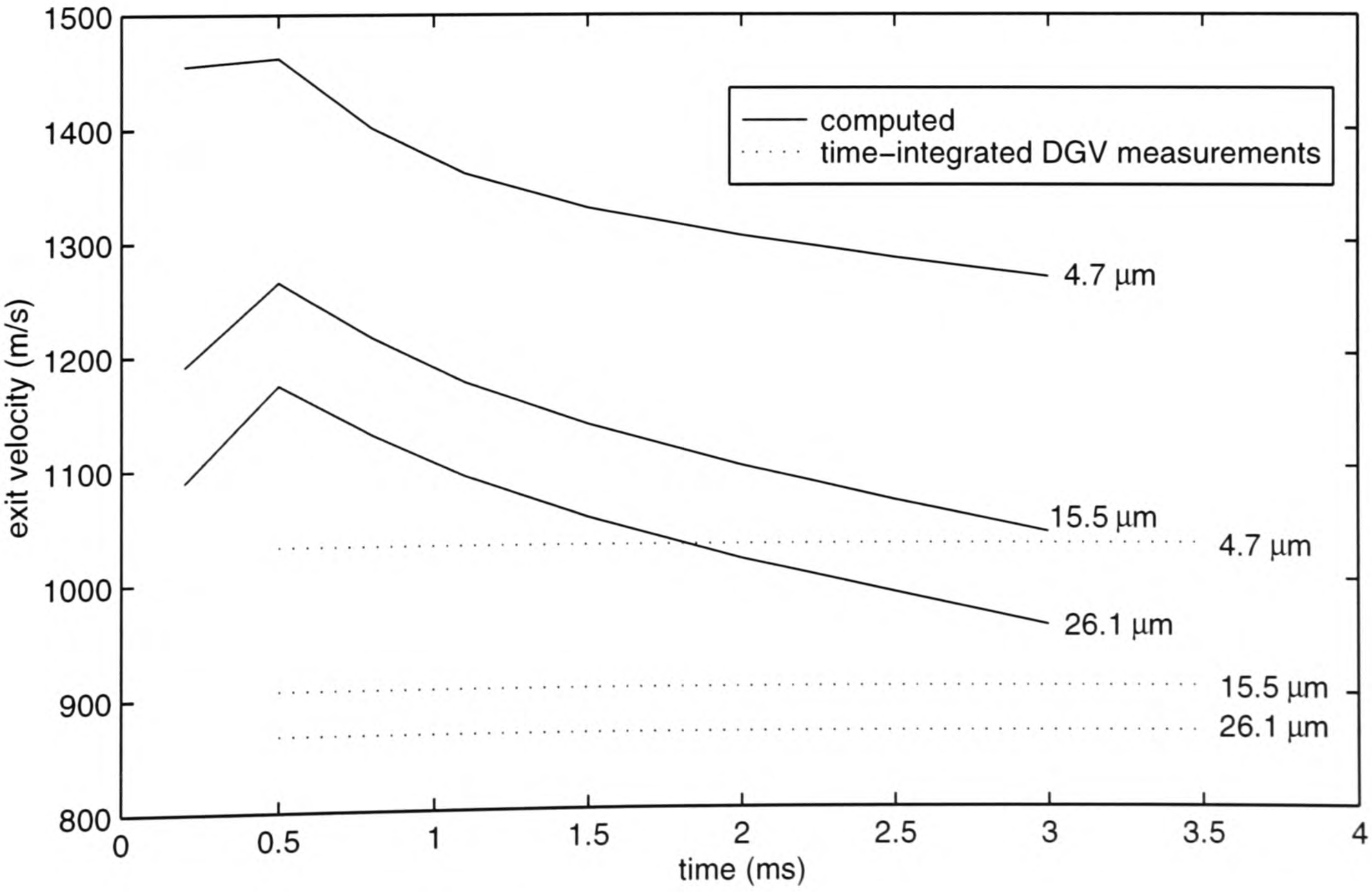


Figure 7.11 Comparisons of computed and measured exit velocity for polystyrene microspheres in the Mach 2.5 contoured nozzle.

A similar set of comparisons of measured and computed exit velocity for the Mach 2.5 contoured nozzle are shown in ***Figure 7.11***. Only the data for the smallest payloads (i.e. 1 mg) have been incorporated in the time-integrated data in this figure, since a significant effect of payload mass was observed in the measurements, and the computations assume isolated particles. The time-varying computations of particle exit velocity exceed the time-integrated measurements by a margin that ranges from 25% to 45% for the 4.7 μm particles, and 12% to 40% for the 26.1 μm particles, over the first 3 ms of flow. The improved agreement between prediction and measurement for this nozzle geometry, in comparison with the conical nozzle, confirms that the shock and separation structures in the conical nozzle are a major cause of disagreement between experiment and computation for that geometry.

The complete experimental data for exit velocity, shown in ***Figure 6.13***, indicate that particle exit velocity rises significantly with decreasing particle size in the Mach 2.5 contoured nozzle. This is reflected in a difference of 160 m/s between the averaged measurements for 1 mg payloads of 4.7 μm and 26.1 μm particles, as shown in ***Figure 7.11***. This sensitivity is stronger in the computed results, with exit velocities spread over a range of 320 m/s.

7.2.5 Additional Predictions

The most interesting applications of the particle velocity prediction code are in exploring the mechanisms which affect particle delivery velocity on a level that is not accessible to experiments. In this section, several sets of results in this category are presented. First, predictions at standard test conditions are presented in a complete form, with detail of several relevant flow variables. Predictions of the variation of the axial velocity profiles, with time and with particle size, are then shown.

A detailed prediction of the gas flow and particle trajectory for a 26.1 μm polystyrene microsphere in the Mach 2.5 nozzle is presented in ***Figures 7.12*** and ***7.13***, as a set of axial profiles of variables which play a part in particle dynamics. These variables are gas, particle and slip velocity, particle acceleration, and gas density (all shown in ***Figure 7.12***), and Reynolds number, Mach number and drag coefficient (***Figure 7.13***).

