A Picosecond Optoelectronic Cross Correlator using a Gain Modulated Avalanche Photodiode for Measuring the Impulse Response of Tissue.

David R. Kirkby B.Sc. M.Sc.
Department of Medical Physics and Bioengineering,
University College London.

Thesis submitted for the degree of Doctor of Philosophy (Ph.D.) at the University of London.

Supervisor: Professor David T. Delpy, FRS.
Second supervisor: Dr. Mark Cope.

April 1999.

Abstract

Human tissue is relatively transparent to light between 700 and 1000 nm in the near infra-red (NIR). NIR spectroscopy is a technique that can measure non-invasively and safely, the optical properties of tissue. Several different types of spectroscopic instrumentation have currently been developed, ranging from simple continuous intensity systems, through to complex time and frequency resolved techniques. This thesis describes the development of a near infra-red time-resolved system, using an inexpensive avalanche photodiode (APD) detector and a microwave step recovery diode (SRD) in a novel way to implement a totally electronic cross-correlator, with no moving parts. The aim of the work was to develop a simple instrument to monitor scattering changes in tissue during laser induced thermal therapy.

The APD was gain-modulated by rapidly varying the bias voltage using electrical pulses generated by the SRD (120 ps full width half maximum (FWHM) and 8 V in amplitude). The resulting cross-correlator had a temporal resolution of 275 ps FWHM - significantly faster than the 750 ps FWHM of the APD when operating with a conventional fixed bias voltage. Spurious responses caused by the SRD were observed, which were removed by the addition of Schottky diodes on the SRD's output, although this slightly degraded the system temporal resolution from 275 to 380 ps FWHM.

The ability of the system to monitor scattering changes was tested using an Intralipid^{TM} phantom containing infra-red absorbing dye. An 800 nm fibre coupled mode-locked (2 ps pulse width) laser source was used with the cross-correlator measuring the temporal point spread function (TPSF) at 5 to 30 mm away from the source fibre. Five different numerical algorithms to derive the scattering coefficient from the measured TPSF were compared. The optimum choice of algorithm was found to depend on whether absolute accuracy or minimum computation time is the most important consideration.

Acknowledgements

I would like to thank the following for their help during this research.

Dr. Klaus Berndt, whose research publications first got the idea for the project going and for very helpful discussions at one of the SPIE conferences.

My supervisor, Professor David Delpy FRS, for first noting the publications of Dr. Berndt and then helping me write the grant proposal, which finally secured funding. His help at every stage of the research was invaluable.

Elizabeth Hillman, for proof reading parts of my thesis and her invaluable help with dispelling doubts I had about the Monte Carlo results. Beth, within a few months of starting her Ph.D., made what was probably the most significant single contribution of anyone else, towards me getting this Ph.D. finished. Thanks Beth.

Barry Hall, Joy Hall and Dr. Jeremy Chipchase, all of whom proof read parts of my thesis. They are of course in no way responsible for any remaining mistakes.

Dr. Jem Hebden, who had the very useful habit of suggesting computer programmes that were appropriate and for helping me implement some of them. Also for the time he spent in meetings with my supervisor and I discussing aspects of this work.

Dr. Mark Cope, for his help on several aspects, especially the phantom construction, noise characteristics of lock-in amplifiers, but also countless other things. Marks superb knowledge base that makes him so popular with his students, was tapped by me on several occasions.

Dr. Paul Ripley, for his advice on PDT and LITT.

Veronica Hollis for her help in using with the CCD spectrometer, which had been put together and optimised by Dr. Roger Springett.

My wife Linda, for her support during the whole of the Ph.D. and who so often met me at Althorne Station for the 2220 or 2301 train, to save me the 1.5 mile walk home - most of which is up hill. Thanks for all your help Lin.

Dedication

This thesis is dedicated to the memory of my late Grandmother, Dorothy May Impey, who would have so much enjoyed to see her Grandson get the title "Doctor".

Chapter 2.

Chapter 3.

Chapter 7.

Design and Construction of a Tissue-like Optical Phantom. 177

Chapter 10.

Discussion and Conclusions.   244

Publications arising from this work.   250

Appendix 1.

Abbreviated data sheets for some electronic devices used.   251

References.   253

List of Tables.

List of Figures

Abbreviations.

APD	Avalanche photodiode
BPF	Band-pass filter
CCD	Charge coupled device
dBm	Power, referenced to 1 mW (0 dBm = 1 mW, -20 dBm = 10 μW)
EM	Electromagnetic radiation
FEM	Finite element modelling
FSD	Full scale deflection
FWHM	Full width half maximum
GaAs	Gallium arsenide
Ge	Germanium
Hb	Deoxygenated haemoglobin
HbO_2	Oxygenated haemoglobin
HeNe	Helium-neon
HPF	High-pass filter
ILP	Interstitial laser photocoagulation
KM	Kubelka Munk
KTP	$KTiOP0_4$
LC	Inductance capacitance
LITT	Laser induced thermal therapy
LT[x]	Laplace transform of x.
MCP	Microchannel plate
MHV/H4	Miniature high voltage (coaxial connector for high voltages)
MRI	Magnetic resonance imaging
Nd:YAG	Neodymium Yttrium aluminium garnet.
NEP	Noise equivalent power
NIR	Near infra-red
NIRI	Near infra-red imaging
NIRS	Near infra-red spectroscopy
PDT	Photodynamic therapy
PIN	P type - Intrinsic - N type (semiconductor diode)
PMT	Photomultiplier tube
PN	P type - N type (semiconductor diode)

RLC	Resistance inductance capacitance
SHV	Super high voltage (coaxial connector for high voltages)
Si	Silicon
SMA	Subminiature connector
SPSF	Spatial point spread function
SRD	Step recovery diode
TPSF	Temporal point spread function
UV	Ultra violet
VSWR	Voltage standing wave ratio

Symbols.

Symbols.	Definition	SI units
(S/N)	Signal to noise ratio	
$\langle i_T^2 \rangle$	Mean-square thermal noise current	A^2
α_e	Electron ionisation rate	m^{-1}
α_h	Hole ionisation rate	m^{-1}
β	Propagation constant of free space	m^{-1}
γ	Exponent in APD gain equation.	
ε_g	Band gap of a semiconductor	J
η	Quantum efficiency	
λ	Wavelength	m
λ_{th}	Cutoff wavelength of a photodiode	m
μ_a	Absorption coefficient	m^{-1}
μ_{eff}	Effective attenuation coefficient	m^{-1}
μ_s	Scattering coefficient	m^{-1}
μ_s'	Transport scattering coefficient	m^{-1}
μ_t	Attenuation coefficient	m^{-1}
ν	Frequency of light	Hz
σ_a	Absorption cross section	m^2
σ_s	Scattering cross section	m^2
$\sigma_s(\vec{r},\vec{e}_s,\vec{e}_i)$	Differential scattering coefficient	m^{-1} sr^{-1}
σ_t	Total (or extinction) cross section	m^2
τ	Time delay	s
τ_c	Lifetime of a carrier in the i-region	s
τ_{mc}	Minority carrier lifetime of a semiconductor	s
τ_F	Lifetime of fluorescence material	s
ϕ_{con}	Contact potential	V
ϕ_n	Phase of nth mode in laser	
χ	Excess noise index in an APD.	
ψ	Radiant energy fluence rate	Wm^{-2}
Ψ	Radiant energy fluence	Jm^{-2}
Ω	Resistance	Ω
ω_n	Angular frequency of nth laser mode	s^{-1}

a	Single scattering albedo	
A	Area	m^2
a_{KM}	Variable used in Kubelka-Munk theory	
A_{KM}	Kubelka-Munk coefficient	m^{-1}
B	Bandwidth	Hz
$B_i(t)$	Basis functions.	
b_{KM}	Variable used in Kubelka-Munk theory	
c	Velocity of EM radiation in a vacuum	$3x10^8$ ms^{-1}
C_j	Reverse biased junction capacitance	F
C_p	Package capacitance	F
D_{laser}	Diameter of a laser beam	m
D_{ph}	Diameter of the active area of a photodiode	m
E(t)	Time dependent electric field	V/m
E_1	Energy stored in a capacitor @$V=V_1$	J
E_2	Energy stored in a capacitor @$V=V_2$	J
E_n	E-field amplitude of nth mode in laser	V/m
E_λ	Extinction coefficient at wavelength λ (nm)	m^{-1}
F	Noise factor of an amplifier	
F(M)	Noise factor of an APD	
$f(\vec{e}_s,\vec{e}_i)$	Scattering phase function	sr^{-1}
$F_1(j\omega)$	Fourier transform of $f_1(t)$.	
f_{3dB}	3 dB bandwidth	Hz
f_{max}	Maximum frequency of signal	Hz
$FT^{-1}(F_1(j\omega))$	Inverse Fourier transform of $F_1(j\omega)$	
g	Mean cosine of scattering angle	
h	Plancks constant	$6.62617x10^{-34}$ Js
I_B	Background current of a photodiode	A
I_D	Dark current of a photodiode	A
I_{eq}	Equivalent current for all noise sources	A
i_{ph}	RMS signal current of a photodiode	A
I_{ph}	Photocurrent	A
i_s	Shot noise current	A
j	$\sqrt{-1}$	
J_{ph}	Photocurrent density	Am^{-2}

k	Boltzmann constant	1.38×10^{-23} J K^{-1}
K	Ratio of hole to electron ionisation rate.	
K$'$	Inverse of K	
$L(\vec{e}_s)$	Energy radiance	Wm^{-2}sr^{-1}
L_{cav}	Length of laser cavity	
L_p	Package inductance	H
l_s	Length travelled before scattering	m
m	Modulation depth	
M	Gain of an APD	
m$'$	Gain ratio of APD	
n	Refractive index	
N_c	Number of carriers generated/m^2/s	m^{-2}s^{-1}
N_d	Doping density of a semiconductor	m^{-3}
N_e	Number of elements in a finite element mesh	
NEP	Noise equivalent power	WHz$^{-1/2}$
N_{par}	Number of particles	
N_{ph}	Number of photons	
$P(\omega)$	Instantaneous optical power	W
P_{avg}	Mean laser power	W
P_d	Power density of optical radiation	Wm^{-2}
$P_{min}(S/N)$	Minimum power for a given S/N ratio	W
P_o	Optical power	W
P_{peak}	Peak power reaching a photodiode	W
q	Charge on an electron	1.6×10^{-19} C
Q	2π energy stored/energy lost	
Q_0	Initial charge	C
Q_{pp}	Charge moved per laser pulse	C
$R(\vec{r},t)$	Residual	
R_{eq}	Equivalent resistance	Ω
R_i	Internal reflectance	
R_i	Input resistance of an amplifier	Ω
R_j	Photodiode parallel resistance	Ω
R_L	Load resistance	Ω
R_{PD_s}	Photodiode series resistance	Ω

R_s	Scattered (or diffuse) reflectance	
R_{sp}	Specular reflectance	
R_{SRD_s}	Dynamic series resistance of SRD	Ω
R_t	Total reflectance	
$r_{xy}(\tau)$	Cross-correlation of signals x and y.	
S_{KM}	Kubelka-Munk coefficient	m^{-1}
$S_N(\vec{r})$	N^{th} Legendre expansion	
T_0	Integration time	s
$t_{10\rightarrow90}$	Rise time	s
$t_{90\rightarrow10}$	Fall time	s
t_a	Time for avalanche to develop	s
T_p	Primary transmittance	
t_s	SRD snap time	s
t_t	Total transit time	s
T_t	Total transmittance	
t_{te}	Electron transit time	s
t_{th}	Hole transit time	s
V	Volume	m^3
V_{100}	Bias on an APD for a gain of 100	V
V_B	Breakdown voltage of a semiconductor device	V
v_{bc}	Blood cell velocity.	ms^{-1}
V_{DC}	DC bias voltage for APD	V
V_M	Bias on an APD for a gain of M	V
V_{mod}	Modulation voltage	V
V_{out}	Output voltage from photodiode	V
V_{ph}	Voltage on photodiode at zero bias	V
V_R	Reverse voltage applied to a photodiode	V
v_s	Saturated velocity of a carrier	ms^{-1}
v_{se}	Saturated electron velocity	ms^{-1}
v_{sh}	Saturated hole velocity	ms^{-1}
w	Width	m
w_{ab}	Width of absorption layer	m
w_{av}	Width of avalanche layer	m
w_d	Width of depletion layer (or i layer)	m

x	Thickness of a sample	m
z_0	Position photons initially scattered at	m
$\Re$	Responsivity	A/W

Chapter 1.

Introduction.

1.1   Purpose.

When a short pulse of light a few picoseconds long, such as can be easily produced with a mode-locked laser[1], is directed into tissue, the light that is detected a distance away will be temporally broadened by a considerable amount. The exiting pulse will be up to a few nanoseconds in duration[2]. The shape of the exiting pulse, known as the temporal point spread function (TPSF), or impulse response of the tissue, depends upon the optical properties of the tissue, which are frequently characterised by three parameters, which are defined fully later - the absorption coefficient μ_a, scattering coefficient μ_s and a scattering anisotropy factor g. Under certain restricted situations, it is possible to determine the absorption coefficient μ_a and the transport scattering coefficient μ_s', where $\mu_s' = \mu_s(1-g)$, from the TPSF. This is very useful, as many applications of optical radiation in medicine require knowledge of μ_a and μ_s' for accurate analysis.

The purpose of this work was to construct an instrument that would be able to measure the TPSF of tissue experimentally and from these measurements deduce the transport scattering coefficient μ_s'. Although there are a number of ways of measuring the impulse response of the tissue, many are very large and expensive in terms of the hardware required. The approach used here is less expensive than many and being entirely electronic, fairly small, with no moving components, it can in principle be made sufficiently rugged and portable to use at the bedside.

1.2   Guide to this thesis.

This thesis consists of several sections. The first section (chapters 1 to 5), introduces optical monitoring in medicine, background work and the necessary terminology. The second section (chapter 6) describes the electronic system developed for determining the impulse response of tissue. This is the main part of the thesis. The third section, chapter 7 describes a tissue-like optical phantom that was constructed to allow measurement with the electronic instrumentation. Chapter 8 describes a Monte Carlo model developed to model light transport in tissue. Chapter 9 shows the experimental results measured with the instrumentation, the computer results calculated using the Monte Carlo model and describes five fitting procedures

from which the optical properties of measured data could be determined. Chapter 10 is a discussion of the work, and discusses further work.

Section 1.3 introduces a small selection of the large number of uses of optical radiation in medicine. Section 1.4 introduces the terminology used in tissue optics. Since for the accurate use of light in medicine, it is essential to understand how light propagates in tissue, section 1.5 discusses some approximate models of light transport in tissue, which predict light distributions given the optical properties of the tissue. Approximate models of light transport in tissue are necessary, since it is not practicable to solve Maxwell's equations, which would be the most rigorous method. Section 1.5 also discusses the inverse problem of estimating μ_a and μ_s or μ_s' from measured data. Section 1.7 discusses briefly some of the many ways of measuring the temporal profile of a light pulse, although section 1.7.6, which describes a cross-correlation technique using a gain-modulated avalanche photodiode is more complete, since this technique is most relevant to this thesis. Section 1.8 shows the basic design decided upon. This is seen to require four major components - a pulsed laser source, an avalanche photodiode, a fast pulse generator and a variable time delay mechanism.

The theory of the pulsed laser source, avalanche photodiode, fast pulse generator and variable time delay mechanism are described in chapters 2, 3, 4 and 5 respectively, along with the reasons for choosing the particular devices used.

The complete instrument design is given in chapter 6. This is the major part of the thesis. The chapter discusses some early prototypes developed, along with the development of the instrument to its current state.

Chapter 7 describes a tissue-like liquid phantom, using Intralipid^{TM} and dye, built to allow testing of the abilities of various algorithms for determining μ_s' from the experimental data. Chapter 8 describes a Monte Carlo model that models what happens when fibres are implanted in tissue, or a tissue-like phantom such as that described in chapter 7.

Chapter 9 shows the results which fall into 3 categories. First there is the simulations performed with the Monte Carlo model described in chapter 8. Secondly there are the experimental results measured on the phantom with the cross-correlator. Finally, the results of five different fitting algorithms are presented, which were developed to estimate μ_s' from the experimental data measured with the cross-correlator.

Chapter 10 discusses the work, in detail and gives an indication of the further work that might be carried out - time and money permitting!

A list of publications arising from this work is given on page 250. Abreviated data

sheets for the most important electronic componenets are given in appendix 1 on page 251.

1.3   Medical uses of optical radiation.

There are a number of medical uses of light, which fall into two main categories - therapeutic or diagnostic. The following few examples are by no means a complete list, but merely illustrate the range of uses of optical radiation in medicine and the importance of the tissue parameters μ_a and μ_s and hence the requirement for hardware capable of measuring them *in vivo*.

1.3.1   Photodynamic therapy (PDT).

Photodynamic therapy (PDT)[3,4,5] is a therapeutic method used for the destruction of tumour tissue, in both hollow organs (eg colon) and solid organs (eg pancreas). A photosensitive drug is first administered that collects in all tissue, but the concentration becomes higher in the tumour, primarily since tumour tissue is fast growing, but is not very efficient at disposing of its waste. Sometime later (typically two to four days), when the healthy cells have flushed out most of the photosensitive drug, light at a suitable wavelength is directed at the tumour site from a laser. The wavelength chosen must have good tissue penetration, and must also activate the photosensitive drug, so the wavelength used depends on the drug selected, but is typically 630 to 650 nm. The photosensitive drug absorbs the light, becoming activated, then produces toxic substances (usually singlet oxygen $^1O_2$), destroying the surrounding tissue. The $^1O_2$ either kills the cancer cell directly, or the damage it does impairs the blood supply to the cancer, which results in cell death by ischaemic processes.

After treatment, the patient is sensitive to the light, and must keep away from modest light for a short period, and strong light for a longer period, that depends on the drug used. For the first generation of drugs used (eg Photofrin), this was up to 3 months. The patient would usually wear a light meter during this time, to monitor the amount of light received. For the second generation of drugs, the time period is much reduced to two or four weeks. (Sunblock does not help, as it stops the ultra violet, not the visible light that activates the drug). Other benefits of the newer generation of drugs, such as Foscan™, include bigger necrosis (not always a good idea - depending on tumour size), better tissue selectivity and a reduced light dose.

The treatment takes several days for the results to be visible to the naked eye, but a more rapid monitor during treatment is desirable. Since PDT requires the presence of oxygen, the photosensitive drug and light, techniques are used to monitor the levels of these 3 important

parameters. Photo-bleaching of the drug occurs during treatment, so monitoring techniques have been developed to monitor for bleaching of the drug, or a fall in the oxygen in the tissue being treated.

PDT has few side effects (the main one being the light sensitivity of the patient), and can be re-treated if and when necessary, making it a flexible therapy.

1.3.2 Laser induced thermal therapy (LITT).

Laser induced thermal therapy (LITT) or interstitial laser photocoagulation (ILP) is a another therapeutic method of destroying tumour tissue, but unlike PDT, is only suitable for solid organs. A laser, typically a 1064 nm Nd:Yag with a power of around 1.5 to 2.5 W[6,7] is coupled to the tumour site by optical fibre, although other wavelengths, such as a 805 nm diode laser have also been used[8]. The tissue absorbs the light and hence rises in temperature. The temperature rise propagates throughout the tissue by thermal diffusion. At least 60°C is required to cause irreversible denaturisation of the proteins. Typically exposure times are of the order of 500-1000 s[7], giving a typical exposure of 1500 J of energy. Thermal necrosis of predictable extent up to 16 mm in diameter is possible[9], although by using more than one optical fibre to feed the power in, it is possible to treat larger areas[6].

LITT or ILP has the disadvantage over PDT in that it is not so selective - the heat dissipates to all surrounding tissue. However, it does have several advantages over PDT. Firstly, there is no patient sensitivity to light after the treatment. Secondly it can be used where the photosensitive drugs do not accumulate well, such as in the breast. It can also be used in the liver, where PDT is not very practical, due to the very high absorption of light. Also, using diffusion sensitive real-time magnetic resonance imaging[10] it is possible to visualise the temporal and spatial development of the temperature rise caused by the laser, something that is not possible with PDT.