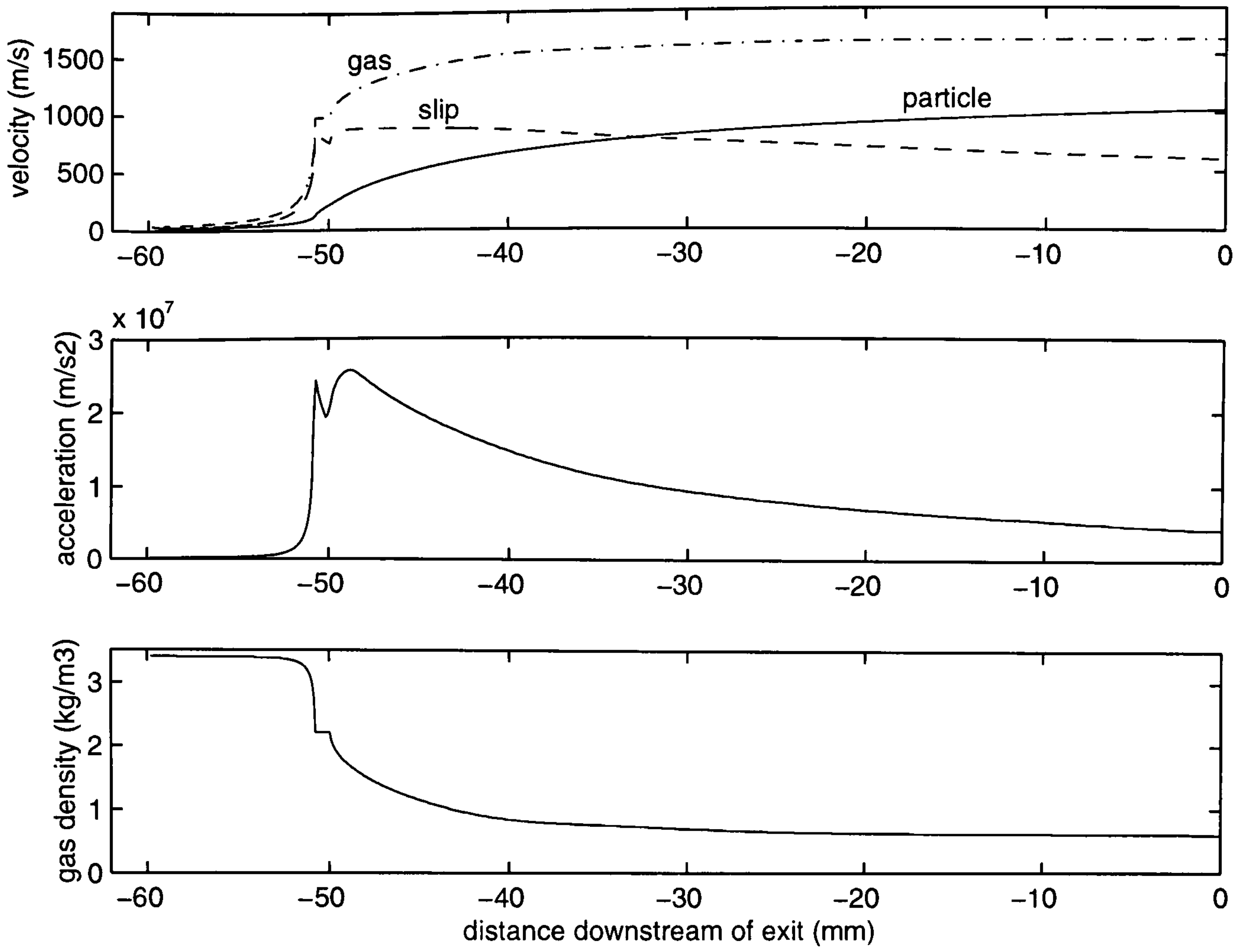


Figure 7.12 Axial variation of gas, particle and slip velocity, particle acceleration and gas density for a 26.1 μm microsphere in the Mach 2.5 contoured nozzle at total pressure 26.4 bar, total temperature 375 K.

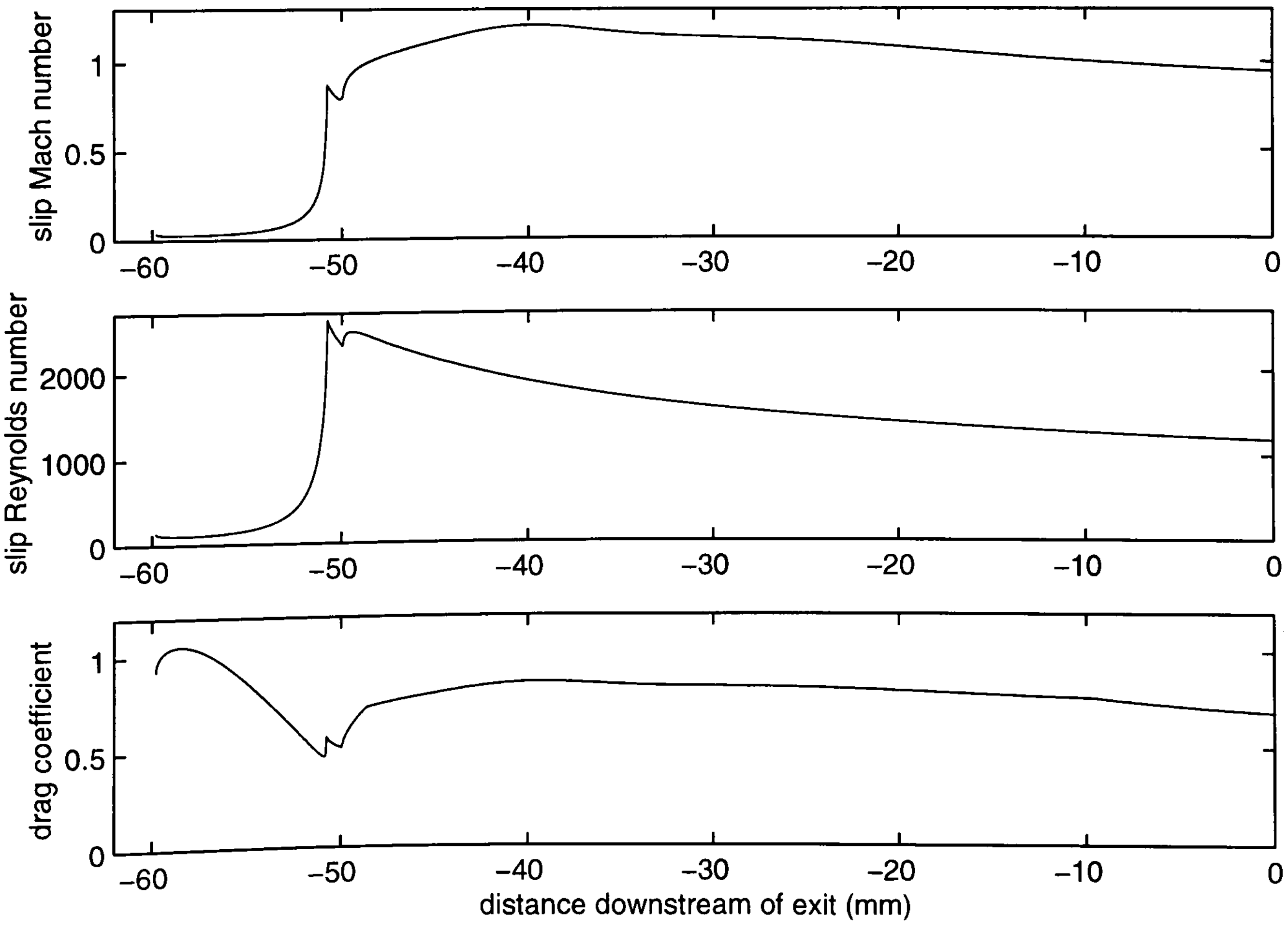


Figure 7.13 Axial variation of slip Mach number, slip Reynolds number and drag coefficient for a 26.1 μm microsphere in the Mach 2.5 contoured nozzle at total pressure 26.4 bar, total temperature 375 K.

Both the particle and the gas accelerate very little through most of the convergent section. As the gas accelerates rapidly in the approach to the throat, a large slip velocity develops and significant particle acceleration begins. The particle achieves 85% of its exit velocity when it has travelled 50% of the distance from throat to exit. Downstream of this point, particle velocity increases less and less rapidly with distance. This is a consequence of decreasing drag, as well as the simple fact that the particle's residence time in this part of the nozzle is short, due to its high velocity. The major causes of decreasing drag in the divergent part of the nozzle are the decay of density with increasing Mach number, and the falling slip velocity, which is due to the near-uniformity of gas velocity in the downstream part of the nozzle.

A similar description of the trajectory of a 26.1 μm particle through the conical nozzle is presented in ***Figures 7.14*** and ***7.15***. Total temperature (310 K) and pressure (12.1 bar) correspond to the state predicted by the reservoir gas dynamics code 0.2 ms after membrane rupture. At these conditions, the computed flow-field contains a normal shock wave approximately 8 mm upstream of the exit. The particle accelerates smoothly and rapidly through the convergent section and the throat, lagging the gas. Again, acceleration drops off rapidly with distance downstream of the throat, and the particle reaches a near uniform velocity as it approaches the shock wave, even though the slip velocity remains constant. As the particle passes through the shock, the particle velocity becomes higher than that of the surrounding gas and a very slight deceleration results.

The effects which dominate particle acceleration are essentially the same for the two nozzles, though they are more pronounced in the conical nozzle. It is clear that drag coefficient is not a major influence, since in both nozzles C_D is a minimum at the throat, quite close to the location of maximum absolute drag. The main reason for the poor velocity gain in the divergent part of each nozzle is low gas density. It appears that the conical nozzle, in particular, makes inefficient use of the available gas energy because of the high Mach number and consequent low density which it induces.