Essenpries[11] showed that following necrosis, there is an increase in the scattering coefficient of the tissue and Benaron et al[12] have also shown that changes in the scattering coefficient can be observed *in vitro* using optical techniques (in this instance, changes due to freezing of the tissue). It is envisaged that by developing instrumentation able to detect changes in μ_s', and measuring in real-time the scattering coefficient at a few sites on the perimeter of a tumour, it would be possible to determine if necrosis has occurred at the perimeter and hence when the treatment should be ceased. Figure 1.1 shows the idea of this method. This thesis describes development of such an instrument.
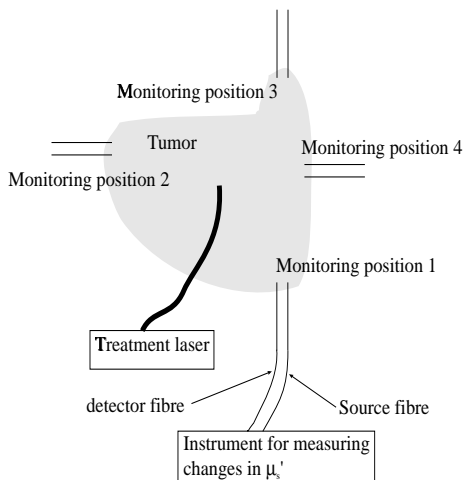
Figure 1.1 Diagram showing how it might be possible to determine when a tumour has developed necrosis, by measuring μ_s' at several perimeter sites.

1.3.3   Near infra-red spectroscopy (NIRS).

Near infra-red spectroscopy (NIRS) is used to determine the concentration of chromophores in tissue, especially the oxygen carrying molecule in blood (haemoglobin) and the enzyme cytochrome oxidase[13]. Haemoglobin has two principle states of oxygenation, known as oxygenated haemoglobin (HbO_2) and deoxygenated haemoglobin (Hb). Since only oxygenated haemoglobin can carry oxygen around the body, it is obviously essential to ensure the blood is sufficiently oxygenated. Unfortunately, in some instances, too much oxygen can also be harmful, as for example in premature babies where it can cause blindness[14,15]. The absorption of haemoglobin as a function of wavelength in the visible region (450 to 700 nm) and in the near infra-red (700 to 1000 nm) is shown in figure 1.2. Note that the scales on the two graphs differ by a factor of 40,
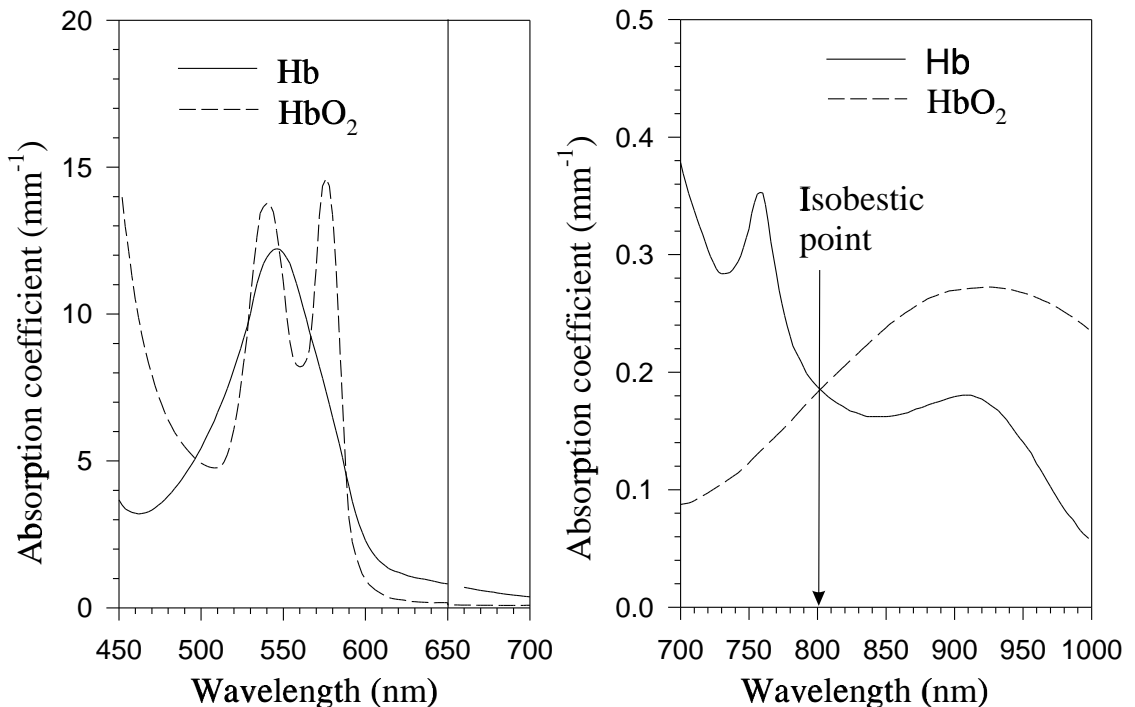
Figure 1.2 Absorption coefficient of haemoglobin, in both the oxygenated form (HbO_2) and the deoxygenated form (Hb) in the visible (left) and NIR (right).

since the absorption in the near infra-red (NIR) is much less than in the visible. Note also that

in the NIR there is an isobestic point between the two spectra at 800 nm. Above 800 nm the absorption of the oxygenated form is greatest, but below 800 nm the greatest absorption is caused by the deoxygenated form. In the visible region (400 to 700 nm) the absorptions are much stronger, which limits light penetration into the tissue, but the differences in the spectra are also more marked, and it is these differences which explain the well known effect that people become more blue when starved of oxygen. There are several isobestic points in the visible region at 500.9, 527.5, 547.5, 572.6 and 585.4 nm. This data is originally from Horecker[16], but is reproduced from the Ph.D. thesis of Cope[17].

1.3.4   Laser Doppler blood flow measurements.

Laser Doppler[18,19] was first used by Riva[20] to measure the blood velocity in the retinal artery of an albino rabbit, using a 10 mW helium-neon (HeNe) laser (λ=632.8 nm). The Doppler power spectrum showed a peak at 3 kHz. Assuming single scattering, the frequency of the Doppler shifted signal has a maximum value given by

$$f_{max} = \frac{2v_{bc}}{\lambda} \qquad\qquad \textbf{1.1}$$

where v_{bc} is the blood cell velocity, but will fall below this if the direction of travel is not directly in line with the laser beam. Stern[21] later showed that the movement of red blood cells would cause laser light directed at the skin to be Doppler[22] shifted as a result of scattering from moving red blood cells. Using a 15 mW 632.8 nm HeNe laser, aimed at the finger tip, he measured the intensity of the Doppler shifted component of light, both before and after occlusion of the brachial artery. The amplitude spectrum was much lower following occlusion and contained less high frequency components. The intensity of the Doppler signal was measured by mixing both the original (un-shifted) and the Doppler shifted laser light on a photomultiplier tube. The photomultiplier tube performs the down conversion from the very high frequencies of light (~10^{14} Hz) to the Doppler baseband ($\leq$ 20 kHz).

There are a number of problems with Doppler based blood flow measurements, which make it difficult to quantify the measurement. Since light will be multiple scattered in tissue, a photon with a large Doppler shift may come from either one interaction with a fast moving blood cell or a number of interactions with slow moving blood cells - there is no way to know for certain. The measurement volume is also difficult to quantify. Some workers have tried to vary the penetration depth by use of two or more different laser wavelengths[23], which will have different penetrations depths or by using a variable coherence length source[24]. Essex[25]

developed a 2D scanner for imaging blood flow and showed images from the hand, chest etc. Although not calibrated, its claimed that by comparing images from an adjacent or opposite part of the body, useful clinical information is obtainable.

1.3.5   Tissue spectrophotometry.

Harrison et al[26] have used visible light in the range 500 to 620 nm to measure non-invasively the relative haemoglobin concentration and absolute oxygen saturation SO_2 of tissue. A xenon lamp was coupled to 18 fibres to illuminate the skin. Twelve detection fibres were used to couple light re-emitted from the skin to a multi channel photodetector, consisting of a diffraction grating and a 1024 element photodiode array. The instrumentation allowed the intensity of the emitted light to be recorded as a function of wavelength. Several isobestic points (500.9, 527.5, 547.5, 572.6 and 585.4 nm) where the absorption is independent of oxygenation state were identified (see figure 1.2). Measurements made at these isobestic points were combined with measurements made at 558.5 nm, where there is maximum difference between the absorption of oxygenated and deoxygenated haemoglobin (again see figure 1.2), to allow calculation of SO_2 using equations

$$SO_2 = \frac{100 \times (OXI + 0.43)}{1.5} \qquad \textbf{1.2}$$

where OXI is an oxidation index given by equation 1.3

$$OXI = \left[\frac{E_{571.8} - E_{560.1}}{11.7} - \frac{E_{560.1} - E_{548.5}}{11.6} \right] \times \frac{100}{HBI} \qquad \textbf{1.3}$$

where E_λ is the extinction coefficient at a wavelength of λ nm. The extinction coefficient is calculated from the absorption at λ according to equation 1.4:

$$E_\lambda = -\log_{10}(1 - A_\lambda) \qquad \textbf{1.4}$$

and the index of haemoglobin content HBI is given by equation 1.5

$$HBI = 100 \times \left[\frac{E_{527.1} - E_{500}}{27.1} + \frac{E_{548.5} - E_{527.1}}{21.4} + \frac{E_{548.5} - E_{571.8}}{23.3} + \frac{E_{571.8} - E_{585.4}}{13.6} \right] \qquad \textbf{1.5}$$

Since visible light is absorbed more strongly than near infra-red light (see figure 1.2), the

penetration depth is shallower using visible light. Hence this method gives information about the skin oxygenation and not that of deep lying tissue, which is the aim of many researchers using near infra-red. The isobestic points need to be determined accurately, although using 5 points reduces the error caused by inaccuracies in their determination compared to the use of 3 points by Feather et al[27].

1.4   Terminology in tissue optics.

A bewildering number of terms are used by many authors to describe physical quantities relevant to tissue optics. Power (W) is usually called the radiant energy flux, but power density (W/m^2) is known as the radiant energy fluence rate. In the field of astronomy, where the properties of space are in many ways similar to those of tissue, although the magnitudes are very dissimilar, different terms are often used. A review of the optical properties of tissue by Patterson[28] lists a number of these, including most of those used in this thesis. As far as possible, the terms used in this thesis are consistent with those of Patterson, although sometimes symbols have been changed, so as to allow a logical and unique set of symbols across the whole thesis. For convenience, the definitions used in this thesis are summarised in table 1.1.

Table 1.1 Definitions of the optical parameters used in this thesis

Name of quantity	Definition	Symbol	Units
Photon number	Number of photons emitted or detected	N_{ph}	
Radiant energy	Energy of photons ($N_{ph}h\nu$)		J
Radiant energy flux	dN_{ph}/dt		W
Radiant energy fluence	dN_{ph}/dA, where N_{ph} are incident on a sphere of cross-sectional area A.	Ψ	J m^{-2}
Radiant energy fluence rate	where N_{ph} are incident on a sphere of cross-sectional area A per unit time.	ψ	W m^{-2}
Energy radiance	Energy transported by photons in direction Ω per unit solid angle per unit time per unit area	$L(\Omega)$	W m^{-2} sr^{-1}

Absorption coefficient	Probability of photon absorption per infinitesimal pathlength	μ_a	m^{-1}
Scattering coefficient	Probability of photon scatter per infinitesimal pathlength.	μ_s	m^{-1}
Attenuation coefficient	$\mu_a + \mu_s$	μ_t	m^{-1}
Mean free path	$1/\mu_t$		m
Absorption cross section	$$\frac{\mu_a V}{N_{par}}$$ where there are N_{par} particles, with μ_a in a volume V.	σ_a	m^2
Scattering cross section	$$\frac{\mu_s V}{N_{par}}$$ where there are N_{par} particles, with μ_s in a volume V.	σ_s	m^2
Total cross section	$\sigma_a + \sigma_s$	σ_t	m^2
Single scattering albedo	$\mu_s/\mu_t = \mu_s/(\mu_a + \mu_s)$	a	
Mean cosine of scattering phase function		g	
Differential scattering cross-section	$$\lim_{r\to\infty} \frac{r^2 I_s(\vec{e}_i,\vec{e}_s)}{I_i}$$	$\sigma(r,\vec{e}_s,\vec{e}_i)$	m^{-1} sr^{-1}
Scattering phase function	Normalised differential scattering coefficient	$f(\vec{e}_s,\vec{e}_i)$	sr^{-1}
Effective attenuation coefficient	$$\sqrt{3\mu_a(\mu_a + (1-g)\mu_s)}$$	μ_{eff}	m^{-1}

Diffusion coefficient	$$\dfrac{1}{3(\mu_a + (1-g)\mu_s)}$$	D	m
Transport scattering coefficient	$(1-g)\mu_s$	μ_s'	m^{-1}
Specular reflectance	Fraction of incident light flux reflected by the irradiated surface	R_{sp}	
Scattered reflectance	Fraction of incident light flux scattered through the irradiated surface	R_s	
Total reflectance	$R_{sp}+R_s$	R_t	
Internal reflectance	Fraction of internal light flux incident on tissue surface and reflected back into the tissue	R_i	
Primary transmittance	Fraction of incident light flux transmitted after scattering	T_p	
Scattered transmittance	Fraction of incident light transmitted without interaction	T_p	
Total transmittance	T_p+T_s	T_t	

1.4.1   Absorption.

If a beam of light is incident on a single particle, some of the light is absorbed and therefore the particle increases in temperature. If a sample of a purely absorbing homogeneous media of thickness dl is placed in the path of a collimated beam of light between the source and a detector, the fraction of light dI/I absorbed by the sample is found experimentally to be directly proportional to the sample thickness. The constant of proportionality is known as the absorption coefficient μ_a.

$$\frac{dI}{I} = \mu_a dl \qquad\qquad \textbf{1.2}$$

If this is integrated for a sample of thickness l, the intensity leaving the sample I, is given by equation 1.3:

$$I = I_0 \exp(-\mu_a l) \qquad\qquad \textbf{1.3}$$

where I_0 is the incident intensity. Note that equations 1.2 and 1.3 assume all the input light reaches the sample. If there is a refractive index mismatch between the sample (with refractive index $n_1 + jk_1$) and its surroundings with refractive index n_2, a fraction of the incident power is reflected, which depends on the relative refractive indices of the two media. For randomly polarised light the fraction of incident light reflected is then[29].

$$\frac{I_{ref}}{I_0} = \frac{n_2 - n_1 - jk_1}{n_1 + n_2 + jk_1} \qquad\qquad \textbf{1.4}$$

The fraction that is transmitted is:

$$\frac{I_t}{I_0} = \frac{n_2}{n_1 + n_2 + jk_1} \qquad\qquad \textbf{1.5}$$

For a tissue/air boundary, where tissue has a refractive index of approximately 1.4[30], approximately 2.8% of the light is reflected at normal incidence. The absorption coefficient may also be expressed as

$$\mu_a = \frac{N_{par}\sigma_a}{V} \qquad\qquad \textbf{1.6}$$

where there are N_{par} particles in a volume V, each with an absorption cross section of σ_a. If a medium contains two or more absorbers, the absorption coefficient is the linear superposition of each of these:

$$\mu_a = \frac{1}{V}\sum_{i=1}^{N} N_{par}^{(i)}\sigma_a^{(i)}$$ **1.7**

where $N_{par}^{(i)}$ is the number of absorbing particles of type i and $\sigma_a^{(i)}$ is the absorption cross section area of the particles. The absorption coefficient can also be expressed as the imaginary part of a complex refractive index. The absorption coefficient may in general be a function of position, time, wavelength etc, and would then be written as $\mu_a(\vec{r},t,\lambda)$.

Sometimes attenuation is expressed in optical densities (OD), where the attenuation in optical densities is given by:

$$Attenuation\ (OD) = \log_{10}\frac{I_0}{I}$$ **1.8**

The absorption of a material is sometimes expressed in optical densities per metre, indicating the amount of attenuation in optical densities for each metre of sample thickness. Since cuvettes of 1 cm thickness are often used to make attenuation measurements of liquids, absorption is often quoted in OD/cm in the literature. To convert from OD/cm to μ_a in mm^{-1}, the OD/cm data should be multiplied by $\log_e(10)/10 = 0.2303$.

1.4.2   Scattering.

When a plane wave interacts with a single particle, with a refractive index different from its surroundings, some of the light is scattered in a direction different from its original path. In the near infra-red, the light scattering in tissue is elastic scattering - ie there is no energy loss and so the wavelength of the scattered light is unchanged. Some tissue, notably the teeth will fluoresce if shorter (UV) wavelengths are directed towards them, with the light scattered being at a longer wavelength (less energy) than the incident light, but these non-elastic scattering effects are negligible with near infra-red light.

In general, the intensity I_s of the scattered radiation in a direction $\vec{e}_s$ depends on the intensity I_i and direction $\vec{e}_i$ of the incident radiation and the radial distance r from the particle. It is characterised by a differential scattering cross section $\sigma_s(\vec{e}_s,\vec{e}_i)$ defined by:

$$\sigma(\vec{e}_i,\vec{e}_s) = \lim_{r\to\infty}\ \frac{r^2 I_s(\vec{e}_i,\vec{e}_s)}{I_i}$$ **1.9**

The scattering cross section σ_s is defined as the integral over all angles of σ.

$$\sigma_s = \int_0^{4\pi} \sigma(\vec{e}_i, \vec{e}_s) d^2 \vec{e}_s \qquad\qquad \textbf{1.10}$$

The phase function $f(\vec{e}_s, \vec{e}_i)$ is defined as

$$f(\vec{e}_i, \vec{e}_s) = \frac{\sigma(\vec{e}_i, \vec{e}_s)}{\sigma_s} \qquad\qquad \textbf{1.11}$$

and is the probability density that an incident photon is scattered from the direction $\vec{e}_i$ to direction $\vec{e}_s$. The mean cosine of the phase function is known the anisotropy factor, normally given the symbol g.

Given a number N_{par} particles in a volume V, with scattering cross section σ_s the scattering coefficient μ_s may be defined:

$$\mu_s = \frac{N_{par}\sigma_s}{V} \qquad\qquad \textbf{1.12}$$

The scattering coefficient μ_s is the probability of a photon scattering per infinitesimal length. It can be seen from equation 1.12 that the scattering coefficient rises linearly with the concentration of particles. However, there is some evidence to suggest[31,32] that the scattering coefficient may not increase indefinitely with increasing particle concentration, as particles may clump together at high concentrations, altering σ_s resulting in a lower scattering coefficient than equation 1.12 suggests. In addition, the assumption that scattering interactions are independent of each other is no longer valid at high particle concentrations, and interference effects need to be taken into account.

If the total attenuation coefficient μ_t is defined as:

$$\mu_t = \frac{N_{par}\sigma_t}{V} = \frac{N_{par}\sigma_a}{V} + \frac{N_{par}\sigma_s}{V} = \mu_a + \mu_s \qquad\qquad \textbf{1.13}$$

which determines the attenuation that will be measured between a source and a detector due to both absorption and scattering, assuming that none of the scattered radiation becomes scattered back into the view of the detector.

The transport scattering coefficient

$$\mu_s^{/} = (1-g)\mu_s \qquad\qquad \textbf{1.14}$$

represents the scattering coefficient of an isotropically scattering medium that would give rise to the same attenuation of a diffuse photon flux. This is used in place of the true μ_s in a number of analytical models, such as the diffusion equation (see section 1.5.3). Consider a photon confined to a 2D space, leaving a source at a position x=0, then scattering twice. The first scatter is of length l_1 and the second of length l_2. If scattering were truly isotropic, the position of a photon after two scattering events would on average be a distance $\sqrt{(l_1^2 + l_2^2)}$ from the source. However, if the two scattering events were in exactly the same direction, the photon would be the longer distance l_1+l_2 from the source. Hence in a scattering medium, which has a non-isotropic scattering phase function, the effective path length is longer, so the effective scattering coefficient is smaller.

A long way from the source in an infinite medium, the intensity drops as[33] exp(-μ_{eff} x), where μ_{eff} is defined in equation 1.15

$$\mu_{eff} = \sqrt{3\mu_a(\mu_a + (1-g)\mu_s)} \qquad\qquad \textbf{1.15}$$

but it is important to note that this does not hold close to the source.

1.4.3 Temporal point spread function (TPSF).

If a short pulse of collimated light of intensity I_0 is directed into a non-scattering medium of thickness x, the light propagates with a velocity v=c/n, where n is the refractive index of the medium and c is the velocity of light in a vacuum. The pulse will exit a time t=x/v=x.n/c later with an amplitude which will be reduced by the absorption in the medium, to I_0 exp(-μ_a x). This is shown in the top half of figure 1.3.

When light is directed into a scattering medium, some light may go directly through if the medium is not too thick, but most will be scattered and so exit at a time later than the x.n/c, calculated earlier. Hence there will be a distribution of transit times as shown in the lower half of figure 1.3. Photons taking the most direct path (shown as a continuous line) arrive at the detector first. Photons taking a longer path (shown dotted) arrive later. Photons taking the longest path (shown dashed) arrive last. The integrated intensity will be much lower than with a pure absorber, since photons spend more time in the material, travelling longer distances and therefore suffering greater attenuation. Hence even the time integrated output intensity will be less than the I_0 exp(-μ_a l) that would result from a pure absorber and in addition some photons
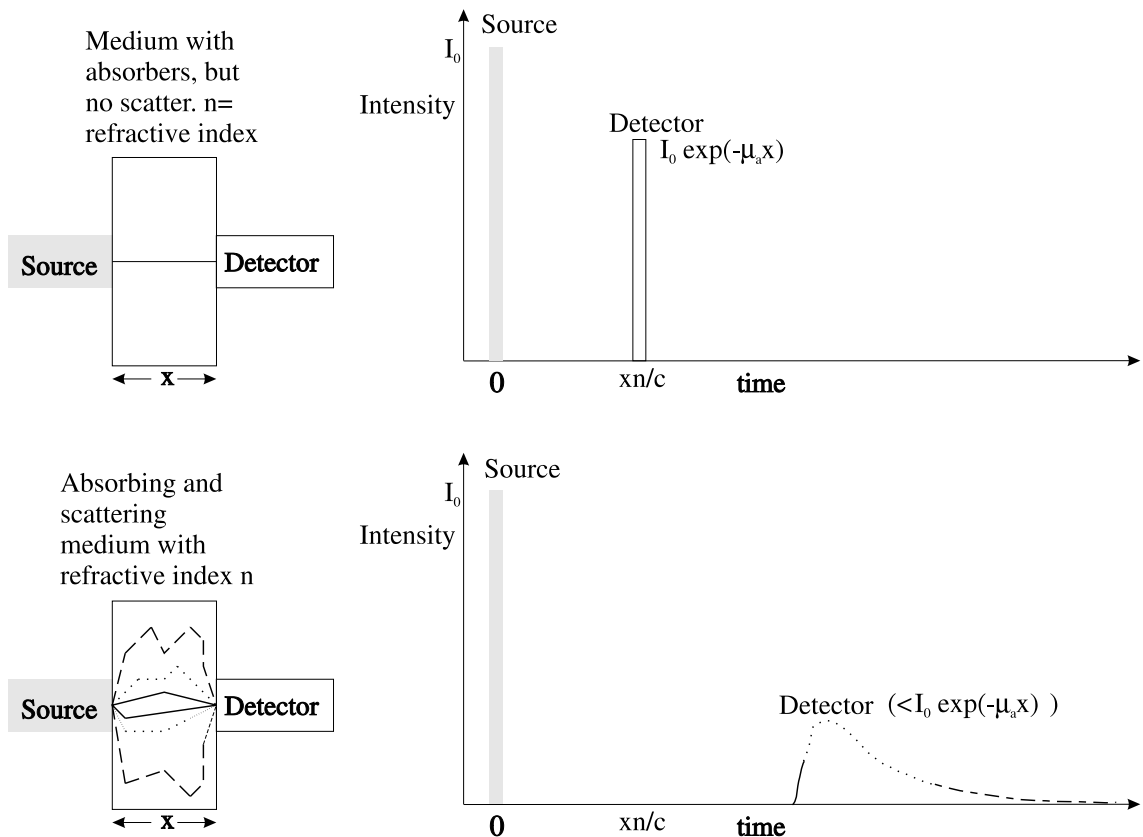
Figure 1.3 Diagram showing the impulse response of a pure absorber (top) and a medium with both absorption and scattering (bottom). Note the scattering causes a distribution of photon transit times.

will be scattered away from the detector.

1.4.4 Typical values of the optical properties of tissue.

A review paper published in 1990 by Cheong et al[34] collects all the published data for the optical properties of mammalian tissue and gives one or more of μ_a, μ_s, μ_t, μ_s', g and μ_{eff} at various wavelengths in the range 193 nm to 2100 nm. Table 1.2 contains a small sample of the data given by Cheong et al and to this has been added *in vivo* data by Mitic et al[35]. In each case the units are mm^{-1}, although Cheong's paper has units of cm^{-1}, so they have been multiplied by 10 for inclusion in table 1.1. Details of the tissue preparation, sample geometry, experimental method as well as the underlying measurement theory and references to the original authors of the data are given in Cheong's paper, but are not given here. Data in the table 1.1 enclosed with parentheses, indicates data not given by Cheong et al or Mitic et al, but which has been calculated from the information in their papers. The parameter μ_{eff} is rarely given for *in vitro* measurements, but it is easily calculated using equation 1.15.

Table 1.2 Optical properties of tissue.

Material	λ	μ_t	μ_a	μ_s	μ_s'	g	μ_{eff}
Aorta	632.8	31.6	0.052	31.6	4.1	0.87	(0.806)
Bladder	632.8	8.94	0.14	8.8	0.352	0.96	(0.455)
Bladder	633	3.07	0.14	2.93	0.264	0.91	(0.412)
Brain (grey matter)	633	5.26	0.158	5.10	0.204	0.96	(0.414)
Brain (white matter)	633	6.28	0.263	6.02	0.722	0.88	(0.882)
Breast (*in vivo*)	800		0.0025		1.03		
Uterus	635	39.4	0.035	39.4	(12.2)	0.69	(1.134)
Whole Blood (HbO$_2$)	665	124.7	0.13	124.6	0.611	0.995	(0.542)
Whole Blood (HbO$_2$)	685	141.6	0.265	141.3	(1.413)	0.99	(1.155)
Whole blood (HbO$_2$)	960	50.8	0.284	50.5	0.384	0.992	(0.766)
Whole blood (HbO$_2$)	665	51.4	0.487	50.9	0.249	0.995	(1.041)
Whole blood (HbO$_2$)	960	67.0	0.168	66.8	0.508	0.992	(0.592)

(units of μ_a, μ_s, μ_s', μ_t and μ_{eff} are mm^{-1}).

1.5   Modelling light transport in tissue.

1.5.1   Why model to obtain μ_a and μ_s.

When we wish to measure the electrical resistance between two points, this is frequently done by applying a fixed current and measuring the voltage drop across the resistance and then determining the resistance directly from the definition of an Ohm.

However it is very difficult to measure the absorption and scattering coefficients of tissue

by such a direct method. It is possible *in vitro* if the tissue is very thin and therefore scattering can be ignored to determine directly the absorption coefficient, using the fact that the output intensity through a sample of thickness x will be related to the input intensity by

$$I_{out} = I_{in} \exp(-\mu_a x) \qquad \qquad \textbf{1.16}$$

but with typical tissue scattering coefficients being of the order of 10 mm^{-1}, photons are scattered on average every 100 μm, so to be fairly certain of no scattering, samples need to be of the order of 10 μm thick. This technique is not suitable for clinical *in vivo* use, although it has been used *in vitro*, but the problems of preparing samples this thin without effecting their optical properties is probably one reason why there is such a wide variation in the optical properties of tissue reported in the literature. If the refractive index of the sample is different from the surrounding medium, a correction for the specular reflection will also be necessary, but this is fairly trivial.

Tissue optical properties are normally estimated by use of a model of light transport in tissue, which is a more practical method, then using the definitions of μ_a or μ_s directly. A number of different models have been developed, each having its advantages and disadvantages. Some of these models will be discussed and the advantages and disadvantages of each considered.

For the purpose of this project, the light source and detector will be fibre optic coupled and be close in proximity to each other. The model used must allow such an arrangement to be modelled.

1.5.2   Kubelka-Munk (KM) model.
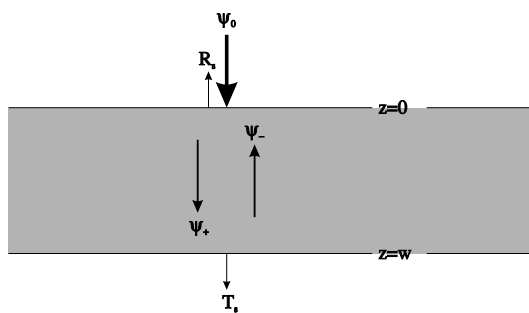
Figure 1.4 Kubelka-Munk model, showing the incident flux ψ_0, forward flux ψ_+ and reverse flux ψ_- in an infinite slab, of finite thickness w.

Kubelka and Munk[36] developed a simple 2 flux theory of light transport in an infinite slab of thickness w, with a source of ψ_0 at z=0, as shown in figure 1.4. A diffuse forward flux ψ_+ propagates in the medium away from the source and a reverse diffuse flux ψ_- propagates towards the source. Although tissue is predominately forward scattering (g=0.9 is typical of tissue), light is clearly scattered away from the normals to the slab faces, but these are

ignored in the simple Kubelka Munk model. Hence the Kubelka Munk model is not an accurate approximation of light transport in tissue, especially near the boundaries. The differential equations describing ψ_+ and ψ_- according to Kubelka Munk theory are:

$$d\psi_+ = -(S_{KM} + A_{KM})\psi_+ dz + S_{KM}\psi_- dz \qquad \textbf{1.17}$$

$$d\psi_- = (S_{KM} + A_{KM})\psi_- dz - S_{KM}\psi_+ dz \qquad \textbf{1.18}$$

where S_{KM} is a modified scattering coefficient and A_{KM} is a modified absorption coefficient. The parameter A_{KM} is a complicated function of the absorption coefficient, but if the radiance is assumed diffuse at all points in the slab, which is clearly not an accurate assumption, then A_{KM} and S_{KM} are the same for both the forward flux and the reverse flux, which allows for simple expressions for A_{KM} and S_{KM}:

$$A_{KM} = 2\mu_a \qquad \textbf{1.19}$$

$$S_{KM} = \mu_s \qquad \textbf{1.20}$$

By using the simple boundary conditions of:

$$\psi_+(z=0) = \psi_0 \qquad \textbf{1.21}$$

$$\psi_-(z=w) = 0 \qquad \textbf{1.22}$$

and making the assumptions mentioned above, relatively simple expressions are obtained for the forward and reverse fluxes:

$$\psi_+(z) = \frac{\psi_0 \; \sinh(b_{KB}S_{KB}z)}{a_{KB} \; \sinh(b_{KB}S_{KB}w) \; + \; b_{KM} \; \cosh(b_{KB}S_{KB}w)} \qquad \textbf{1.23}$$

$$\psi_-(z) = \frac{\psi_0(a_{KM} \; \sinh(b_{KM}S_{KM}z) \; + \; b_{KM} \; \cosh(b_{KM}S_{KM}z))}{a_{KM} \; \sinh(b_{KM}S_{KM}w) \; + \; b_{KM} \; \cosh(b_{KM}S_{KM}w)} \qquad \textbf{1.24}$$

where

$$a_{KM} = \frac{S_{KM} + A_{KM}}{A_{KM}} \qquad \textbf{1.25}$$

and

$$b_{KM} = \sqrt{a_{KM}^2 - 1} \qquad\qquad \textbf{1.26}$$

This makes it possible to obtain expressions for the scattered (or diffuse) reflectance R_s, which is the fraction of the incident light scattered through the irradiated surface, and the scattered (or diffuse) transmittance T_s, which is the fraction of incident light which is transmitted after scattering:

$$R_s = \frac{\psi_-(z=0)}{\psi_0} = \frac{\sinh(b_{KM}S_{KM}w)}{a_{KM}\,\sinh(b_{KM}S_{KM}w)\,+\,b_{KM}\,\cosh(b_{KM}S_{KM}w)} \qquad\qquad \textbf{1.27}$$

$$T_s = \frac{\psi_+(z=w)}{\psi_0} = \frac{b_{KM}}{a_{KM}\,\sinh(b_{KM}S_{KM}w)\,+\,b_{KM}\,\cosh(b_{KM}S_{KM}w)} \qquad\qquad \textbf{1.28}$$

If we substitute equations 1.19, 1.20, 1.25 and 1.26 into 1.27 and 1.28 we have just two unknowns (μ_a and μ_s) which can be obtained by measuring the two parameters R_s and T_s. Hence the absorption and scattering coefficients of the tissue can be obtained, using just a CW light source and a measurement of the diffuse reflectance and the diffuse transmittance. However, the simplicity of this theory must be weighed against the fact that the assumptions of an infinite medium and a diffuse radiation everywhere are incorrect, and in the particular case, where the source and detector are very close together, the detected radiation will not be diffuse at all, although it will be well away from the source and detector fibres.

1.5.3   The diffusion equation approximation of the radiative transfer equation.

One of the most common mathematical methods for describing light transport in tissue is by use of radiative transport theory. The equation of radiative transfer can be derived by considering the radiant energy balance in an elemental volume of tissue, which can be stated as:

$$\begin{vmatrix}\text{change of}\\\text{radiance}\\\text{with}\\\text{time}\end{vmatrix} + \begin{vmatrix}\text{net change}\\\text{due to}\\\text{energy}\\\text{flow}\end{vmatrix} + \begin{vmatrix}\text{radiance loss}\\\text{due to}\\\text{absorption}\\\text{and scattering}\end{vmatrix} - \begin{vmatrix}\text{gain in radiance}\\\text{due to scatter}\\\text{from all}\\\text{other directions}\end{vmatrix} + \begin{vmatrix}\text{gain in}\\\text{radiance}\\\text{due to}\\\text{source}\end{vmatrix} \qquad \textbf{1.29}$$

This can be expressed mathematically as:

$$\frac{n}{c}\frac{\partial I}{\partial t}+\vec{e}_s.\nabla I(\vec{r},t,\vec{e}_s)+\mu_t(\vec{r})I(\vec{r},t,\vec{e}_s)=\int_0^{4\pi}\mu_s(\vec{r})f(\vec{e}_i,\vec{e}_s)I(\vec{r}_s,t,\vec{e}_i)d^2\vec{e}_i+s(\vec{r},t,\vec{e}_s)$$ **1.30**

This integro-differential equation can be expanded as an infinite set of coupled differential equations. A truncated subset of these equations, known as the P_N approximation gives $(N+1)^2$ coupled differential equations which can then be used to approximate the radiative transfer equation. If N=1, there are four equations and these can then be reduced to what is known as the diffusion equation or the P_1 model. The time independent diffusion equation is:

$$\nabla^2\psi(\vec{r})-\frac{\mu_1}{D}=\frac{-S_0(\vec{r})}{D}+3\nabla.S_1(\vec{r})$$ **1.31**

where $\psi(r)$ is the energy fluence rate (W m^{-2}), D is the diffusion coefficient and is given by:

$$D=\frac{1}{3\mu_{tr}}=\frac{1}{3(\mu_a+(1-g)\mu_s)}$$ **1.32**

and $S_0(r)$ and $S_1(r)$ are the first two coefficients in the Legendre expansion of the source $s(r,\Omega)$.

The time-independent diffusion equation may be extended to a time-dependent equation, which allows for the calculation of the $\psi(r,t)$.

$$\frac{n}{c}\frac{\partial}{\partial t}\psi(r,t)-D\nabla^2\psi(r,t)+\mu_a\psi(r,t)=s(r,t)$$ **1.33**

where $s(r,t)$ is the source and n is the refractive index, which is usually taken as 1.4 as this value is typical of tissue.[30]

For the case when the source $s(\vec{r},t)$ is a delta function at t=0 ($\delta(0,0)$) and the geometry is an infinite medium, the solution to this time-dependent diffusion equation is[37]:

$$\psi(r,t)=\frac{c}{n}(4\pi Dtc/n)^{-3/2}\exp(-\frac{nr^2}{4Dct}-\mu_a tc/n)$$ **1.34**

Figure 1.5 shows some graphs of equation 1.34 at various source-detector spacings.

If the input is not a pulse of light, but a modulated light source, modulated at a frequency ω, the frequency-dependent intensity at a distance r is given by equation 1.35[38].

$$I(r,\omega)=\frac{\exp(-j\omega t)(1+\alpha d)(\exp(-\alpha r)}{2(2\pi)^{3/2}r^2}$$ **1.35**

where α is defined in equation 1.36.

$$\alpha = \sqrt{\frac{3(\mu_a c + j\omega)(\mu_a + \mu_s')}{c}} \qquad \mathbf{1.36}$$

If we substitute for $\omega = 0$, α simplifies to μ_{eff} (see equation 1.15) and we get the integrated intensity, for an infinte medium, as would be measured on a photodiode.

$$I(r) = \frac{(1 + \mu_{eff} r)\exp(-\mu_{eff} r)}{(2\pi)^{3/2} r^2} \qquad \mathbf{1.37}$$

The same result can be obtained by integrating equation 1.34 between t=0 and t=∞.
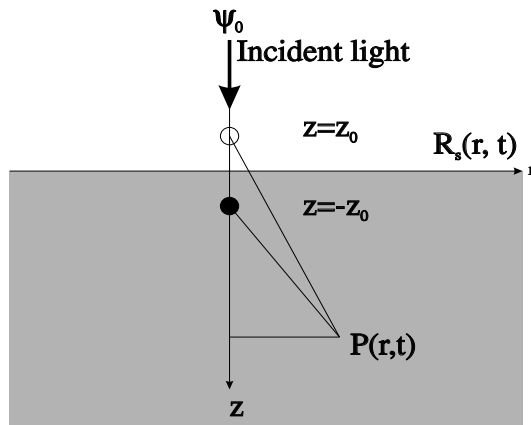
Figure 1.6 Geometry for calculation of the diffuse reflectance $R_s(r,t)$ in a slab infinitely wide and infinitely thick.
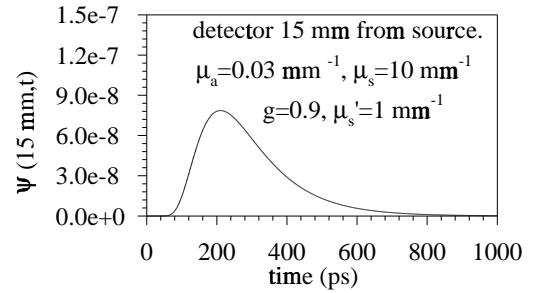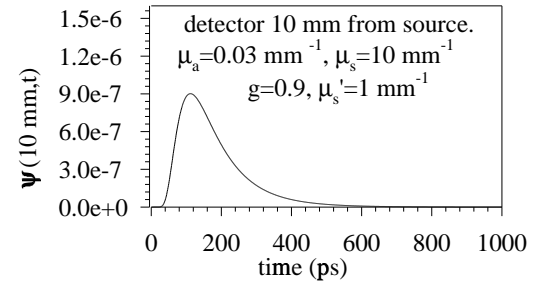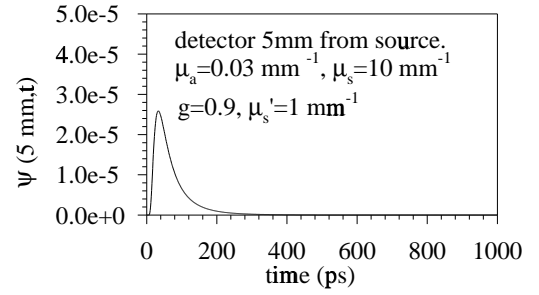
Figure 1.5 Graphs showing the form of the diffusion equation, for an infinite medium.