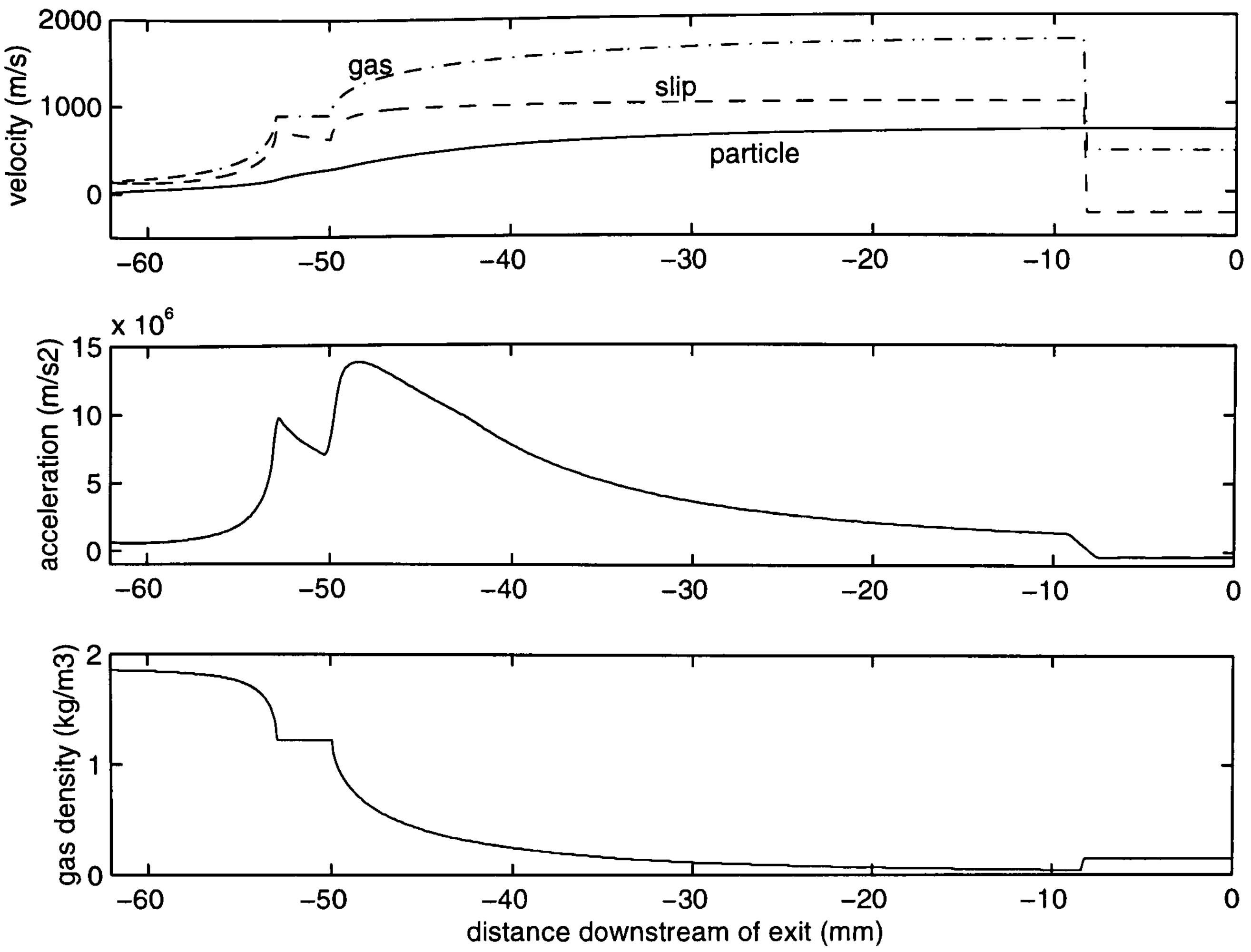


Figure 7.14 Plots of gas and particle velocity, drag force and gas density for a 26.1 μm microsphere in the conical nozzle at total pressure 12.1 bar, total temperature 310 K.

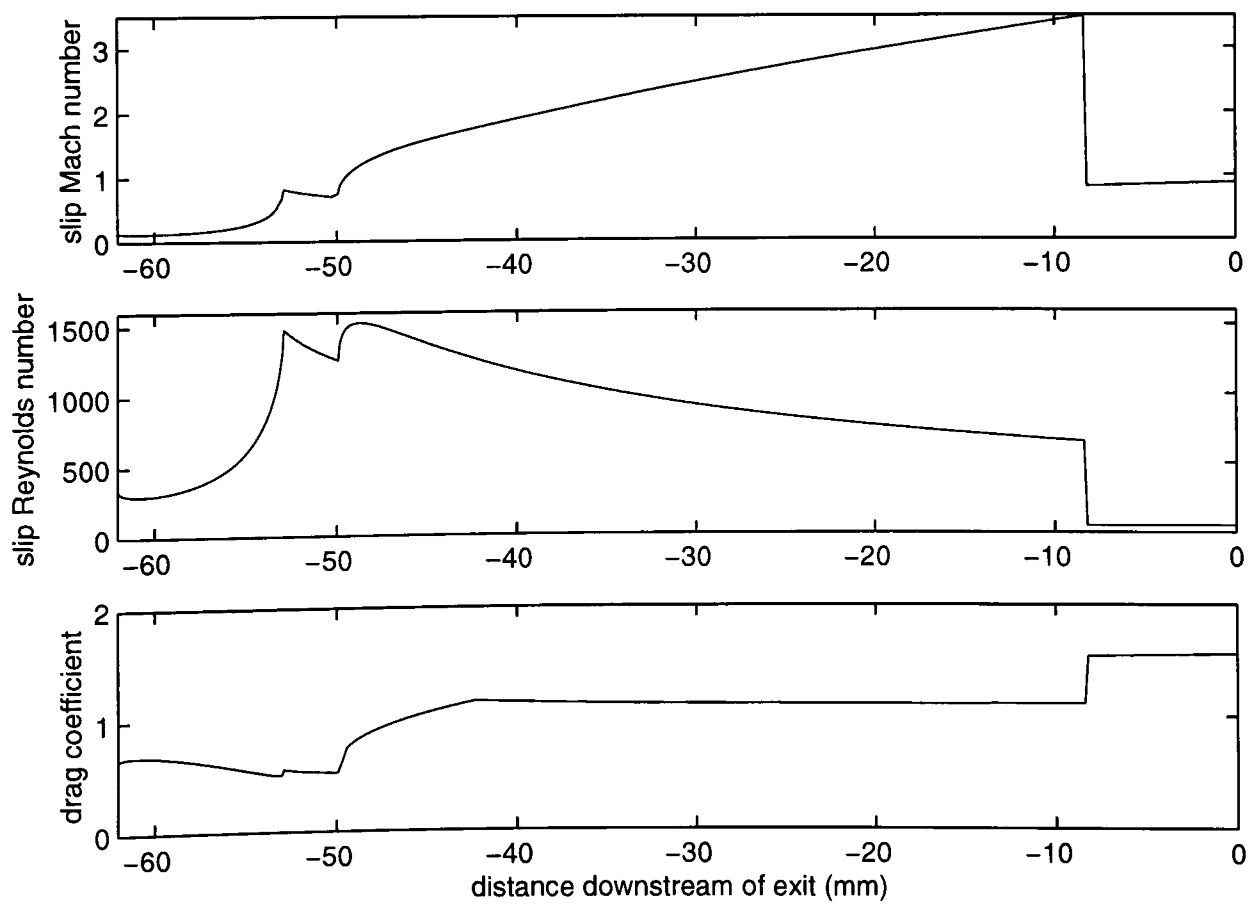


Figure 7.15 Plots of slip Mach number, slip Reynolds number and drag coefficient for a 26.1 μm microsphere in the conical nozzle at total pressure 12.1 bar, total temperature 310 K.

The following simple consideration of isentropic flow illustrates the roles of density and gas velocity. For fixed total conditions, gas velocity tends towards an upper bound as Mach number increases, because of decreasing static temperature and speed of sound. This is expressed in ***Eq. 7.20***.

$$\lim_{M \to \infty} u = \sqrt{\frac{2}{\gamma - 1}} \sqrt{\gamma R T_0} \tag{7.20}$$

For helium, this limit on gas velocity is just 1.73 times the stagnation speed of sound. Gas density, on the other hand, continues to fall towards zero as Mach number tends to infinity. As Mach number is increased to very high values, therefore, falling density outweighs gains in gas (and slip) velocity, with a detrimental effect on particle acceleration. Specifically, as pointed out by Prof. Brian Bellhouse, the product ρu^2 (which appears in the definition of drag coefficient) has a global maximum at $M = \sqrt{2}$. Under fixed total conditions, this is the optimum Mach number for acceleration of a resting particle (ignoring variations of drag coefficient). The situation is more complex for a moving particle, when slip velocity is not the same as gas velocity.