Patterson et al[39] used equation 1.34 to find an analytical expression for the pulse shape in an infinite homogeneous slab (figure 1.6). They assumed the incident photons were initially scattered at a depth z_0 given by equation 1.38:

$$z_0 = \left[(1-g)\mu_s\right]^{-1} \qquad \mathbf{1.38}$$

then found an expression for the diffuse reflectance $R_s(r,t)$ which would be measured at a distance of r away from the source on the surface of the infinite slab:

$$R_s(r,t) = (4\pi Dc/n)^{-3/2} z_0 t^{-5/2} \exp(-\mu_a ct/n) \exp\left(-\frac{n(r^2 + z_0^2)}{4Dct}\right) \qquad \mathbf{1.39}$$

In order to satisfy the boundary conditions, an infinite number of photon sources and an infinite number of photon sinks were necessary. Only one of each are shown in figure 1.6, with the photon source shown as a filled back dot and the photon sink shown as an empty circle. For the case when $r^2 \gg z_0^2$, this gives the slightly simpler expression

$$R_s(r,t) = \frac{c}{n}\left(\frac{4\pi Dc}{n}\right)^{-3/2} z_0 t^{-5/2} \exp(-\mu_a tc/n)\ \exp\left(-\frac{nr^2}{4Dtc}\right) \qquad \textbf{1.40}$$

The differential of the log of this with respect to time is:
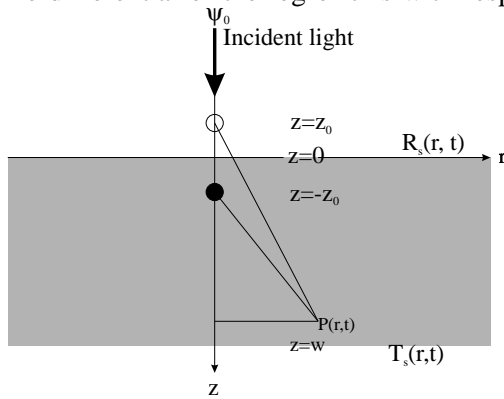
$$\frac{d}{dt}\log_e R(r,t) = -\frac{5}{2t} - \mu_a c/n - \frac{nr^2}{4Dct^2} \quad \textbf{1.41}$$

Figure 1.7 Geometry for calculation of the time-resolved reflectance and transmittance of an infinite slab of finite thickness w.

At large values of t, the absorption coefficient could easily (at least in theory) be found from the log slope of the TPSF, as

$$\lim_{t\to\infty} \frac{d}{dt}\log_e R_s(r,t) = -\mu_a c/n \qquad \textbf{1.42}$$

In a more practical sense, at large t, the amount of available light is small (see figure 1.5), making the measurements very noisy, and so of limited accuracy. The transport scattering coefficient could also be determined from the temporal profile of $\log_e R_s(r,t)$, by noting that at the time that the TPSF reaches a peak t_{peak}, the slope is zero and so by solving equation 1.41, we can obtain an expression for μ_s':
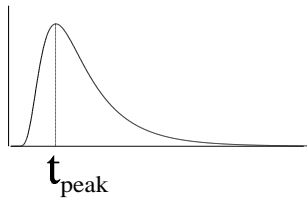
$$\mu_s' = \frac{1}{r^2}(4\mu_a c^2 t_{peak}^2/n^2 + 10ct_{peak}/n) - \mu_a \qquad \textbf{1.43}$$

Figure 1.8 Definition of t_{peak} - the temporal position of the TPSF peak.

Patterson et al also derived expressions for the time-resolved reflectance $R_s(r,w,t)$ and the time resolved transmittance $T_s(r,w,t)$ of a slab of finite thickness w, as shown in figure 1.7. The expression for the diffuse reflectance $R_s(r,w,t)$ is given by equation 1.44

$$R_s(\vec{r},w,t)=(4\pi Dc/n)^{-3/2}t^{-5/2}\exp(-\frac{\mu_a ct}{n})\exp(-\frac{r^2 n}{4Dct})$$

$$\times\left[z_0\exp(-\frac{z_0^2 n}{4Dct})-(2w-z_0)\exp(-\frac{n(2w-z_0)^2}{4Dct})+(2w+z_0)\exp(-\frac{n(2w+z_0)^2}{4Dct})\right] \qquad \textbf{1.44}$$

and for the transmittance $T_s(r,w,t)$ is given by equation 1.45:

$$T_s(\vec{r},w,t)=(4\pi Dc/n)^{-3/2}t^{-5/2}\exp(-\frac{\mu_a ct}{n})\exp(-\frac{r^2 n}{4Dct})$$

$$\times\left[(w-z_0)\exp(-\frac{(w-z_0)^2 n}{4Dct})-(w+z_0)\exp(-\frac{n(w+z_0)^2}{4Dct})\right. \qquad \textbf{1.45}$$

$$\left.+(3w-z_0)\exp(-\frac{n(3w-z_0)^2}{4Dct})-(3w+z_0)\exp(-\frac{n(3w+z_0)^2}{4Dct})\right]$$

Note that equations 1.39, 1.40, 1.44 and 1.45 all give positive outputs for any t>0, when in fact for t< rn/c, no photons can have ever arrived. As such, the early part of these predicted TPSF are unphysical and so must be ignored. This is a major limitation of the diffusion equation.

Arridge et al[38] have derived solutions to the diffusion equation for time, frequency, phase and the mean arrival time of photons for a number of cases including the infinite medium, infinite slab (figure 1.6), semi-infinite half-space (figure 1.7), 2D circle, finite and infinite cylinders and a sphere. The mean time of the TPSF <t> for an infinite medium was given by:

$$<t>=\frac{d^2}{2(\gamma^2+d(\mu_a c/n)^{1/2}\gamma)} \qquad \textbf{1.46}$$

where $\gamma^2=vD$. The expression for mean time <t> of the infinite half-space (figure 1.6) is identical to that of an infinite medium if the source-detector spacing d is replaced by the radial distance between source and detector r.

1.5.3.1 Solution of the diffusion equation by finite-element modelling (FEM).

The diffusion equation has been shown to be capable of predicting both the time-independent and time-dependent behaviour of light in tissue, on the condition that $\mu_s \gg \mu_a$. However, even for simple geometries, such as an infinite cylinder or sphere, analytical approximations are difficult to derive. For more challenging geometries, such as are required to solve questions about light transport in the head, analytical solutions have never been derived, and probably never will be. The finite-element modelling (FEM) technique has however been used to solve the problem of extending the diffusion equation to more complex geometries than

is possible by purely analytical means.

Finite-elements has been widely used in a number of fields of engineering fields since the 1950's[40], although its use in tissue optics is much more recent. Livesley[40], gives an introduction to the finite-element method and discusses applications in both electrical, mechanical and civil engineering, whereas Silvester et al[41] concentrates on electrical engineering. In a book edited by Itoh on the use of numerical techniques for microwave engineering, a chapter by Davies[42] discusses how finite-elements, and a related method finite-difference, are used for solving electromagnetic field problems. Kirkby[43] has shown how a small (66-line) computer programme may be used to solve Laplaces equation[44] numerically using the technique of finite-difference. This programme was able to find the capacitance per metre, inductance per metre, velocity of propagation and impedance of a boxed microstrip. Director et al[45] has used the finite-difference technique to simulate the thermal behaviour of tissue and Yao et al[46] have used the technique to determine the optical properties of tissue from measurements with a frequency domain spectrometer. It is in general much easier to use the finite-element method rather than the finite-difference method when the boundary is irregular, as the finite-element model can use arbitrary shaped triangles (2D model) or tetrahedrals (3D model) to approximate closely any physical boundary. It is much more difficult to achieve this flexibility using the finite-difference method.

The finite-element method is now briefly discussed. It is first assumed that we have some known excitation or source s, and an unknown and wanted field ψ, with ψ and s being related by some operator L, which may involve differentiation, integration etc.

$$L\psi = s \qquad \qquad \textbf{1.47}$$

The unknown ψ is then approximated by

$$\psi = \sum_{i=1}^{N} \psi_i b_i(\vec{r}) \qquad \qquad \textbf{1.48}$$

where the $b_i(\vec{r})$ are some known *basis functions*, typically delta functions or polynomials. The problem is then to choose the coefficients $\psi_1, \psi_2 \ \psi_3 \ .. \ \psi_N$ which best approximates the true ψ. An error residual $R(\vec{r})$ is defined as:

$$R(\vec{r}) = L\psi - s(\vec{r}) = L\sum_{i-1} N\psi_i b_i(\vec{r}) - s(\vec{r}) \qquad\qquad \textbf{1.49}$$

which is zero only when we have a precise solution, which is usually an impossible dream for finite N.

A second set of functions, known as *weight functions* are then chosen and the coefficients ψ_i adjusted so the error residual is orthogonal to each of the weight functions w_i. If the scalar product of x and y is written as $<x,y>$, we obtain the following N equations with N unknown ψ's.

$$\left\langle L\sum_{i-1}^{N} \psi_i b_i(\vec{r}) - s(\vec{r}), w_j(\vec{r})\right\rangle \qquad \text{for } j=1,2,3...N \qquad\qquad \textbf{1.50}$$

which can be written as a matrix equation

$$[L][\psi] = [s] \qquad\qquad \textbf{1.51}$$

where   [ψ] denotes the unknown column vector containing the elements ψ_1, ψ_2 ... ψ_N

[**s**] denotes the known column vector with elements $\langle s(\vec{r}), w_j(\vec{r})\rangle$

[**L**] is the known N x N matrix with $(i,j)^{th}$ element $\langle L\, b_i(\vec{r}), w_j(\vec{r})\rangle$

This can be solved using normal matrix methods. In the Galerkin method, which is a common method on which to base finite-elements, the basis and weight functions are identical.

The large and possibly complex shaped domain Ω is then broken into a number of smaller elements - usually triangles for 2D problems and tetrahedrals for 3D problems. For the purpose of discussion, it is assumed the domain is 2 dimensional and the elements are triangles. Curved boundaries are usually approximated by the straight edges of the triangles. The entire domain Ω is subdivided into N_e non-overlapping elements, with no gaps between the elements. Each triangle has a number (N_n) of nodes - typically just 3 (one at each corner), but 6, 10 or more may be used. ψ is then approximated over <u>each</u> individual element as in equation 1.48.

Figure 1.9 shows 3 elements in a typical FEM mesh, with each element having 10 nodes. The total number of unique nodes (N_N) is less than the product $N_e N_n$ (22 instead of 30 in figure 1.9), since many nodes are common to two or more triangles. With a typical mesh having many more in the middle
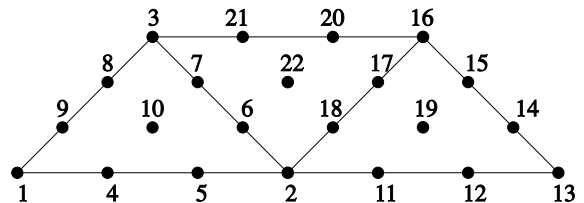
Figure 1.9 3 elements in a typical finite element mesh. Here each element has 10 nodes, although there are 22 nodes in total - not 30, as some are shared between elements.

than at the edges, the number of shared nodes increases. A global matrix **G** is then constructed to allow all the ψ's to be evaluated, one for each of the N_N nodes.

The global matrix **G** has mostly zero elements, since matrix element corresponding to nodes that do not lie in any one element are zero. Referring to figure 1.9, we see elements $G_{1,11}=G_{1,12}=G_{1,13}=G_{1,14}=G_{1,15}=0$. Elements $G_{1,1},G_{1,2},G_{1,3},G_{1,4},G_{1,5},G_{1,6}$ etc are non-zero. Since node 2 is common to all three elements, $G_{2,1}$, $G_{2,2}$, $G_{2,3}$ ... $G_{2,22}$ are all non-zero. This sparsity of the matrices is crucial to the finite-element method, as it allows sparse matrix solving algorithms to be used, which are much faster than those used for dense matrices and use less memory, since there is no need to store the zeros.

Arridge et al[47] described a finite-element model for light transport in tissue by solving the time-dependent diffusion equation. The aim was to predict the photon flux given the tissue parameters - ie a forward model. Using a finite-element solution to the diffusion equation, Schweiger et al[48] tackled the more difficult inverse problem of determining the optical properties of the tissue. Forward data was generated by the FEM model, then an iterative method was used to reconstruct the images. It was found that the mean time of flight data was best for reconstructing the scattering coefficient, whereas intensity data was better for absorption coefficient. Schweiger et al[49] modified the basic FEM code to be able to use different types of boundary conditions, to more closely represent the experimental conditions being modelled. This model is covered in greater detail in Dr. Schweigers Ph.D. thesis[50]. These FEM models used 3 nodes per element ($N_n=3$) and delta functions for both the base and weighting functions, which assumes the flux ψ varies linearly over individual elements. Using more nodes per element, perhaps 6 or 10 by using 2nd or 3rd degree polynomials as basis functions would increase the fitting accuracy over an element, but for a constant computation time, larger elements would be necessary.

1.5.4   Monte Carlo models of light transport in tissue.

Monte Carlo models are used to simulate the passage of individual photons through the tissue. Photons are considered as ballistic particles, which start from the source and then diffuse away. In its simplest form, photons are launched into the tissue, one by one. Each photon then travels a random length before being scattered. On average this length will be $1/\mu_s$, but the exact length is determined by use of a random number. It can be shown that the probability density function of the path length is given by[51],

$$P(l) = \mu_s \exp(-\mu_s l) \qquad\qquad \textbf{1.52}$$

so the probability of the photon being scattered before travelling a length l is

$$p(l) = \int_0^l P(l)\,dl = 1 - \exp(-\mu_s l) \qquad\qquad \textbf{1.53}$$

where p(l) obviously takes on values in the range $0 \le p(l) \le 1$. Hence we can find a number of random scattering lengths, by setting p(l) to a random number between 0 and 1. Therefore, if we have a random number N_{rand} uniformly distributed between 0 and 1, which represents the probability, then

$$N_{rand} = 1 - \exp(-\mu_s l) \qquad\qquad \textbf{1.54}$$

so the length of the photon travels before being scattered l_s will be:

$$l_s = \frac{-\ln(N_{rand})}{\mu_s} \qquad\qquad \textbf{1.55}$$

This gives a distribution of lengths, which are identical to the distribution that should occur in practice.

After being scattered, the trajectory of the photon is changed, again using a random number, but this time relating this to the physically expected distribution of scattering angles. The scattering distribution can be predicted from Mie theory[52], Henyey-Greenstein[53], or better still by using experimentally measured data on the scattering phase function, such as that given by van der Zee[54]. This process is continued until the photon exits, or becomes so weak that its intensity will be negligible, after which the whole process is repeated for more photons until such time as the statistics of the exiting photons are sufficient, or one decides to abandon the problem until a faster computer is available!

If there are two or more regions with different optical refractive indices, the change in the photons direction as it crosses the boundary must be taken into account. When the photon exits, the total path length travelled in each region (if more than one) is recorded, along with position of exit, direction of exit, and any other information that is necessary for the study. From the total path length travelled in each region and the absorption coefficient in each region, it is possible to determine the output intensity due to one photon, by use of the following equation

$$I = I_0 \exp \sum_{i=1}^{N} -\mu_a^{(i)} l_i \qquad\qquad \textbf{1.56}$$

where there are N regions of absorption coefficients $\mu_a^{(1)}$, $\mu_a^{(2)}$ ... $\mu_a^{(N)}$ and the path lengths in the regions are l_1, l_2 ... l_N. Therefore, it is not necessary to repeat calculations if only the absorption coefficient is changed, as long as the path lengths in the regions are known. However, any change of scattering coefficient or geometry requires a separate simulation.

The temporal profile of the exiting light is simple to calculate, since the time a photon would take to arrive at a position is simply given by

$$t = \frac{1}{c} \sum_{i=1}^{N} n_i l_i \qquad\qquad \textbf{1.57}$$

Since research often leads one to ask questions that were originally not considered, it is generally best to save as much information about the photon paths as possible, even if no immediate need for such data is known, since it requires very little extra time to record the information and disk space is relatively cheap nowadays. Once the information is collected, selecting only relevant data from all that is collected is a simple task. However, if one vital piece of information, that is later found to be necessary, was not recorded, it becomes necessary to repeat the simulations which may take months or even years of CPU time.

1.5.5 Comparison of modelling methods and why Monte Carlo was selected.

The Kubelka Munk theory has been shown to be fundamentally flawed and so while it is simple, it can not be expected to be very accurate.

The diffusion equation is an excellent predictor of light propagation in tissue, but analytical solutions, except for the simplest of geometries, are impossible to obtain. However, when coupled to a FEM model to obtain numerical values, it is very powerful. However, it can not be expected to be valid until the light has become diffuse, which is at a distance of approximately $1/\mu_s'$.

Although Monte Carlo modelling is a very slow technique, it really is the only acceptable method when data is required close to the source and detector, where other models just break down. In addition, the Monte Carlo method can deal with arbitrarily complex measurement geometries and can even take account of the detector geometry and spatial sensitivity characteristics. Hence for this project it was decided a Monte Carlo model would be

developed to accurately model the expected TPSF. By fitting the measured TPSF's against the predicted TPSF using the accurate Monte Carlo model, it should be possible to determine the optical properties of the material measured.

For this project to work it is obviously essential, that it is possible to estimate μ_s' from a measurement of the light distribution. This is known as an *inverse problem*, as opposed to a *forward problem* of calculating the light distribution, given knowledge of the tissue optical properties. Bohren and Huffman compare the *inverse problem* to one of estimating the size and shape of a dragon, from its footsteps, which is difficult. The *forward problem*, of estimating the size and shape of a dragon's footprints, given that you know what the dragon looks like, is easier. There are numerous papers which solve the *forward problem* of calculating light distributions, given the optical properties. A few authors have attempted the inverse problem of the sort described here and two specific examples are worth further discussion here.

The first by Farrell et al[55] starts from an analytical expression for the diffuse reflectance R_s at a radius r from a source incident on a semi infinite medium (see figure 1.7), which was a complicated non-linear equation of the form:

$$R(r)=f(r,\mu_{eff},\mu_t^/)$$ **1.58**

The authors developed a method, that whilst not exactly solving the problem that needs to be tackled here, gives a hint of how the problem may be solved. Equation 1.58, like equation 1.39 is derived from the diffusion equation, but whereas equation 1.39 assumes a short pulse of light, equation 1.58, assumes a constant level of illumination, and as such has no time-dependence. However, instead of using intensity as a function of time, 800 sets of data of reflected intensity $R_s(r)$ as a function of radius r for various tissue optical parameters typical of mammalian tissue were generated using equation 1.58. Random noise was then added, to simulate typical *in vivo* measurements. The noisy data was then fitted to equation 1.58, using a non-linear least squares fitting algorithm to estimate the tissue optical parameters μ_{eff} and μ_t and hence μ_a and μ_s'. The RMS error on the estimate of the accuracy of the optical properties was 10%, with the algorithm taking 4 s per calculation. A neural network algorithm was then written, to perform the same task of estimating the tissue parameters. The neural network was trained with 800 sets of noisy data, which took two hours. After training, the neural network was some 400 times faster than the non-linear least squares fit, estimating μ_a and μ_s' in 10 ms, with an RMS error of 7% on the optical properties. It should be noted that the computer used by the authors was said to be a MicroVAX 1170, although no such computer exists! However, a modern PC is considerably

more powerful than any of the MicroVAX's, so today it will probably be unnecessary to use a neural network purely for speed of computation, although the increase of accuracy may warrant it. However, although Farrell et al do state that although the RMS error is smaller with the neural network, it is not significantly smaller, so even this benefit is debatable.