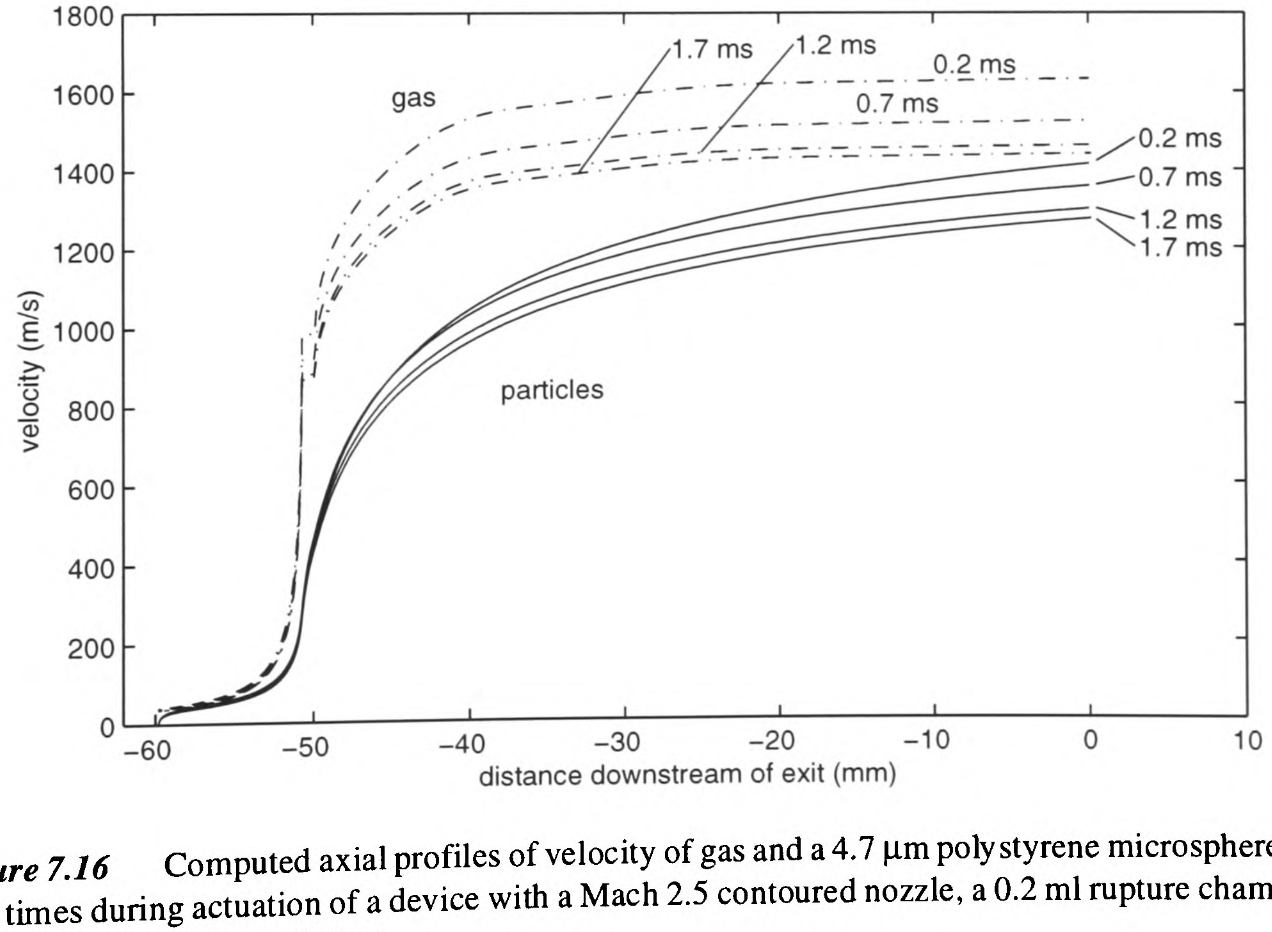


Figure 7.16 Computed axial profiles of velocity of gas and a 4.7 μm polystyrene microsphere at various times during actuation of a device with a Mach 2.5 contoured nozzle, a 0.2 ml rupture chamber and a 5 ml cylinder, filled to 60 bar.

In the predictions presented in the remainder of this section, the variations of the particle velocity profile with time and with particle size are examined for both the Mach 2.5 contoured nozzle and the conical nozzle. In ***Figure 7.16***, axial profiles of the velocity of a 4.7 μm sphere in the Mach 2.5 nozzle are shown for various times in the flow history. The velocity varies in a complex manner with time, because of the complexity of the computed variations of total temperature and pressure. As shown in ***Figure 7.6***, total temperature initially rises, causing gas velocity to increase and density to decrease. Total pressure also rises at the beginning of the flow, competing against the effects of total temperature by tending to raise gas density. Total temperature reaches a maximum first, and begins to fall while total pressure continues to rise. As total temperature stabilises at ambient temperature, total pressure decays monotonically. The decay in particle velocity from 0.2 to 1.2 ms, apparent in ***Figure 7.16***, is due largely to the fall in gas velocity associated with a large drop in total temperature. Between 1.2 ms and 1.7 ms, falling total pressure outweighs falling total temperature, and causes a net decrease in density, with a slight drop in particle exit velocity.

To assess particle size effects, the 4.7 μm and 26.1 μm polystyrene spheres have been modelled, as well as an imaginary 50 μm polystyrene sphere, included simply to extend the range of test conditions. Computed axial velocity profiles for these particles in the contoured nozzle are shown in ***Figure 7.17***. As expected, particle velocity tends towards gas velocity as particle diameter is reduced.

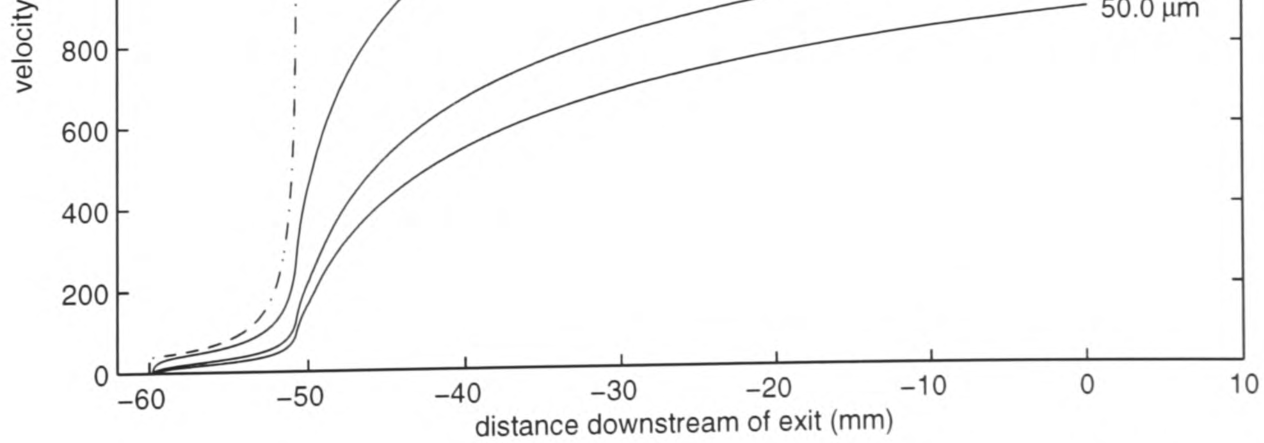


Figure 7.17 Axial profiles of gas velocity and the velocity of polystyrene microspheres, delivered through a Mach 2.5 contoured nozzle 0.2 ms after the beginning of flow. The simulated device has a 0.2 ml rupture chamber and a 5 ml cylinder, filled to 60 bar.

In the conical nozzle, the velocity of a 4.7 μm particle falls off monotonically with time, as shown in ***Figure 7.18***. Between 0.2 and 0.7 ms, gas velocity upstream of the shock drops slightly because of falling total temperature. From then on, gas velocity is practically constant, while the particle's velocity in the region of supersonic gas flow falls off slightly with time. This is because of the decay in density, which is due in turn to the decay in total pressure. The primary cause of the velocity decay at exit is the upstream drift of the normal shock, the point at which particle deceleration begins, again caused by decaying total pressure.