With the complex geometries involved in this project, it is unlikely that an analytical expression of the form in equation 1.58 will ever be found, but it would be easy to generate several hundred data sets for training a neural network, using the Monte Carlo computer program. Once this is done, the optical properties of tissue could be estimated with the aid of the neural network. It should be noted that the method used by Farrell to derive the optical properties is a simple measurement of continuous reflectance versus distance, not time-dependent data as would be generated with this instrument. However, there is every reason to believe the same principles could be applied. Instead of supplying $R_s(r)$ as inputs to the neural network, we would supply $\psi(r,t)$. As such, it is believed that assuming the time dependent data can be measured, estimating the optical properties of the tissue from the data should not be too difficult. However, fitting the full TPSF can be expected to take longer than fitting just the reflectance.

Marquet et al[56] and Bevilacqua et al[57] have both described a method of estimating μ_a
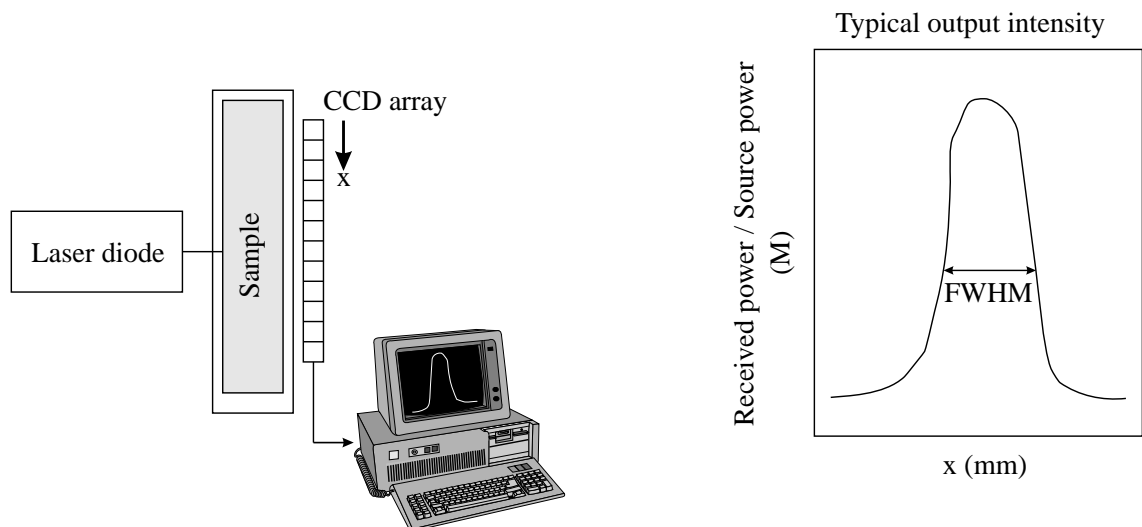
Figure 1.10 System developed for estimating μ_a and μ_s' by measuring two parameters M and FWHM.

and μ_s' from measurements made with the hardware shown in figure 1.10. This uses a sample of tissue in a cuvette with a laser diode (λ=670 nm) illuminating the tissue. The transmitted light intensity was detected on a CCD camera and the maximum intensity at the peak M and the full width at half-maximum of the intensity (in mm) was recorded with a CCD camera. The expected maximum intensity and full width at half maximum were then estimated using a Monte Carlo

model for various tissue optical properties, assuming a geometry identical to the measurement system. By comparing the measured data to the Monte Carlo data, they could estimate both μ_a and μ_s' directly. This was said to have an accuracy of 5% for μ_a and 20% μ_s', although for some tissue types, especially brain white matter, the technique fails to work satisfactorily.

The basic method of comparing observed data to Monte Carlo simulations could in principle be extended to this work, although the form of the data (peak signal and full width at half maximum) will of course be different. Perhaps the full width at half max, mean time, or skew of the TPSF could be compared.

1.6 Hardware for finding μ_a and μ_s without measuring the TPSF.

Before discussing how the TPSF can be measured then analysed to obtain μ_a and μ_s, it should be realised there are other ways to obtain the tissue optical parameters. One method, that of a CW based CCD system, was described in the last section, but others will now be considered. It is especially important to consider these alternatives, as time-domain methods are probably the more complex, and therefore may not be justified.

1.6.1 Simple CW hardware.

Patterson et al[58] have proposed calculating the effective attenuation coefficient of tissue by a straightforward measurement of the spatial variation of the reflectance intensity at the surface of the tissue. They show that as long as the source and detector and sufficiently well spaced for the diffusion approximation to be valid, the following approximation is valid for a semi-infinite half-space

$$R(r) = \frac{\exp(-\mu_{eff})}{r^2} = \frac{\exp(-\mu_a(\mu_a+(1-g)\mu_s)}{r^2} = \frac{\exp(-\mu_a(\mu_a+\mu_s'))}{r^2} \qquad \textbf{1.59}$$

where r is the source-detector spacing. Matcher et al[59] described a spatially resolved spectrometer (SRS) using this principle. This had three laser diode sources (745, 770 and 820 nm) and four physically separate photodiode detectors with integral charge amplifiers as shown in figure 1.11. Unfortunately, this has three main shortcomings. Firstly, it is necessary to know one of μ_a or μ_s' in order that the other may be calculated. However, since the main aim here is to measure changes in μ_s' when μ_a will remain relatively fixed, this would not be a major limitation. Secondly it is unsuitable for fibre coupling and finally although the method works well in homogeneous media, a considerable error is observed in layered media, the reasons for
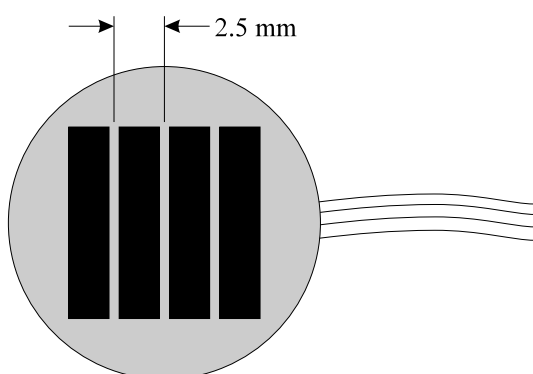
Figure 1.11 Detector of the spatially resolved spectrometer (SRS) described by Matcher et al[59] and based on a principle developed by Patterson et al.

which are not currently known.

1.6.2 Frequency domain phase fluorometer.

In 1984, Berndt[60] described a system (see figure 1.12) for measuring the fluorescence lifetime of dyes, by measurement of the phase shift of the detected light with respect to that of the light source illuminating the dye. Although the purpose for which this hardware was developed was very different from ours, it will be described in detail for two reasons. Firstly, other workers[61,62,63,64] using different hardware, for measuring the same phase shift, have developed methods for obtaining μ_a and μ_s from these measurements and secondly, later developments of this system are particularly relevant to the work described in this thesis.
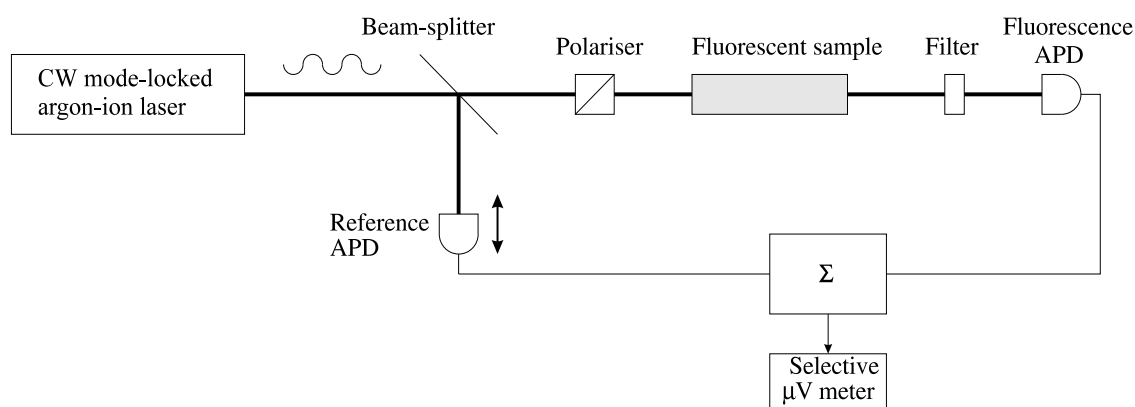
Figure 1.12 Principle of the phase fluorometer described by Berndt in 1984[60] using a sinusoidally modulated laser. Thick lines are optical, thin lines are electrical connections.

The dye was illuminated with a mode-locked argon-ion laser which was run at a reduced current to lower the magnetic field strength in the plasma so as to produce two stable modes oscillating in the laser, separated in frequency by an integer multiple of the fundamental mode spacing of 125 MHz[*1]. The effect of these two modes was to give at least 98% modulation of the laser output at a frequency of either 500 MHz, 625 MHz, 750 MHz or 875 MHz - selectable by altering the current in the laser. By using two avalanche photodiode (APD) detectors - one

---

[*1] The oscillating modes in lasers are discussed more fully in chapter 2

(the fluorescence APD) detecting the fluorescence decay and another (the reference APD) detecting a sample of the laser output, Berndt was able to estimate the fluorescence lifetime of the dye, by measuring the phase shift θ between the two signals, using a μV meter. When the two signals are 180° out of phase, the detector indicated a null. Assuming a single exponential behaviour of the fluorescent dyes, the lifetime τ_F was calculated using

$$\tau_F = \frac{\tan \theta}{2\pi f} \qquad \textbf{1.60}$$

where f was the modulation frequency used (500 MHz to 875 MHz). The optimum frequency to use f_{opt} was related to the fluorescence lifetime:

$$f_{opt} = \frac{2\pi}{\tau_F} \qquad \textbf{1.61}$$

If the shape of the tissue TPSF were a single exponential decay, it could in principle be measured in this way, since the dye lifetimes are not too different from the expected durations of the TPSF in tissue. However, under some circumstances the shape of the TPSF's can be predicted analytically (eg equation 1.39), and this shows that it will not be a single exponentially decaying signal. Berndt later described how both time-constants in a double exponential decay
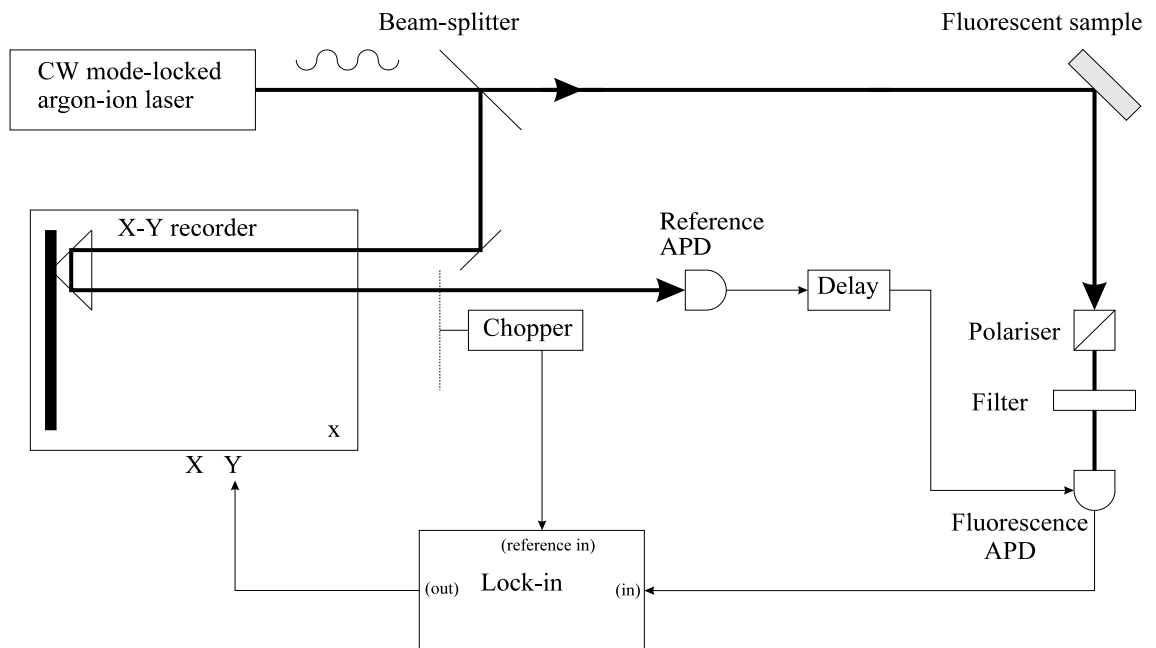
Figure 1.13 System developed by Berndt for the measurement of fluorescence lifetimes, using sinusoidal modulation of the APDs bias voltage and hence gain-modulation.

could be determined using the instrument, but again this does not help with our application. We also show in section 1.6.3 how both μ_a and μ_s can be determined in an infinite homogenous volume, from the phase shift of the signal passing through tissue, using a phase-resolved optical spectrometer.

In 1985 Berndt[65,66] described another method of measuring the phase shift, using gain-modulation of the avalanche photodiode detector. Although the theory of avalanche photodiodes is covered fully in chapter 3, for now it may be assumed that an APD is an optical detector which produces an output current in direct proportion to the incident optical power on its active surface. The constant of proportionality depends on a few physical constants and the internal gain. The gain of an APD depends on the bias voltage as is shown for a typical APD in figure 1.14. In the time-resolved instrument described by Berndt, the principle of which is shown in figure 1.13, the sinusoidal signal from the reference APD is used to modulate the gain of the fluorescence APD. Instead of the signal from the reference APD being added to the signal from the fluorescence APD, the sinusoidal signal from the reference APD is used to gain-modulate the fluorescence APD.
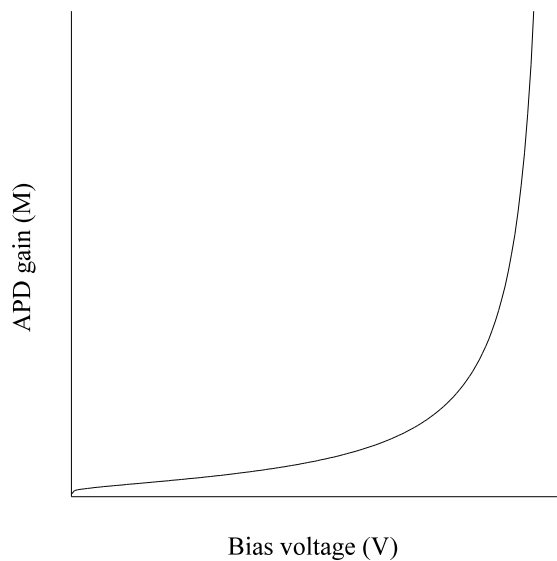
Figure 1.14 Diagram showing the typical relationship between bias voltage and gain for an APD.

The variable time delay mechanism used was ingenious, although it would certainly not be suitable for a portable application. A corner cube reflector was mounted on the arm of an XY recorder. As stepper motors in the XY recorder moved the arm, so the position of the corner cube reflector was moved and so the path length was altered while the lock-in output was recorded.

1.6.3   Frequency domain phase measurements for tissue optics.

Recently a number of phase resolved optical spectrometers have been developed, with the aim of obtaining μ_a and μ_s' from measurements of phase shift and modulation depth of the detected light when the source is a sinusoidally modulated light source of the form:

$$I = I_0(1 + m\,Sin(\omega t)) \qquad\qquad \textbf{1.62}$$

where I is the instantaneous detected intensity, I_0 is the peak intensity, m the modulation depth and ω is the modulation frequency. A number of parameters can be measured, including the phase shift through the sample, the loss of modulation depth etc, as shown in figure 1.15.
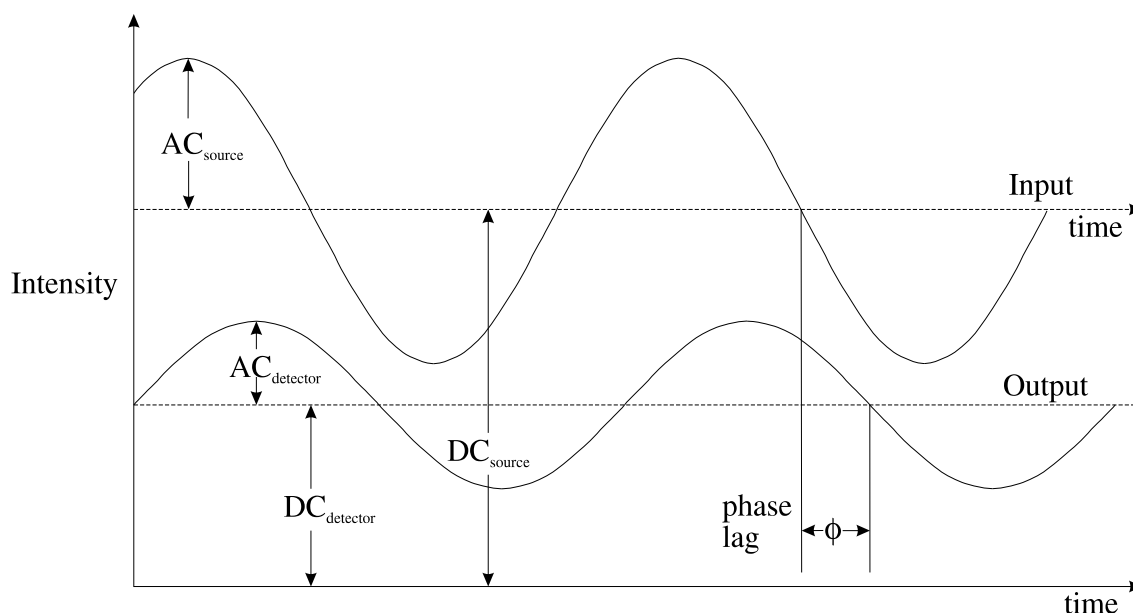
Figure 1.15 Graphs showing the reduced amplitude and modulation depth of an optical signal after passing through a scattering medium.

A system for obtaining these measurements was first described by Chance et al[67]. This used a radio transmitter designed for the American 220 MHz amateur radio band to modulate a laser diode at a fixed frequency and a photomultiplier tube (PMT) gain-modulated at a slightly offset frequency as the detector. A PMT[68,69,70,71] is an optical detector that has an internal gain mechanism, that means one incident photon produces more than one electron in the external circuit. Duncan et al[61] described a similar, but wide-band optical spectrometer whose frequency could be varied over the range 1 to 500 MHz. A simplified version is shown in figure 1.16. Four laser diodes are each capable of being modulated up to 500 MHz, by a low phase noise oscillator. A typical measurement would be made at a modulation frequency of 200 MHz, so the following explanation uses this as an example. The optical output is passed through the sample and is detected on a PMT. A second oscillator, 10 kHz higher than the first, at 200.01 MHz, is mixed with the output of the 200.00 MHz oscillator, then low-pass filtered, to enable a 10 kHz reference frequency to be obtained. This is used to drive the reference input of a lock-in amplifier. The current from the PMT is passed into the RF+DC port of a bias-tee. The signal that
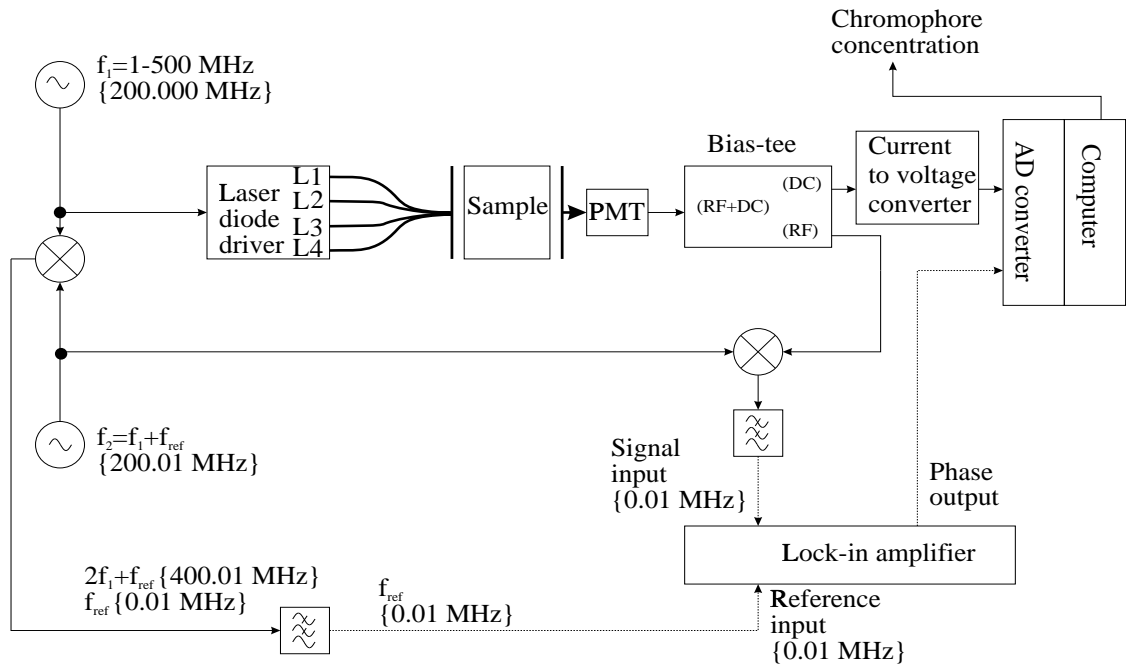
Figure 1.16 Optical spectrometer especially built for evaluation of tissue properties. RF connections are shown solid, low frequency connections are dotted. Optical fibre connections are shown thick.

emerges from the RF port of the bias-tee, which is modulated at 200.00 MHz is mixed with a 200.01 MHz signal from the second oscillator to provide down conversion of the RF frequency signal to 10 kHz, where a lock-in amplifier can be used as a phase detector. The lock-in amplifer is used at a very short time constant (10 μs) and so the lock-in is used for its phase measuring abilities, not for resolving signals buried in the noise, which is often the use of lock-in amplifiers. The phase shift and amplitude of the 10 kHz AC component of the PMT signal is then measured on the lock-in amplifier and passed to a computer. The DC component of the PMT current (usualy less than 5μA) which comes from the DC port of the bias-tee, is converted to a voltage with a current-to-voltage converter and then measured on a 12-bit A/D board fitted in the computer. There are some added complications created by the fact the lasers are multiplexed, although these are not shown in figure 1.16.

Using the diffusion equation, Fishkin et al[72] analysed theoretically the propagation of light in an infinite strongly scattering medium, where the source was sinusoidally modulated. They found that the photon density propagates with wave-like properties, with a phase velocity that depends on the optical properties of the medium and the source modulation frequency, but independent of the distance from the source - assuming the distance was large enough for the diffusion equation to apply. Assuming no absorption ($\mu_a=0$), the wavelength is then

$$\lambda = 2\pi \sqrt{\frac{2vD}{\omega}} \qquad\qquad \textbf{1.63}$$

where v is the velocity of propagation of the electromagnetic wave in the medium (c/n), D is the diffusion coefficient (equation 1.32) and ω the angular frequency of the source modulation. The propagation constant β is simply $2\pi/\lambda$, so β is given by equation 1.64:

$$\beta = \frac{2\pi}{\lambda} = \sqrt{\frac{\omega}{2vD}} \qquad\qquad \textbf{1.64}$$

The phase velocity v_{ph} is given by equation 1.65.

$$v_{ph} = \frac{\omega}{\beta} = \sqrt{2vD\omega} \qquad\qquad \textbf{1.65}$$

In the more useful regime of non-zero absorption and large scattering, equations were derived for the three experimentally determined quantities in a uniform infinite medium:

$$\phi = r\left(\frac{v^2\mu_a^2 + \omega^2}{v^2 D^2}\right)^{1/4} \sin\left[\frac{1}{2}\tan^{-1}\left(\frac{\omega}{v\mu_a}\right)\right] \qquad\qquad \textbf{1.66}$$

$$\log_e(DC_{detector}) = -r\sqrt{\frac{\mu_a}{D}} + \log_e\left(\frac{S}{4\pi vD}\right) \qquad\qquad \textbf{1.67}$$

$$\log_e(AC_{detector}) = -r\left(\frac{v^2\mu_a^2 + \omega^2}{v^2 D^2}\right)^{1/4}\cos\left[\frac{1}{2}\tan^{-1}\left(\frac{\omega}{v\mu_a}\right)\right] + \log_e\left(\frac{Sm}{4\pi vD}\right) \qquad\qquad \textbf{1.68}$$

where S is a fluence of the source in photons/second, and m is the modulation depth of the source. Equations 1.66 to 1.68 are linear functions of the source to detector separation r, but are non-linear functions of the tissue optical properties μ_a, μ_s, v and g and the source modulation frequency ω. A non-linear least squares fitting algorithm is likely to be necessary to determine the optical properties from measurements made with a frequency domain phase system.

For small changes of absorption in tissue, such as occur due to changes in oxygenation, the change in measured attenuation is directly proportional to the change in absorption, the

constant of proportionality being known as the differential pathlength (DP). Currently the most accurate method of determining the DP and hence determining the change in absorption from the change in measured attenuation, is by measuring the mean time of flight of the photons[2].

$$<t>=\frac{\int\limits_{-\infty}^{+\infty} h(t)t\, dt}{\int\limits_{-\infty}^{+\infty} h(t)\, dt}$$ **1.69**

Duncan et al[61] has shown that in a non-dispersive, attenuating medium, there is a simple relationship between the phase lag ϕ and the mean time of flight $<t>$,

$$<t>=\frac{\phi}{\omega}$$ **1.70**

although this relationship changes at frequencies above approximately 200 MHz due to the dispersive nature of tissue.

Figure 1.17 shows the modulation attenuation, modulation depth and phase measured on the head of a human fetus during labour at University College Hospital, with a later development of the instrument shown in figure 1.16. The upper graph of the figure 1.17 shows the contractions, measured on a HP-8040A cardiotocgraph.

Kohl et al[73] have shown one method of obtaining the optical properties of tissue using a frequency domain optical spectrometor such as that in figure 1.16. By making 2 measurement types from 3 possible ones (intensity, modulation depth and phase) before and after a change in either optode spacing or absorption, and looking at the ratios of changes in intensity, modulation depth or phase, it is possible to calculate μ_a and μ_s'. The changes in absorption caused by the change of volume of blood in the head, caused by the contractions shown in figure 1.17 can be used for this purpose.

1.7   Measurement of the temporal point spread function (TPSF).

There are a number of ways of determining the TPSF of optical signals. A number of these, will be briefly discussed.

1.7.1   Measurement of the TPSF with a photo-detector and sampling oscilloscope.

The simplest and most direct method of measuring the temporal profile of a repetitive
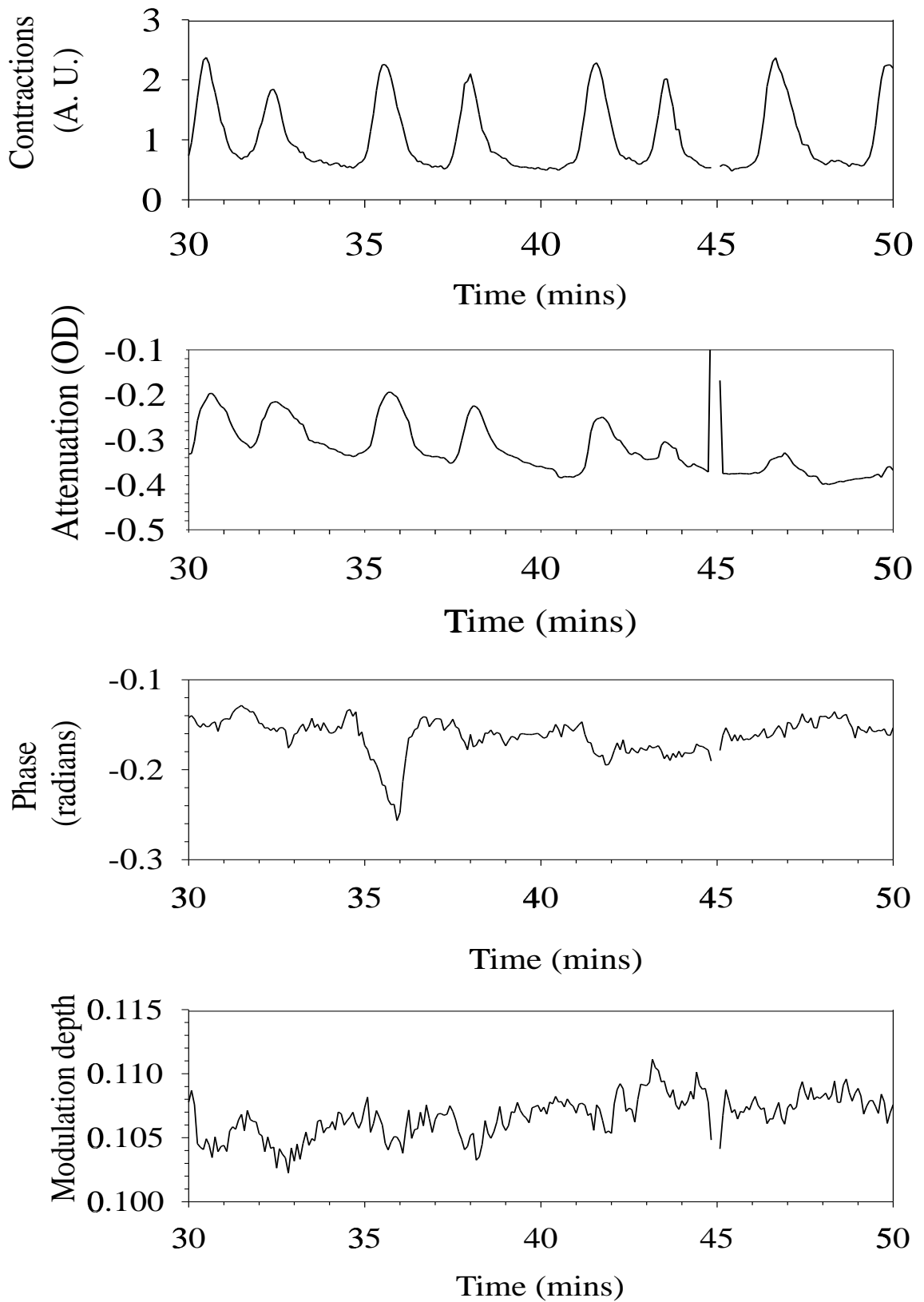
Figure 1.17 Measurements of attenuation, phase and modulation depth measured on a human fetus during labour. Data from Kirkby and Chipchase (unpublished).

optical pulse is to illuminate a properly biased photodetector with the light and view the resulting current waveform on a high-speed sampling oscilloscope. However, for the intended application, this is not a suitable method for a number of reasons.

(1) The high speed oscilloscope required is very expensive. Currently Hewlett-Packard's HP-54124T (7 ps risetime, 50 GHz bandwidth) which is the present state of the art in sampling oscilloscopes, costs £41,000 (April 1996 price).

(2) The intensity of light available from fast (small area) photodiodes after passing through tissue is too low for this method. Fast photodiodes have necessarily very small active areas, and so the signal amplitudes would be very low. However, with the gradually increasing powers from pulsed semiconductor lasers, this may be possible in the near future.

1.7.2   Measurement of the TPSF with a streak camera.

Another direct method of measuring the temporal profile of a laser pulse is to use a streak camera. Although called streak cameras for historical reasons, they have little similarity to a normal camera. Streak cameras are currently made by a number of manufacturers including Hamamatsu Photonics, Photonics Analysis Ltd and Cordin. Figure 1.18 shows a typical streak camera. Incident light enters through a small slit and is focused by a lens onto the photocathode of the streak tube. Typically an S-25 photocathode (wavelength range 400 nm to 900 nm) would
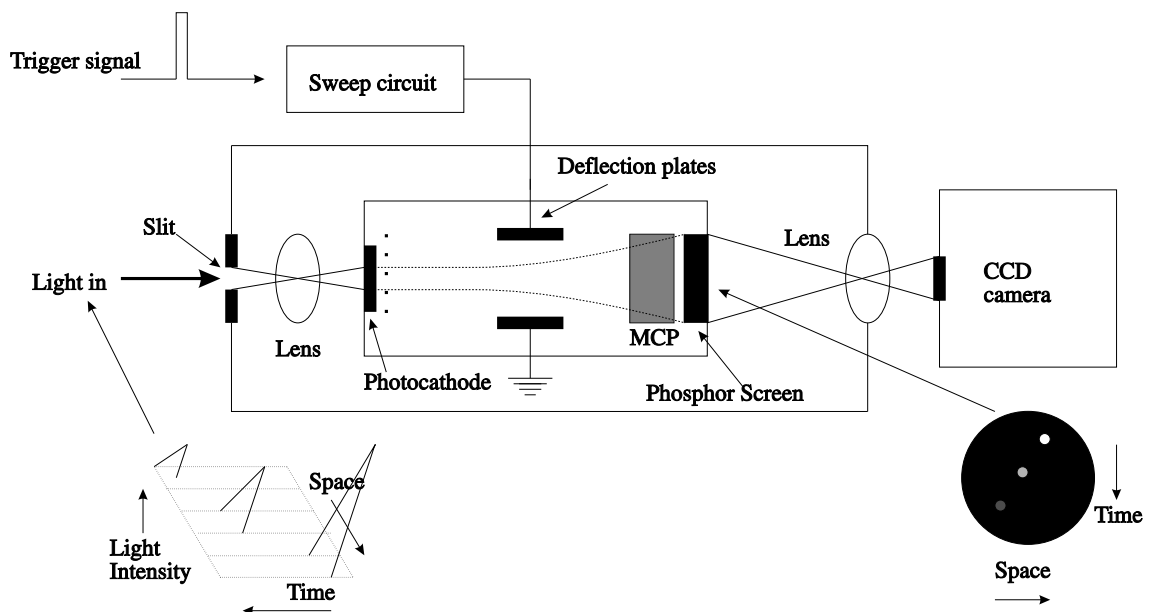
Figure 1.18 Diagram showing a typical streak camera, used for measuring the temporal profile of a light pulse. MCP = Microchannel Plate intensifier.

be used for measuring the wavelengths of light necessary here. This photocathode emits electrons -ideally one per incident photon. The electrons are than deflected by the sweep voltage applied to the electro-static deflection plates, before reaching a microchannel plate (MCP) which has an internal gain mechanism and produces up to 10,000 electrons for each incident electron. The electrons emitted from the MCP reach the phosphor screen (typically a type P-20) and cause the phosphor to emit light. The phosphor screen is viewed by a readout system, which is typically a charge coupled device (CCD) camera. Since a time dependent deflection voltage is applied to the deflection plates, the phosphor screen shows time on the vertical axis and the intensity of the input light as an intensity variation on the screen. The data from the CCD camera is usually read into a computer for analysis. The time axis is usually non-linear, due to the sinusoidal waveform used on the deflection plates, although this is easily corrected in software. A shutter is fitted to the streak camera, so measurements of the dark signal may be made - these are then subtracted in software.

When the input signal is of a low-intensity, but repetitive, it is possible to increase the signal to noise ratio by averaging the data for many input pulses. In this case, it is necessary to synchronise the deflection voltage to the input by using a reference pulse from the laser.

1.7.3   Measurement of the TPSF with a sampling optical oscilloscope.

The optical oscilloscope is similar to the streak camera, but unlike the streak camera, it is only capable of resolving data from one input point rather than a slit, although some models, such as the Hamamatsu OOS-01, allow a second sampling head to be attached, so two signals can be sampled simultaneously.

Figure 1.19 shows the operating principles of a typical sampling optical oscilloscope. The light to be measured is first focused on a photocathode. The type of photocathode would depend on the wavelength of light, but for use in the near infra-red, a type S-25 would be suitable. The photocathode creates electrons, which are accelerated towards a phosphor screen. A sweep voltage, is applied to two vertical deflection plates, which moves the electron beam vertically with time. As the electrons pass the vertical deflection plates they are deflected at high speed past a narrow slit on the sampling slit plate. Only those electrons that pass the slit plate reach the phosphor screen, so only a small section of the input waveform is sampled. The sampled electrons cause the phosphor screen to emit light, which is detected with a photomultiplier tube and then amplified before driving the vertical deflection plates of a cathode ray tube (CRT). The sampling process is repeated many times, with a different time delay
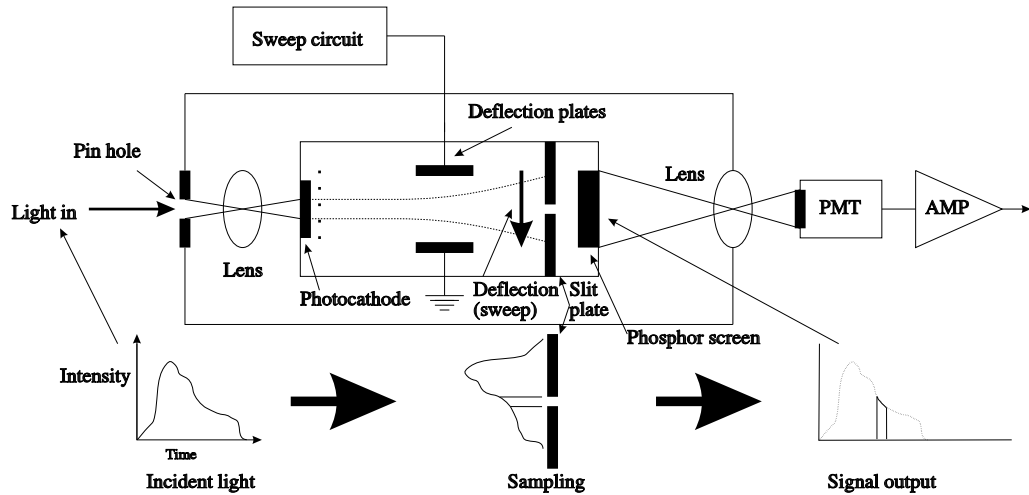
Figure 1.19 Diagram showing the principle of operation of a typical optical sampling oscilloscope.

between the voltage on the deflection plates and the input signal. The full temporal profile of the input may thus be determined.

The following differences should be noted between the streak camera (see section 1.7.2) and the sampling optical oscilloscope.

1) The optical oscilloscope has one input channel, using one streak tube. A second input channel requires a second streak tube. In contrast, the streak camera can have many tens of inputs arrayed along the slit, to allow time-resolved spatially resolved data to be collected.

2) The optical oscilloscope uses a sensitive high-gain photomultiplier tube to provide gain in the conversion from light to electricity, whereas the streak camera uses a microchannel plate to provide optical amplification before the detector, which is a CCD.

3) The streak camera may measure the whole TPSF in a single shot if the light intensity is sufficiently high.