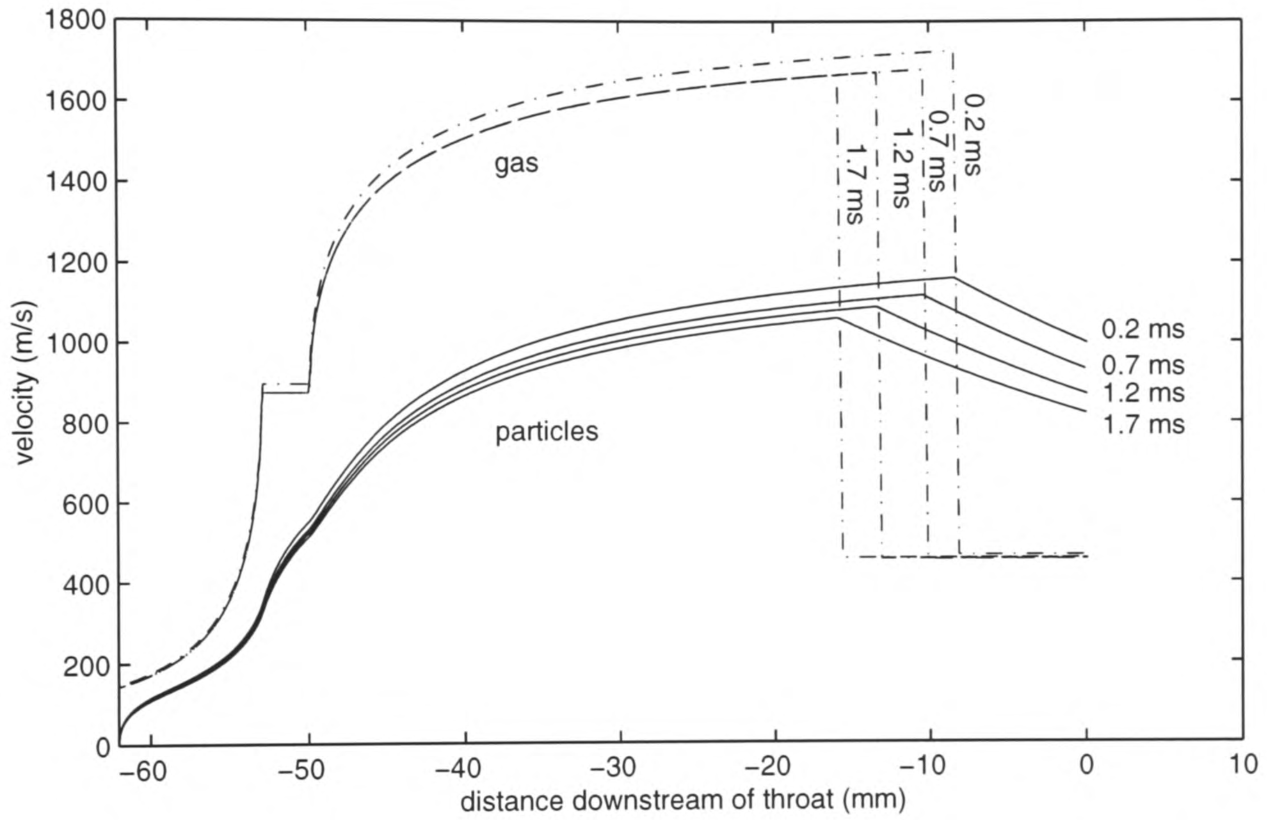


Figure 7.18 Computed axial profiles of velocity of gas and a 4.7 μm polystyrene microsphere at various times during actuation of a device with a conical nozzle, a 0.2 ml rupture chamber and a 5 ml cylinder, filled to 60 bar.

Velocity profiles for particles of three sizes in the conical nozzle are shown in ***Figure 7.19***. As expected, the smallest particles accelerate most rapidly under the influence of positive slip velocity. As gas density decays near the downstream end of the nozzle, the 26.1 μm and 50 μm particles almost entirely cease to accelerate. Downstream of the shock, their slip velocity becomes negative, but quite low. This, in combination with their large inertia, overrides the increase in gas density due to the shock, and there is no appreciable particle deceleration. In contrast, the velocity of the 4.7 μm particle falls by almost 200 m/s over the 8 mm distance from the shock to the exit. Its high drag-to-mass ratio and large negative slip velocity on entry to the subsonic zone, and the high density of the subsonic gas, contribute to this deceleration. The properties of the 4.7 μm particle which give it high acceleration in upstream regions of the nozzle, and its resulting high velocity at the shock front, are precisely the factors which cause its rapid deceleration downstream of the shock.

Even though the time-resolved DGV results suggest that the treatment of shocked flow in the conical nozzle is simplistic, the numerical model can be considered to represent the important physical phenomena of the flow. In this context, the predictions depicted in ***Figure 7.19*** suggest two interesting qualitative results. They illustrate how particles of various sizes can be delivered from the conical nozzle at approximately equal velocities, as observed in time-integrated DGV (§6.3). The shock in the computed solutions considerably reduces the discrepancy which builds up between the 4.7 μm and 26.1 μm particles in the supersonic flow zone. Secondly, these results show that it is possible for particles to move through the downstream part of the nozzle at a higher velocity than local gas. This behaviour was hypothesised in §4.4 to explain certain features of nozzle pressure measurements made in the presence of particles.

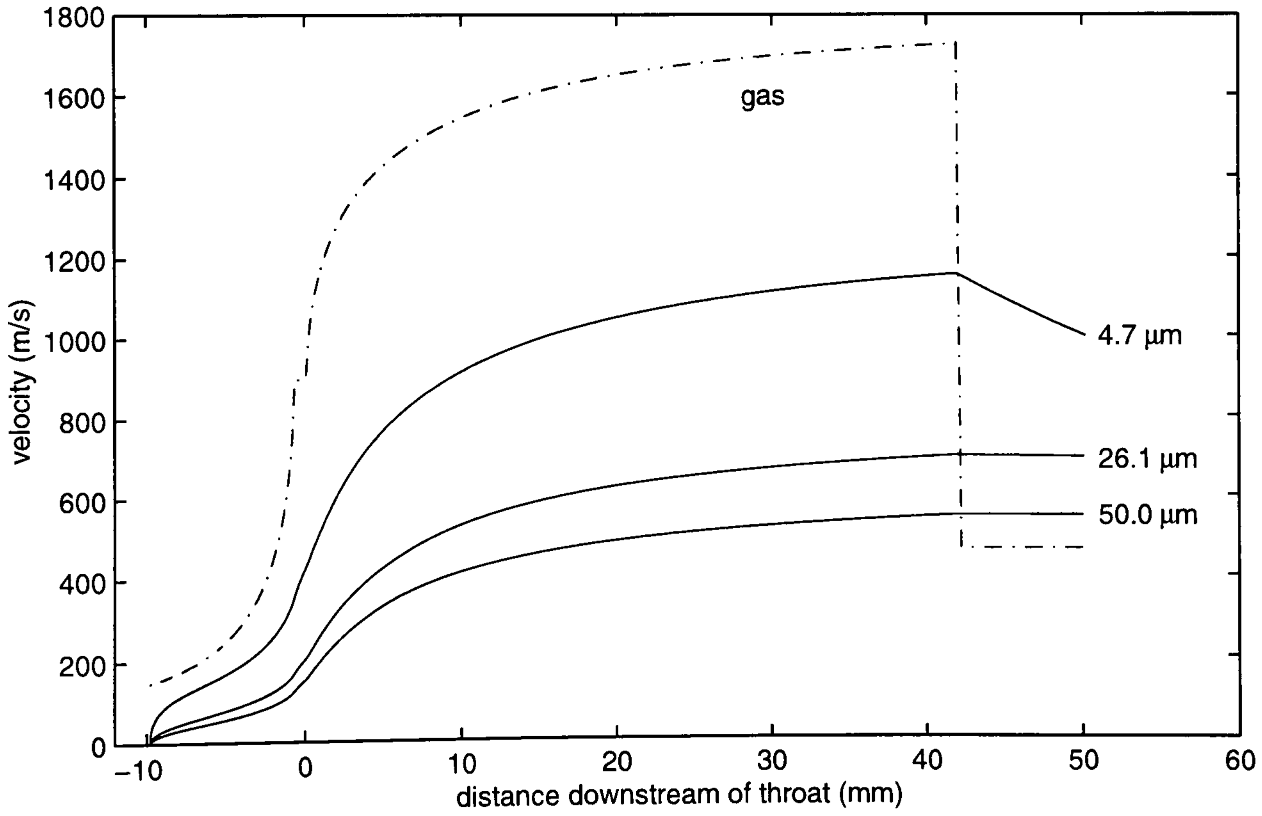


Figure 7.19 Axial profiles of velocity of gas and polystyrene microspheres, delivered through a conical nozzle 0.2 ms after the beginning of flow. The flow-field was computed for a fill pressure of 60 bar.