1.7.4   Temporal response of optical pulses by intensity auto-correlation using second harmonic generation.

A fairly common method of measuring the temporal profile of laser pulses is an indirect method whereby the intensity auto-correlation of the signal is obtained by using a non-linear crystal to generate a second harmonic of the input. Typical non-linear crystals are $LiIO_3$, $KNbO_3$ and $KTiOPO_4$ (also known as KTP). A typical setup is shown in figure 1.20. The laser pulse with an electric field E(t) is incident on the non-linear crystal, as well as a delayed version $E(t+\tau)$. Hence the resultant electric field reaching the non-linear crystal is $E(t)+E(t+\tau)$. This is
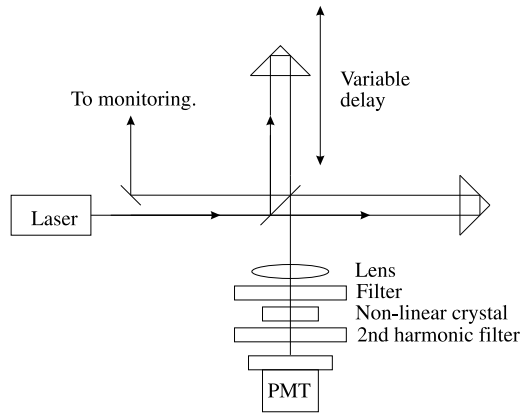
Figure 1.20 Typical experimental setup for measuring the temporal profile of a laser using the auto-correlation function with second harmonic generation.

multiplied in the non-linear crystal to give $(E(t)+E(t+\tau))^2$, ie a second harmonic at $\lambda/2$. The PMT output current is proportional to the optical power or the square of the resultant E-field intensity. As the variable delay τ is changed, the signal measured is:

$$I(\tau) = \int |[E(t)+E(t+\tau)]^2|^2 \; dt \qquad \textbf{1.71}$$

From this it is possible to determine the temporal characteristic of the laser pulse. However, the simple intensity auto-correlation is symmetrical even if the laser pulse is asymmetrical, as can be easily seen from the definition of the auto-correlation

$$R(-\tau) = \int_{-\infty}^{\infty} f(t)f(t-\tau)dt = \int_{-\infty}^{\infty} f(x)f(x+\tau)dx = R(\tau) \qquad \textbf{1.72}$$

although it is possible to use higher order correlations to exactly determine the pulse shape and duration.

This method has a very good temporal resolution down to tens of femtoseconds, but the two beams in the crystal need to meet accurately, which is not really suitable for tissue, although some work using this technique has been performed in scattering media[74], using the cross-correlation method. One of the other problems with this method, is that the optical powers needed to cause sufficient non-linearity in the crystal is quite high[75].

1.7.5   Measurement of the TPSF using time-correlated single-photon counting.

Another method of measuring the TPSF of light is that of time-correlated single-photon counting[76], which is a system first developed for time domain fluorescence spectroscopy. This commonly uses a photomultiplier tube as an optical detector, although germanium APDs[77,78,79,80], silicon APDs[80,81] and InGaAs[80] APD's have also been used. These various semiconductor detectors have the advantage that the wavelength range that may be detected is wider than with PMTs.
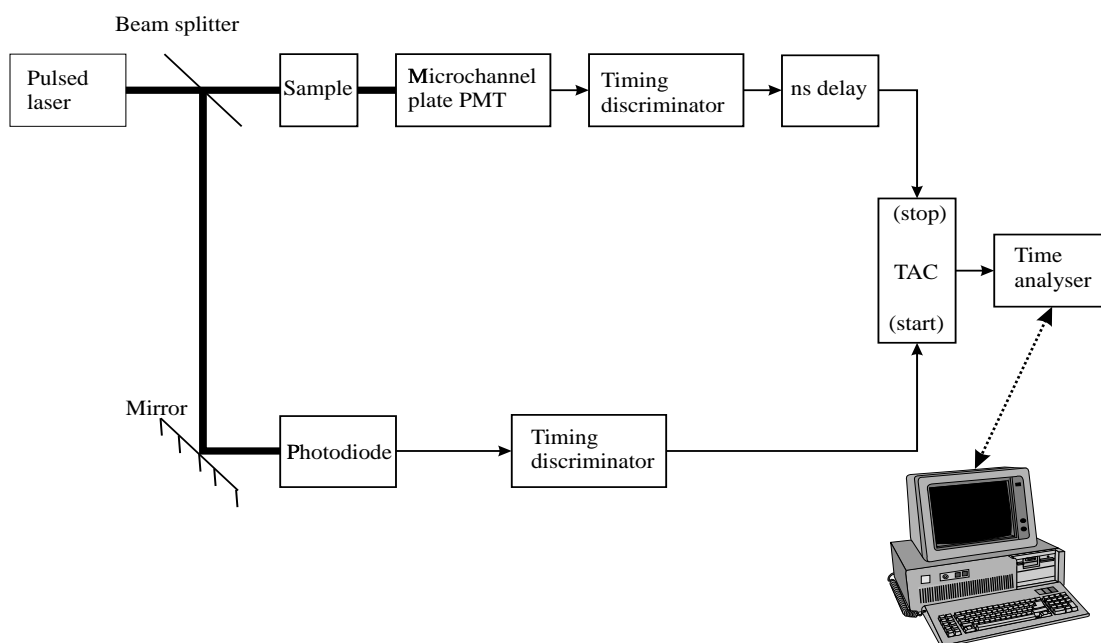
Figure 1.21 Simplified block diagram of a time-correlated single-photon counting system for measuring the temporal profile of light emitted from a fluorescent material or tissue. Optical paths are drawn as thick lines.

A simplified block diagram of a time-correlated single-photon counting system using a PMT as the main detector is shown in figure 1.21[82]. A pulsed light source is split into two components by a beam splitter. A small fraction goes to the sample, but the majority is converted into a short duration electrical pulse by a photodiode. The output of the photodiode drives a discriminator, which converts the analogue pulses into digital pulses, which by convention are usually negative going. The negative going pulses from the discriminator drive the start input of a time to amplitude converter (TAC). This starts a constant current flowing into a capacitor in the TAC.

The sample emits photons, which are collected on a single photon counting photomultiplier tube. This produces a pulse of current, which drives another discriminator, then a delay line, which is used to trigger the stop input of the TAC. This stops the current flowing into the capacitor of the TAC.

Since the voltage across a capacitor of capacitance C (Farads), which is initially uncharged, is related to the charging current I(t) by equation 1.73

$$V = \frac{1}{C}\int_0^t I(t)\, dt \qquad\qquad \textbf{1.73}$$

then for a constant current I flowing for a time t, the voltage across the capacitor is

$$V=It \qquad \textbf{1.74}$$

so the capacitor in the TAC develops a voltage proportional to the time between the start and stop pulses. This process is repeated and so a profile of the distribution of light exiting the sample is obtained, by analysing the voltages on the capacitor. This analysis is performed using a pulse height analyser.

Since the gain experienced by an photoelectron is not fixed in the PMT, but there is a noise associated with the gain, the analogue output pulses that the PMT produces are not all of the same amplitude. The discriminator converts these pulses to a digital signal, such that the digital output occurs at a constant fraction (typically 20%) of the pulse height. This allows precise timing of the digital pulse. If a simple comparator were used instead of a constant fraction discriminator, the time of the start and stop pulses that reach the TAC would depend on the gain of the PMT, which has noise associated with it.

If the probability of detecting more than one photon after the laser pulse is high, the TAC will respond to the first of the photons and so give a distorted temporal profile, skewed to short times. Hence this condition must be avoided. This is done by ensuring the light level leaving the sample is sufficiently low that the probability of receiving a photon is less than 0.01 per pulse. Hence the probability of receiving two photons (which causes the distortion) is less than 0.01^2, which is less than 1% of the single photon probability.

A nanosecond delay line is incorporated between the discriminator and the stop input of the TAC, to insure that the stop input does not occur too quickly. There is a minimum time, known as the dead time, before the TAC can be stopped.

Since only 1% or less of the incident pulses produce a detection pulse, when this is combined with the dead time mentioned above, the system as described would spend a long time giving no useful information, as 99% of the *start* pulses would produce no *stop* pulse. To avoid this, the start and stop inputs are often reversed, as shown in figure 1.22, which means any start pulse coincides with a useful detection photon. However, the number of start pulses is now much lower than the system shown in figure 1.21, so the dead time is not a problem.

Figure 1.23 shows a diagram of 32 channel time-domain imaging system under development at UCL. The output from a mode-locked Titanium:Sapphire laser (Ti:Sap) is split with a beam-splitter into two signals. A small fraction of the light generates reference pulses on a fast photodiode (PD), whereas the remainder is coupled to a 32 position fibre switch. Light
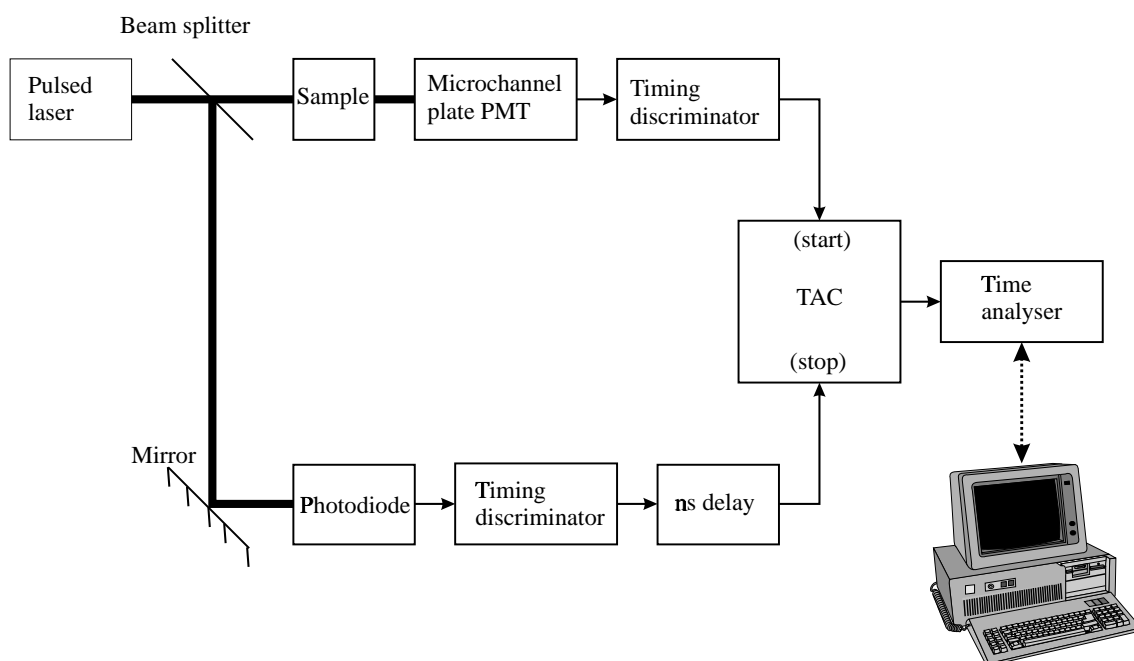
Figure 1.22 Reversed *start* and *stop* connections of a time-correlated single-photon counting system, which has reduced dead-time compared to that in figure 1.21.

FS: Fibre Switch, VOA: Variable Optical Attenuator, PF: Polymer Fibre, LPF: Long-Pass Filter, MCP-PMT: Multichannel Plate-Photomultiplier Tube, PA: Pre-Amplifier, CFD: Constant Fraction Discriminator, PTA: Picosecond Time Analyser, PD: Photodiode, PTD: Pico-Timing Discriminator
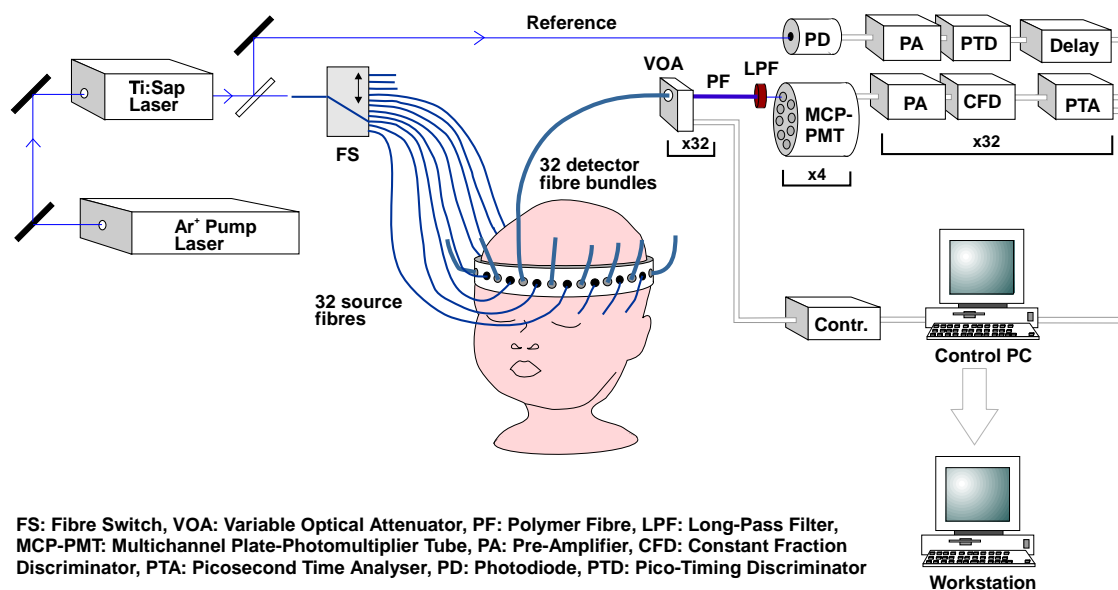
Figure 1.23 Diagram of a 32 channel time-resolved imaging system, currently under development at University College London.

from this is coupled into the sample, and detected on 32 detector fibres. The light in each detection fibre is attenuated by 1 of 32 variable optical attenuators, before being passed to a long-pass filter (LPF) that passes near infra-red, but attenuates daylight. The light then passes

to one of 4 microchannel plate photomultiplier tubes (MCP-PMT), each of which has 8 detection cathodes, which convert light into a series of electrical pulses. There is statistical variation on the amplitude of these pulses, so a constant fraction discriminator (CFD) is used to convert the electrical signals into digital pulses that are generated at a time when the pulse amplitude is a constant fraction of the peak amplitude. The picosecond time analyser (PTA) then measures the arrival time of the photons. The above system is very big, requiring a large 19" rack, plus a trolly for a refrigeration system to keep the MCP-PMTs at the correct temperature, but looks very promising. Some preliminary images of an optical phantom, using only 8 of the 32 channels, have been published[83].

1.7.6   Temporal response of optical signals using a cross-correlation method.

It may be shown that under certain conditions, measurement of the cross-correlation of the optical signals at the input and the output of a linear, time-invariant[*1] system can allow the TPSF or impulse response to be determined. Standard control theory[84,85,86,87] states that the output y(t) of such a system is the convolution of the input x(t) and the system impulse response h(t). That is,

$$y(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} h(t-\tau)x(\tau)d\tau \qquad \textbf{1.75}$$

Figure 1.24 shows the inputs and output in both the time and frequency domain of a linear, time invariant, causal system. It can be shown[84,85,87] that an equivalent relationship exists in the frequency domain, whereby the Fourier transform of the output $Y(\omega)$ is the product of the Fourier transform of the input $X(\omega)$ and the Fourier transform of the system response $H(\omega)$.
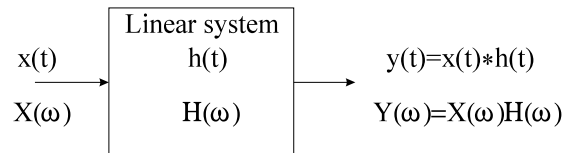
Figure 1.24 Diagram showing the time domain and frequency domain relationships between the input and output signals in a linear, time-invariant, causal system.

$$y(t) = x(t) * h(t) \Leftrightarrow Y(\omega) = X(\omega)H(\omega) \qquad \textbf{1.76}$$

The cross-correlation R_{xh} of the signals x(t) and h(t) has a similar, but different definition

---

[*1] Linear means that if input $x_1(t)$ gives an output $y_1(t)$ and an input $x_2(t)$ gives an output $y_2(t)$, then the input $x_1(t)+x_2(t)$ will give an output $y_1(t)+y_2(t)$. Time invariant means if an input x(t) gives an output y(t), an input x(t+T), gives an output y(t+T).

to the convolution integral and is defined as[85]:

$$R_{xh}(\tau) = \int_{-\infty}^{\infty} x(t)h(t+\tau)dt \qquad \textbf{1.77}$$

where τ is a continuous time shift in the range $-\infty \le \tau \le \infty$. The Fourier Transform pair for the cross-correlation is:

$$\int_{-\infty}^{\infty} x(t)h(t+\tau)dt \Leftrightarrow H(\omega)X^*(\omega) \qquad \textbf{1.78}$$

where the $X^*(\omega)$ indicates the complex conjugate of $X(\omega)$. If the input signal $x(t)$ is an even function of time[*1], then the Fourier transform is real, so $X^*(\omega)=X(\omega)$. For an input $x(t)$ consisting of a delta function at $t=0$ ($x(t)=\delta(0)$), this condition is met so the cross-correlation R_{xh} of $x(t)$ and $h(t)$ is the same as the convolution of $x(t)$ and $h(t)$.

$$y(t)=x(t)*h(t)=R_{xh}=\int_{-\infty}^{+\infty} h(t-\tau)x(\tau)d\tau = \int_{-\infty}^{+\infty} x(\tau)h(t+\tau)d\tau \qquad \textbf{1.79}$$

Hence by probing a region of tissue with an even input signal $x(t)$ and taking the cross-correlation of this input signal and the output $y(t)$, the impulse response $h(t)$ of the tissue can be measured.

Hence by measuring the cross-correlation of the input and output signals to a region of tissue, the impulse response $h(t)$ of the tissue is measured. Note that since this was derived for arbitrary signals, it is not dependent on any signal being either periodic or non-periodic, only that $x(t)$ is even.

1.7.7 Hardware for TPSF measurement using a cross-correlation technique.

In 1987[88] Berndt described a further development of the systems described in section 1.6.2 for measuring dye lifetimes. This newer system, was a cross-correlator and therefore gave extra information compared to the earlier phase fluormeters, which are unable to determine the complete TPSF. Since this method was the inspiration for this project, this will be described in detail. In this system, (see figure 1.25), the laser was run with all modes locked, so instead of

---

[*1] $x(t)$ is said to be even if $x(t)=x(-t)$, then the Fourier transform of $x(t)$ is even and real. If $x(t)=-x(-t)$, then $x(t)$ is odd and its Fourier transform is odd and imaginary.
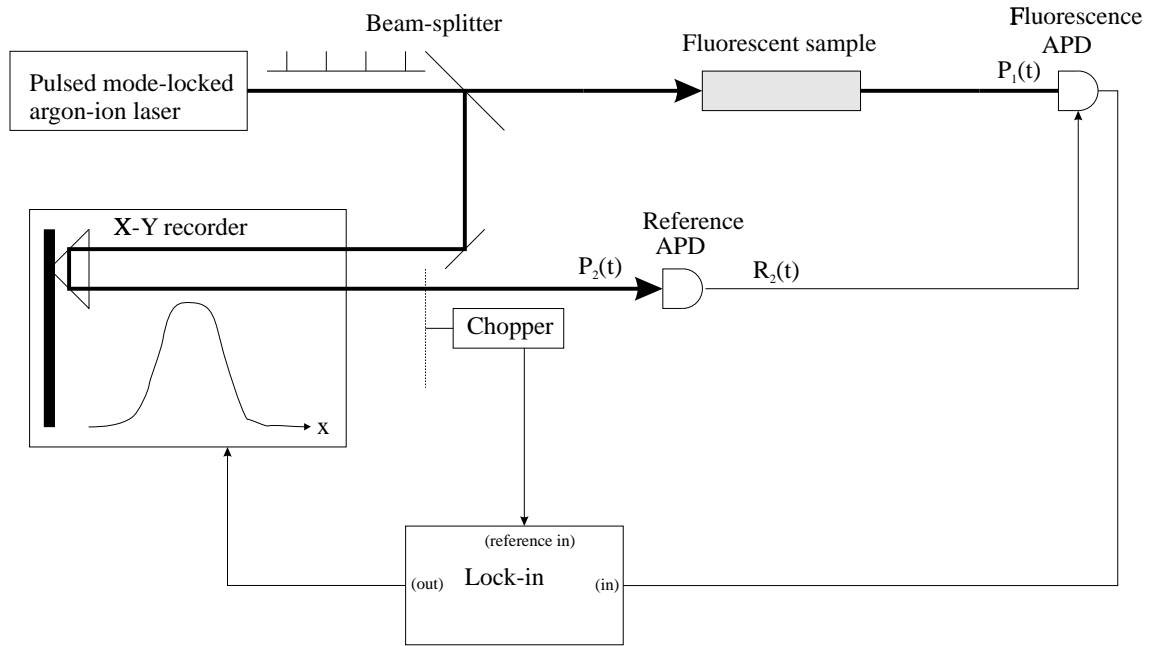
Figure 1.25 Cross-correlator developed by Berndt[88] for measurement of the lifetime of fluorescence dyes. This is similar in principle to the system developed for this thesis.

producing an amplitude modulated output, its output was a series of pulses, 110 ps FWHM, at a rate of 125 MHz. A beam splitter placed in the output splits some of the light to the sample and some on to a silicon APD, type TIXL 55 (Texas Instruments), via an optical chopper and a time delay mechanism, consisting of a corner cube reflector placed on the carriage of an XY chart recorder. The signal reaching the reference APD generated pulses of 320 ps FWHM at 1.2 - 1.5 V in amplitude, with which to gain-modulate the fluorescence APD.