7.3 Conclusions

Mathematical models of flow in the cylinder and rupture chamber, and of gas and particle flow in the nozzle, have been developed and used as the basis of predictive software codes. The main output of the code for reservoir flow consists of time histories of pressure and temperature in the rupture chamber, which can be used as an input to the nozzle flow model. This model's predictions have been tested directly against measurements of rupture chamber pressure in a device with a conical nozzle. Agreement is excellent for times later than 0.2 ms after the beginning of flow, when quasi-steady flow is established. An empirically tuned discharge coefficient, used to describe the loss of total pressure in the cylinder outlet valve, has been shown to be a realistic approximation to the effects of friction and heat transfer in the valve.

The predictions of the reservoir flow model have drawn attention to a cause of overall device inefficiency, due to the fact that the pressure of the gas in the rupture chamber is effectively determined by mass flow rates through the cylinder outlet valves and the nozzle. Incoming gas from the cylinder is therefore forced to adjust to this pressure level, often with a substantial loss in total pressure.

The nozzle flow model predicts gas and particle velocity, and all other gas properties, in the quasi-steady stages of flow. The results of the particle velocity model tend to overpredict the velocities measured by DGV. In the case of the Mach 2.5 contoured nozzle, this is probably largely due to friction in the nozzle, and the absence of a model of particle loading of the gas. Analysis of flow in the conical nozzle is also compromised by the lack of an accurate quantitative model of shock structure and separation in the nozzle. However, predicted trends in the dependence of exit velocity on particle size show similarities to experimental results.

The predictions have provided some insight into the mechanisms of particle acceleration. They have drawn attention to the role of gas density as a factor in particle acceleration. In particular, as nozzle Mach number increases to very high values, decreases in density outweigh gains in velocity and cause reduced particle drag. It has been demonstrated that a shock system in the conical nozzle can tend to reduce the dependence of particle exit velocity on particle size by negating the rapid acceleration of small particles in the upstream section of the nozzle. This is a possible explanation for the experimental result that particle exit velocity is independent of particle size for the conical nozzle. It has also been shown that a shock in the nozzle can lead to particle velocities which are higher than gas velocities at exit.

Chapter 8 Conclusions

The primary motivation for this research was to characterise and explain the flow of gas and particles in devices for transdermal powdered drug delivery, through an experimental and numerical investigation of prototype devices. The results of this effort are summarised in §8.1. Conclusions regarding Doppler Global Velocimetry, arising from the application of the experimental technique to the flows of interest in this work, are outlined in §8.2. Directions for future research are suggested in the third and final section of this chapter.

8.1 Description and Understanding of Gas and Particle Dynamics

The flow of gas and particles in prototype drug delivery devices has been studied by means of pressure measurement, optical particle detection, particle velocity measurement and numerical simulation. A device fitted with a conical nozzle, closely resembling future production devices, has been the subject of the most comprehensive investigations. Contoured nozzles of two types have also been studied. The flow in the contoured nozzles displayed behaviour which was not observed in the conical nozzle device, providing a more general basis for the discussion of gas and particle dynamics in these drug delivery devices.

The gas flow histories observed in the contoured and conical nozzles are similar in some general features. In both nozzles, flow is initiated by a shock wave which is formed in membrane rupture and propagates down the nozzle. This primary shock opens a starting process of typical duration 200 μs, during which an unsteady wave system brings about a transition from the initial shock tube-like flow to a relatively slowly varying nozzle flow. This quasi-steady flow is driven by the pressure ratio across the nozzle, and is sustained for several milliseconds while the reservoir system drains.

The character of the long-term quasi-steady flow is determined by the nozzle geometry. In the contoured nozzle, flow is supersonic and overexpanded for up to 8 ms, and static pressure in the nozzle falls as reservoir pressure and total pressure decay. This period comes to an end when a shock system drifts upstream from the exit plane. According to time-integrated DGV measurements, the exit velocity of polystyrene microspheres from the contoured nozzle ranges from 740 m/s to 1050 m/s, and increases as either particle size or payload mass is decreased. The highest measured velocities for silica particles approach 1200 m/s. The magnitude of these velocities suggests that a significant fraction of the payload, if not all, is entrained in the quasi-steady supersonic flow. The insensitivity of lidocaine particle velocity to nominal particle size is probably due to disintegration of lidocaine particles inside the device. The effect of payload mass, for microsphere payloads, suggests that the particles have a significant influence on the gas flow-field.

In the conical nozzle, on the other hand, supersonic quasi-steady exit flow does not occur. Pressure in the rupture chamber only very briefly (or never, in many cases) exceeds the estimated minimum total pressure required to drive fully supersonic flow. Nozzle pressure measurements confirm this, suggesting that a system of shocks extends over a substantial downstream region of the nozzle, drifting upstream as total pressure decays.

Drug particles are entrained in both the transient starting flow and the quasi-steady nozzle flow in the conical nozzle. DGV measurements of the instantaneous velocity fields both inside and outside the nozzle indicate that the earliest particles are delivered in a wide cloud with typical velocities between 200 and 400 m/s. The transient starting process dominates the dynamics of these leading particles. The leading cloud is followed by a narrow, high-speed stream of particles which is embedded in the quasi-steady nozzle gas flow. Particles in this flow regime first emerge at a velocity of around 700 m/s, which decays to 400 m/s over the first millisecond of flow. This stream pours into the rear (upstream) face of the leading particle cloud, and imparts some momentum to it. The broad spectrum of particle delivery velocities, due to the disparity between the leading cloud and the quasi-steady stream, and the temporal decay and non-uniform transverse profile of velocity in quasi-steady stream, is an undesirable feature. It implies that a substantial proportion of the payload cannot be delivered to the skin at an optimal velocity (whatever that may be) for the desired interaction with biological tissue. DGV maps of the flow inside the nozzle clearly show separation of the high-speed particle jet from the nozzle wall, typically occurring at a distance of 15 to 30 mm upstream of the exit. In conjunction with the results of pressure measurements, this strongly suggests shock-induced flow separation.

Time-resolved measurements of particle velocity for silica, lidocaine and 4.7 μm polystyrene spheres are all approximately similar. This is consistent with time-integrated measurements of exit velocity, which are scattered between 300 and 500 m/s, and are independent of particle size and type. This insensitivity of velocity to particle size is due in part to generally low gas exit velocity in the conical nozzle, which allows time for particles of all types to approach the gas velocity. Numerical predictions (based on a simplified shock model) suggest that the shock system may also have a role in producing this particular effect. The small particles which accelerate most rapidly in the upstream region of the nozzle are also the quickest to decelerate on passing through a shock, because of their initially high absolute velocity and large (negative) slip velocity, as well as their high drag-to-mass ratio. Nozzle pressure measurements in the presence of payload indicate a period of low static pressure as particle delivery comes to an end, suggesting that particles do indeed travel faster than the gas in the downstream stretch of the nozzle. The appearance of particle-dependent features in the nozzle pressure history again suggests that the particles have a significant effect on the gas flow.

Numerical modelling of mass flows from the cylinder to the rupture chamber and through the nozzle has enabled prediction of quasi-steady rupture chamber pressure variations with excellent accuracy. These results draw attention to the mismatch of cylinder valve and nozzle throat mass flows as a major source of inefficiency, manifested as a loss of total pressure. Flow in the valve itself adds to these losses, with strong frictional effects and perhaps heat transfer effects. The consequence of this for the conical nozzle device at standard operating conditions is that rupture chamber pressure never rises above membrane burst pressure (approximately 20 bar), even though the initial cylinder pressure is 60 bar. The failure to attain supersonic exit flow in the conical nozzle is a result of this low total pressure.

8.2 Doppler Global Velocimetry

Time-integrated DGV, a technique previously used almost invariably in the study of steady flows, has been applied here to strongly unsteady short-duration flows. Theoretical analysis has shown that the velocity measured by DGV in this mode is not a true average over time or over particles, but a weighted average over particles, biased towards the brightest (normally, the largest) and slowest particles. An appropriately formulated comparison of time-resolved and time-integrated measurement results for a particular test case indicated that the time-integrated measurement underestimates the particle-averaged velocity by a margin of the order of 20% or less of the maximum velocity in the field. The time-integrated technique has proven very useful, providing a single-shot measurement which is representative of every particle in the flow. For example, it has enabled assessment of the influence of particle size and payload mass on particle exit velocity.

Time-resolved DGV has been demonstrated as an extremely powerful tool in this application, having revealed the structure and transient development of the particle cloud with great clarity. Time-resolved DGV measurements were in reasonable agreement with velocity estimates based on tracking of the particle cloud front from shot to shot, within the limits of uncertainty of those estimates. A second systematic comparison of time-integrated and time-resolved DGV data, designed to account for the biases associated with time-integrated measurement of unsteady flow, demonstrated excellent agreement between the two sets of results. The applicability of DGV to an internal flow, bounded by a curved window, has also been demonstrated.

8.3 Recommendations for Future Work

It is recommended that time-resolved DGV be adopted as a central element of any future research programme into the drug delivery devices. This technology has been shown to be a highly powerful tool for exploring the mechanisms which underlie device operation. Particle Image Velocimetry (PIV) is worthy of evaluation as a means of measuring drug particle velocities, though it is possible that the high particle concentrations observed in raw flow-field images would pose problems for PIV.

There is ample scope for more extensive application of DGV to the measurement of the velocity field inside the nozzle. The application of DGV to internal flows, in general, is an important topic which merits further investigation as an issue in the development of the technique. For example, practical limits on window curvature and thickness could be evaluated.

The behaviour of the gas phase in the two-phase flow is of fundamental importance. Its overall structure has been inferred in the current work from point pressure measurements and the indirect evidence of the particle velocity field. However, visualisation of the gas flow field itself would be a valuable complement to particle velocity measurements. One approach would to be to explore the use of DGV with light-scattering particles of sufficiently low size and/or density to follow the gas flow with an acceptably small velocity lag. The silica particles used in the present work appear to be promising in this regard, although more work is required. Alternatively, the well-established shadowgraph, schlieren or Mach-Zehner interferometry techniques could provide qualitative visualisation of shocks and the air-helium contact surface in the drug delivery devices. Time-resolved schlieren photography has already been used successfully for this purpose in recent work by Dr. Mark Kendall.

In the interests of simplicity in this first study, all the drug delivery devices tested have been fired into free air. In reality, devices fire onto the skin at a typical distance of 10 mm from the nozzle exit plane. Flow features associated with impingement of the gas-particle jet on the skin must cause some deceleration of the particles immediately before impact. Moreover, current designs for production devices feature a porous shroud around the downstream part of the device, which seals against the skin and forms a closed chamber around the device exit and the target region. Its functions are to contain unwanted particle efflux and attenuate the primary shock (which otherwise registers audibly as an uncomfortably loud bang). The impact of this design feature on the flow-field is as yet unknown, but it can be expected to influence the nozzle flow structures by causing an increase in effective back pressure. The silencer and the impingement flow are essential features of working devices, which may have significant effects on particle delivery velocity. They should therefore be incorporated in future studies.

The basic numerical models developed in this research could be developed further. For example, the quasi-steady particle dynamics model could be modified, without a fundamental increase in complexity, to account for the influence of the particle phase on the momentum of the gas and the influence of the boundary layer. However, it seems unlikely that numerical models of this simplicity and compactness will ever achieve useful predictions of the transient or multi-dimensional aspects of the flow. In the longer term, therefore, the application of full computational fluid dynamics methods is recommended.

Ultimately, all current and future research into the fluid dynamics of transdermal powder drug delivery devices must be directed towards improvement of the technology. Work presented in this thesis points specifically to the non-uniformity of particle delivery velocity and nozzle flow separation as undesirable characteristics of a particular current device, which call for rectification. However, it is hoped that the insights gained in this work can help to guide the next stages of research in a more general way. Perhaps an appropriate long-term goal for ongoing research into transdermal drug delivery gas-powder flows is the development of a proven master device configuration, along with quantitative design methods to adjust the basic device to the demands of each individual particular therapeutic application.

References

1 Bellhouse B. J., Burkoth T. L., Hewson G., Longridge D. J., Muddle A. G., Sarphie D. F., *Transdermal and Transmucosal Powdered Drug Delivery*, in: Bruck S. D. (ed.), *Critical Reviews in Therapeutic Drug Carrier Systems*, Begell House, New York (in preparation)

2 Sanford J. C., Wolf E. D., Allen N. K., *Method for Transporting Substances into Living Cells and Tissues and Apparatus Therefor*, US Patent 4945050, 1990

3 Sarphie D. F., *Gas Propulsion of Microprojectiles for the Transformation of Biological Cells,* D. Phil. Thesis, University of Oxford, 1992

4 Bellhouse B. J., Sarphie D. F., Greenford J. C., International patent WO94/24263, 1994

5 —, *Prospectus*, PowderJect Pharmaceuticals plc, 1997

6 Ganderton D., *Current Innovations in Drug Delivery*, Drug Design and Delivery 2:1-7, 1987

7 Ranade V. V., Hollinger M. A., *Drug Delivery Systems*, CRC Press, Boca Raton, 1995

8 Meyers J. F., *Development of Doppler Global Velocimetry as a Flow Diagnostics Tool*, Measurement Science and Technology 6:6, 1995

9 Ainsworth R. W., Thorpe S. J., Manners R. J, *A New Approach to Flow-Field Measurement — a View of Doppler Global Velocimetry Techniques,* Journal of Heat and Fluid Flow 18:116-130, 1997

10 Anderson J. A., *Modern Compressible Flow*, McGraw-Hill, New York, 1992

11 Shapiro A. H., *The Dynamics and Thermodynamics of Compressible Fluid Flow*, Ronald, New York, 1953

12 Stodola A., *Steam and Gas Turbines*, McGraw-Hill, New York, 1927

13 Adamson T. C., Messiter A. F., *Analysis of Two-Dimensional Interactions Between Shock Waves and Boundary Layers*, Annual Review of Fluid Mechanics 12:103-138, 1980

14 Meier G. E. A., Szumowski A. P., Selerowicz W. C., *Self-excited Oscillations in Internal Transonic Flows*, Progress in the Aerospace Sciences 27:145-200, 1990

15 Szumowski A. P., Obermeier F., Meier G. E. A., *Oscillation Modes of Laval Nozzle Flow*, Experiments in Fluids 18:145-152, 1995

16 Biedron R. T., Adamson T. C., *Unsteady Flow in a Supercritical Supersonic Diffuser*, AIAA Journal 26:11, 1988

17 Chisnell R. F., *The motion of a shock wave in a channel, with applications to cylindrical and spherical shock waves*, Journal of Fluid Mechanics 2:3, 1957

18 Whitham G. B., *On the Propagation of Shock Waves Through Regions of Non-uniform Area or Flow*, Journal of Fluid Mechanics 4:4, 1958

19 Smith C. E., *An analytic study of the starting process in a hypersonic nozzle*, in: Geidt W. H., Saloman L. (eds.), *Proceedings of the 1964 Heat Transfer and Fluid Mechanics Institute*, Stanford University Press, Stanford, 1964

20 Smith C. E., *The Starting Process in a Hypersonic Nozzle,* Journal of Fluid Mechanics 24:4, 1966

21 Amann, H. O., *Experimental Study of the Starting Process in a Reflection Nozzle*, Physics of Fluids 12:5 Supplement I, 1969

22 Amann H. O., Reichenbach H, *Unsteady Flow Phenomena in Shock-Tube Nozzles*, in: Bershader D., Griffith W. (eds.), *Recent Developments in Shock Tube Research : Proceedings of the Ninth International Shock Tube Symposium*, Stanford University Press, Stanford, 1973

23 Tocarcik-Polsky S., Cambier J.-L., *Numerical Study of Transient Flow Phenomena in Shock Tunnels*, AIAA Journal 32:5, 1994

24 Gvozdeva L. G., Zhilin, Y. V., *Formation of a Quasistationary Jet within a Nozzle During its Start-up*, Fluid Mechanics (translation by Plenum, New York from Russian *Mekanika Zhidkosti i Gaza*) 1:76-82, 1977

25 Britan A. B., Vasil'ev E. I., *Investigation of the Starting Process in the Shaped Nozzle of a Large-Diameter Shock Tube*, Fluid Dynamics (translation by Plenum, New York from Russian *Mekhanika Zhidkosti i Gaza*) 21:5, 1986

26 Rudinger G., *Fundamentals of Gas-Particle Flow*, Elsevier, Amsterdam, 1980

27 Soo S. L., *Multiphase Fluid Dynamics*, Science Press / Gower Technical, Beijing, 1990

28 Anilkumar A. V., Sparks R. S. J., Sturvetant B., *Geological implications and applications of high-velocity two-phase flow experiments*, Journal of Vulcanology and Geothermal Research 56:145-160, 1993

29 Rogue X., Rodriguez G., Haas J. F., Saurel R., *Experimental and Numerical Investigation of the Shock-Induced Fluidization of a Particles Bed,* Shock Waves 8:29-45, 1998

30 Rogue X., Rodriguez G., Haas J. F., Saurel R., *Shock-induced Fluidization of a Dense Bed of Particles,* 21st International Symposium on Shock Waves, Great Keppel Island, Australia, 1997

31 Bode R., Ferch H., Fratzscher H., *Technical Bulletin Pigments no. 11: Basic Characteristics and Applications of AEROSIL*, Degussa, Frankfurt, 1982

32 Ainsworth R. W., Dietz A. J., Nunn T. A., *The Use of Semi-Conductor Sensors for Blade Surface Measurements in a Model Turbine Stage*, Journal of Engineering for Gas Turbines and Power 113:261-268, 1991

33 Roberts M. J., *Needle-less Drug Delivery*, Fourth Year Project Report, Department of Engineering Science, University of Oxford, 1997

34 Komine H., Brosnan S. J., Litton A. B., Staeppers E. A., *Real-Time, Doppler Global Velocimetry*, AIAA-91-0337, 1991

35 Ainsworth R. W., Thorpe S. J., *The Development of a Doppler Global Velocimeter for Transonic Turbine Applications*, ASME 94-GT-146, 1994

36 Thorpe S. J., Ainsworth R. W., Manners R. J., *The Development of a Doppler Global Velocimeter and its Application to a Free Jet Flow*, ASME/EALA 6th International Conference on Laser Anemometry Advances and Applications, Hilton Head Island, 1995

37 Thorpe S. J., Ainsworth R. W., Manners R. J., *Time-Averaged Free-Jet Measurements using Doppler Global Velocimetry,* ASME Symposium on Laser Anemometry and Experimental and Numerical Flow Visualisation, San Diego, 1996

38 Durst, F., Melling, A. and Whitelaw, J. H., *Principles and Practice of Laser-Doppler Velocimetry,* Academic Press, London, 1976

39 Adrian, R. J., *Particle-Imaging Techniques for Experimental Fluid Mechanics,* Annual Review of Fluid Mechanics 23, 1991

40 Meyers J. F., Lee J. W., Cavone A. A., Suzuki, K. E., *Investigation of the Vortical Flow above an F/A-18 Using Doppler Global Velocimetry*, in: *Laser Anemometry Advances and Applications*, SPIE 2052, 1993

41 Thorpe S. J., *A Study of Doppler Global Velocimetry in its Application to Aerodynamic Flows*, D.Phil. thesis, University of Oxford, 1996

42 Manners R. J., Thorpe S. J., Ainsworth R. W., *Image Processing Techniques for Doppler Global Velocimetry*, I. Mech. E. International Seminar on Optical Methods and Data Processing in Heat and Fluid Flow, London, 1996

43 Thorpe S. J., Quinlan N. J., Ainsworth R. W., *The Characterisation and Use of a Pulsed Neodymium YAG Laser in DGV Applications*, 8th International Conference on Laser Anemometry – Advances and Applications, Rome, 1999

44 Roehle I., Schodl R., *Applications of Three Dimensional Doppler Global Velocimetry to Turbo Machinery and Wind Tunnel Flows*, 7th International Conference on Laser Anemometry – Advances and Applications, Karlsruhe, 1997

45 —, *CCD Camera Manual for Wright Instruments Peltier Cooled CCD Cameras*, Wright Instruments Ltd., Enfield, 1994

46 Beesley, M. J., *Lasers and Their Applications*, Taylor and Francis, London, 1971

47 —, *Model 164 General Laboratory Ion Laser Instruction Manual*, Spectra Physics, Mountain View CA

48 —, *Pulsed Nd:YAG Lasers — User's Manual*, Spectra Physics, Mountain View CA, 1993

49 —, *Model 6530 Instruction Manual*, Spectra Physics, Mountain View CA, 1992

50 Hercher M., *The Spherical Mirror Fabry-Perot Interferometer*, Applied Optics 7:5, 1968

51 Bevington P. R., *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York, 1969

52 Van Wylen G., Sonntag R. E., *Fundamentals of Classical Thermodynamics*, Wiley, New York, 1985

53 Gray D. E. (ed.), *American Institute of Physics Handbook 3rd ed.*, McGraw-Hill, New York, 1972

54 Smith M. W., Northam G. B., Drummond J. P., *Application of Absorption Filter Planar Doppler Velocimetry to Sonic and Supersonic Jets*, AIAA Journal 34:3, 1996

55 Quinlan N. J., Ainsworth R. W., Bellhouse B. J., Manners R. J., Thorpe S. J., *Application of Doppler Global Velocimetry to Supersonic Gas-Particle Flows in Drug Delivery,* 7th International Conference on Laser Anemometry – Advances and Applications, Karlsruhe, 1997

56 Bellhouse B. J., Quinlan N. J., Ainsworth R. W., *Needle-less delivery of drugs, in dry powder form, using shock waves and supersonic gas flow*, 21st International Symposium on Shock Waves, Great Keppel Island, 1997

57 Quinlan N. J., Thorpe S. J., Ainsworth R. W., *Time-Resolved Doppler Global Velocimetry of Gas-Particle Flows in Transdermal Powder Drug Delivery,* 8th International Conference on Laser Anemometry – Advances and Applications, Rome, 1999

58 Brandt O., Rajathurai A. M., Roth P., *First Observations on the Break-up of Particle Agglomerates in Shock Waves*, Experiments in Fluids 5:2, 1987

59 Strecker J. J. F., Roth P., *Particle Breakup in Weak Shock Waves: Preliminary Observations*, Journal of Aerosol Science 23, Supplement 1, 1992

60 White F. M., *Fluid Mechanics*, McGraw-Hill, Singapore, 1986

61 Incropera F. P., DeWitt D. P., *Fundamentals of Heat and Mass Transfer*, Wiley, New York, 1990

62 Liou M.-S., *Solutions of One-Dimensional Steady Nozzle Flow Revisited,* AIAA Journal 26:5, 1988

63 Henderson C. B., *Drag Coefficients of Spheres in Continuum and Rarefied Flows,* AIAA Journal 14:6, 1976

64 Bailey A. B., Hiatt J., *Sphere Drag Coefficients for a Broad Range of Mach and Reynolds Numbers,* AIAA Journal 8:11, 1970

65 Kurian J., Das H. K, *Studies of Shock Wave Propagation in Gas-Particle Mixtures,* 21st International Symposium on Shock Waves, Great Keppel Island, Australia, 1997

66 Igra O., Takayama K., *Shock Tube Study of the Drag Coefficient of a Sphere in a Nonstationary Flow,* 18th International Symposium on Shock Waves, Sendai, Japan, 1991