The weak light scattered from the sample is then detected by the second APD, which is a high speed (>200 GHz gain-bandwidth product), type BPW 28 made by AEG Telefunken, which will be called the fluorescence APD.

The pulses from the reference APD were used to change the bias voltage of the fluorescence APD from approximately 186.5 to 187.7 V, which changed its gain from approximately 414 to approximately 2000.

If the laser pulse to be measured is of the form $P_1(t)$, then the primary (unmultiplied) photocurrent is:

$$i_p(t) = S_1(t) + I_B + I_D + N(t) \tag{1.80}$$

where $S_1(t)$ represents the flourescence APD response to the light pulse $P_1(t)$, I_B is the current due to background radiation, I_D is the dark current and $N(t)$ is the noise current. During blocking

of the laser pulse to the reference APD, the gain of the flourescence APD is a constant M_0 and so the multiplied APD current is:

$$M_0(S_1(t)+I_B+I_D+N(t)) \qquad \textbf{1.81}$$

During the open periods of the chopper wheel, the APD gain depends on both the time t and the delay τ, such that:

$$M(t,\tau)=M_0(1+mS_2(t+\tau)) \qquad \textbf{1.82}$$

where $S_2(t)$ is the response of the reference detector to the optical pulses $P_2(t)$ and m is the fractional increase in gain. The lock-in amplifier performs two tasks:

1) The APD current is averaged with a time period T by the lock-in, so that:

$$I(\tau)=\frac{1}{T}\int_0^T i_p(t)M(t,\tau)dt \qquad \textbf{1.83}$$

2) The lock-in ouput is proportional to the difference between the APD currents in the high and low gain states:

$$\Delta I(\tau)=I(\tau,m\neq 0)-I(\tau,m=0) \qquad \textbf{1.84}$$

which gives

$$\Delta I(\tau)=\frac{M_0 m}{T_1}\int_0^{T_1} S_1(t)S_2(t+\tau)dt \qquad \textbf{1.85}$$

This is the cross-correlation[87] of the two periodic signals $S_1(t)$ and $S_2(t)$. Hence we have proved that the system developed by Berndt measures the cross-correlation and that this allows measurement of the TPSF.

A more intuitive, less mathematical explanation of how the cross-correlator works will now be given. It may be considered that the fluorescence APD has two gain states - one high the other low. For most of the time the APD's gain is low, so its integrated output current is simply the time averaged intensity of the TPSF. For a short duration the APDs gain is high, so its integrated output current is then the time averaged intensity of the input, but with more weight attached to a small section of the TPSF. The lock-in amplifier performs the integration in both cases, but then subtracts one signal from the other. So the output of the lock-in amplifier

is a sample of the TPSF, over the region where the APD gain was high. Since the time period for which the APD gain is high may be altered with the time delay mechanism, so the section of the TPSF that is sampled can be altered. This is in essence the principle of operation.

Another cross-correlation technique has been described by Choi[75] using a Michelson interferometer, and a low speed (~5MHz bandwidth) silicon photodiode detector. The system measured the pulse width from a mode-locked laser to be 258 fs ± 44 fs. Unfortunately, the system uses several mechanical choppers, and several mirrors that must be aligned very accurately. It is not easy to see how Choi's system could be adapted for clinical use.

1.7.8 Other methods of TPSF Measurement.

There are a number of other methods of measuring the temporal profile of light, but many are not suitable for tissue optics, due to the scattering. A review paper by Hebden et al[89] discusses a number of methods.

1.8 Design methodology for the instrument.

We have seen there are essentially 3 methods by which the optical properties of the tissue, μ_a and either μ_s or μ_s' can be obtained *in vivo*.

(1) Using CW intensity systems. This is certainly the cheapest, since the electronics can be slow (time constants of the order of a second or more) and therefore inexpensive. Using the diffuse reflectance R_s as a function of radial distance r from the detector is one method that looks attractive and a prototype instrument based on this technology can be found in our laboratory. However, it would be difficult to adapt to the *in vivo* measurement via small optical fibres.

(2) Using frequency domain spectrometers which have an amplitude modulated light source. The phase, amplitude and modulation depth of the detected optical signal are then measured - not just the intensity as with CW systems. This is more expensive than the CW systems, as the source must be modulated (adding complexity) and the detection electronics are much more complex. Above approximately 500 MHz, standard photomultiplier tubes are unusable and so considerably more expensive microchannel plate PMT's are necessary, or very small area photodiodes.

(3) Using a time-domain approach. The time domain equipment can easily measure all the data that the CW systems can, since it is easy to integrate the received signal over time to have a measure of the total light. Via the Fourier transform, the time-domain approach is equivalent to the frequency-domain methods, but since very short pulses (<100 ps) are much easier to generate from lasers than good quality amplitude modulation with similar (10 GHz) frequency
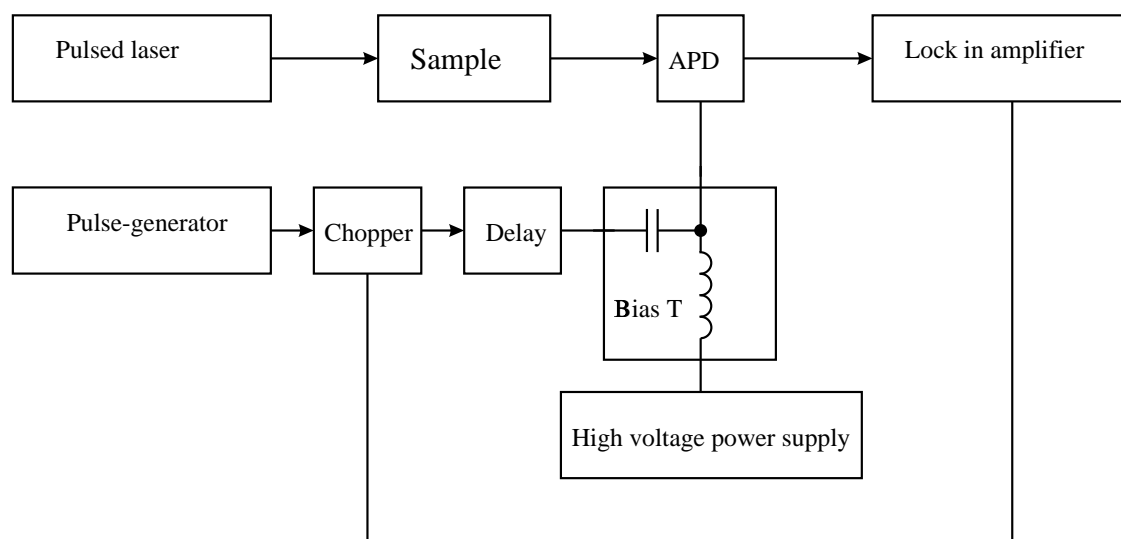
Figure 1.26 Diagram showing the approach considered as a possible design for the cross-correlator. At this stage, the form that the pulse generator, delay and chopper would take were unknown.

components, the time domain approach seems in principle to offer the most data, even though some of the electronics must handle the very fast (100 ps) signals. Of course extra data is not worth obtaining if it can not be used, but the work by Schweiger et al[48] shows that time domain data is better for reconstruction of μ_s than intensity data measured with an amplitude modulated laser source, so there would appear to be an advantage in going to time-domain methods, even though they are more complex. Frequency domain methods are often thought to be cheaper than time domain methods, but this may not be an accurate assumption when frequency domain methods are compared to the time-domain method pioneered by Berndt et al, which has few expensive components, so this approach was considered the best. Figure 1.26 was considered a possible design for an electro-optic cross-correlator. The pulse generator must provide pulses that are synchronised with the laser pulses, but at this stage the form the pulse generator would take was unknown. The choice of the pulse generator, chopper and delay line were all interdependent, as it was essential that they were compatible.

1.9 Initial design specifications of the instrument.

Once the basic design methodology was decided - i.e. a time domain system based on the cross-correlation technique, the detailed specifications were considered. The design specifications are broken down into a few main areas - portability, cost, temporal resolution, sensitivity and a few minor considerations.

The instrument should run from a standard 230 V, 50 Hz, single-phase mains outlet -

battery operation was not considered necessary. The instrument should be fibre coupled, for both connections to the light source and the detector. For maximum tolerance to mechanical shocks, the instrument should have no moving mechanical components and have no optical components requiring precise alignment. The instrument must be capable of being moved easily.

With a streak camera currently costing of the order of £100,000 the system must be less expensive too produce.

Although it will be impossible to predict the exact effect that the temporal resolution of the instrument has on the accuracy with which μ_s' changes are estimated, without performing a detailed time-consuming Monte-Carlo analysis, we know that typical TPSF's are of the order of 1 ns in length, so really a temporal resolution of about 100 ps would be desirable, so that the measured data is not adversely effect by the instrument response. However, a longer response would not be too serious, at it should be possible to deconvolve the instrument response from the measured data.

A number of other features are desirable. These include:

1) Device must not easily be damaged by excess light, or operator error.

2) The data should be as real-time as possible, since if used as an aid during laser induced thermal therapy, it is essential that the data is available quickly.

Chapter 2.

Short Pulsed Laser Sources.

2.1   Introduction.

Although it is in principle possible to use a multi-mode continuous wave (CW) laser to obtain the impulse response of tissue, relying on the mode noise to generate high frequency modulated light[90], better S/N ratio is obtainable if a pulsed laser is used. For this project, a short (<50 ps) pulsed light source is an essential requirement. The purpose of this chapter is not to explain the principle of operation of lasers. Numerous books[91,92,93,94,95,96,97] are available to do this. Here, methods of obtaining short pulses from lasers are described, with a particular emphasis on semiconductor lasers - a subject given very little treatment in most text books, with one notable exception[1], although there are numerous research papers published in this field. These techniques for short pulse generation include gain-modulation, mode-locking and Q-switching. However, before dealing with pulse generation in particular, a short review of laser theory is presented, mainly to define the symbols used.

Consider first a simple laser with 3 energy bands - usually known as a 3-level laser, although each band will contain a number of different energy levels, so there will many more than 3 different levels. Atoms can be in the ground state with an energy ε_0 or one of two excited states with energies ε_1 or ε_2, such that $\varepsilon_0 < \varepsilon_1 < \varepsilon_2$, as in figure 2.1.

Boltzmann statistics shows that at thermal equilibrium, the number of atoms N_0 in the ground state is greater than the number of atoms N_1 at energy ε_1 and that N_1 is greater than the number of atoms at energy ε_2 which is N_2, so $N_0 > N_1 > N_2$. In a laser, atoms in the ground state are pumped to the ε_2 state by applying energy, then they rapidly decay to ε_1, building up a population inversion between ε_1 and ε_0, such that $N_1 > N_0$. Decay from ε_1 to the ground state results in emission of radiation at a longer wavelength (less energy) than the pump energy.

The energy bands ε_0, ε_1 and ε_2 are not single levels, but each are a range of discrete energy levels. If ε_2 was a single energy level, pumping would be extremely difficult and inefficient, as the pump energy would need to be precisely ε_2-ε_0. If ε_1 - ε_0 was a single level, mode-locking (see section 2.3) would be impossible.

Decay from ε_1 to ε_0 can be due to either spontaneous emission or stimulated emission

and in general will consist of both. When the laser material is placed in an optical resonator, consisting in its simplest form of a totally reflecting mirror and a partially reflecting mirror, light will be emitted from the partially reflecting mirror. As photons pass through the laser medium, they can stimulate further emission, resulting in a predominance of stimulated emission, which is what defines laser action, as is obvious from the acronym laser - **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. The fact that one incident photon can stimulate

Figure 2.1 Diagram showing the energy levels in a laser with 3 energy bands.

more than one photon to be emitted, implies the laser material has gain. Like any oscillator, whether it mechanical, electrical or optical, the losses in a laser cavity must be small, otherwise there is no optical gain and hence no lasing action. Hence for CW laser operation, the reflectivity of the mirrors must be sufficiently high to allow laser operation. It can be shown[1] that the gain of the laser material increases as the population inversion increases, which can be used to allow gain-modulation.

If the pump energy is continuously applied, such as through a high voltage discharge in a HeNe laser, argon-ion laser or CO_2 laser, the output can be continuous. For this work, a CW laser is not suitable, so some method must be used to convert the output into a pulsed waveform. Such methods are now discussed.

2.2   Q-switched Lasers.

The Q of a laser system is defined as:

$$Q = 2\pi \frac{\text{energy contained within cavity}}{\text{energy loss per cycle}} \qquad \textbf{2.1}$$

If the laser is pumped while the losses in the cavity are high, a large population inversion will build up - ie $N_2 \gg N_1$. This will give a laser medium with very high gain, but the system will not lase, due to the high cavity losses. If the losses in the cavity are suddenly reduced, the high gain medium will give a short, high power laser pulse, with a much larger peak power than the laser

could provide if run CW. A number of methods are available for switching the Q of the laser cavity. These will be considered.

2.2.1   Q-switching with a rotating mirror.

Conceptually the simplest method of Q-switching is to make one of the end mirrors rotate rapidly. For most of the time, the alignment of two mirrors will be such that the loss will be high and hence the Q low. This will allow a large population inversion to develop. At the instant that the two mirrors are aligned, the Q will be high and a large output pulse will be developed. Although this method was the first developed for Q-switching, it does not give the performance in terms of peak power of some of the newer methods of Q-switching.

2.2.2   Active Q-switching using an electro-optic modulator.
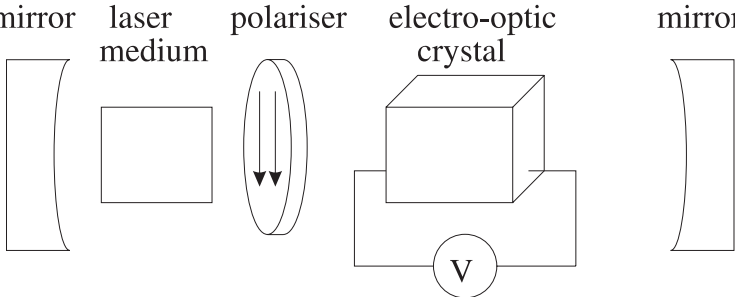
Another method of Q-switching is to place an electro-optic modulator in the laser cavity, as shown in figure 2.2. Light leaving the laser medium is randomly polarised, but is then vertically polarised by the polariser. With a DC voltage applied to the electro-optic modulator, it acts as a quarter-wave plate, converting the vertically polarised light entering its

Figure 2.2 Electro-optic crystal used to Q-switch a laser.

left-hand side to circular polarisation. The hand of the polarisation is reversed after reflecting from the mirror. Light entering the right hand side of the electro-optic modulator is then converted to horizontally polarised light, which will not pass through the vertical polariser. This creates a high cavity loss and hence low Q. If the voltage is removed from the electro-optic modulator it ceases to rotate the polarisation and hence the cavity Q is high.

2.2.3   Passive Q-switching using a saturable absorber.

A more common method of Q-switching is to use a saturable absorber (dye), with an absorption characteristic similar to that shown in figure 2.3. The dye is usually pumped through a small jet, which sprays the liquid through the laser beam. At the beginning of the pump cycle the incident intensity on the dye is low. The dye is highly absorbing, which gives a low cavity Q. Later on during the pump cycle, the irradiance is much higher, which bleaches the dye. Hence

the dye becomes transparent, giving a sharp increase in cavity Q. This increase in Q gives rise to an overall gain and so a high amplitude output pulse.

2.2.4 Q-switching of semiconductor diode lasers.

Q-switching offers the highest peak powers of any laser diode pulse generation method, by a few orders of magnitude.

2.3 Mode-locking.

As is shown in figure 2.1, the number of energy levels in a laser are quite large, not just the 3 or 4 usually used to explain laser operation. It can be seen that there is a possibility of a large number of laser transitions from

Figure 2.3 Absorption as a function of incident light irradiance for a saturable absorber.

energy level ε_1 to ε_0 each with a slightly different energy, since both ε_1 and ε_0 are energy bands, not just levels. In a typical multi-mode laser, each of these possible laser transitions can give rise to a different longitudinal mode. Expressed mathematically, the total electric field $E(t)$ as a function of time can be written as:

$$E(t) = \sum_{n=1}^{N_M} E_n e^{(j\omega_n t + \phi_n)} \qquad \textbf{2.2}$$

Figure 2.4 The output of a mode-locked laser, with 10 modes locked is seen to be a periodic train of pulses, with a peak power 10 times the average power.

where E_n, ω_n and ϕ_n are the amplitude, frequency and phase of the nth mode. Usually these parameters are all time varying so the modes are incoherent and the total irradiance I is simply the sum of the individual irradiances:

$$I = \sum_{n=0}^{N_M} E_n^{\,2} \qquad \textbf{2.3}$$

If however the modes are forced to have the same relative phase to each other, such that $\delta_n=\delta$, then the total irradiance is now found by adding the individual electric fields, rather than the irradiances. Hence the total electric field can be written as:

$$E(t)=E_0 e^{-j\phi}\sum_{n=0}^{N_M-1} e^{j\omega_n t}$$

2.4

Assuming all modes have the same amplitude of oscillation, it can then be easily shown[69] that the irradiance as a function of time is given by summing a geometric series with the result that:

$$I(t)=E_0^2\frac{\sin^2(\frac{N_M\pi ct}{2L_{cav}})}{\sin^2(\frac{\pi ct}{2L_{cav}})}$$

2.5

where L_{cav} is the length of the laser cavity in m and c is the velocity of light (ms^{-1}). The optical output reaches a peak every T seconds given by:

$$T=\frac{2 L_{cav}}{c}$$

2.6

This means the laser cavity loss has to be modulated at a frequency:

$$f=\frac{c}{2 L_{cav}}$$

2.7

Figure 2.4 shows a diagram of the output of a laser with a cavity length of 1.829 m with 10 modes, all assumed to be of the same amplitude locked together. The laser output is seen to consist of a series of pulses, each separated by 12.195 ns, with a peak power of N_M times the power in any one mode. The irradiance reaches a maximum whenever $N_M \pi c t / L_{cav}=\pi p$ where p is an integer (corresponding to t=0, $2L_{cac}/c$, $4L_{cav}/c$, $6L_{cav}/c$ etc). Assuming all modes are of equal amplitude, the full width at half maximum pulse width τ_w is[91]:

$$\tau_w\approx\frac{1}{B_{osc}}=\frac{1}{N_M f}$$

2.8

where B_{osc} is the bandwidth of the oscillator. However, in the more practical case where the gain

of the material has a Gaussian shape, so the modes near the centre oscillate with a larger amplitude and the pulse width is given by[91]:

$$\tau_w = \frac{2\log_e 2}{\pi B_{osc}} = \frac{0.441}{B_{osc}} = \frac{0.441}{N_M f}$$

2.9

so the pulse width is narrower than that predicted by equation 2.8. This should be compared with the almost constant output level of a non-locked, multi-mode laser, where the output is essentially constant with small random variations with time.

2.3.1   Active mode-locking.

Although a laser has a tendency to mode-lock, some form of active mode-locking is usually used if a laser is intended to produce short repetitive pulses. If an electro-optic or acousto-optic modulator is placed in the laser cavity, with its transmission normally low but increased to a high value every $2L_{cav}/c$ seconds, the laser can be mode-locked. If a wave packet arrives at the modulator when its transmission is high, the wave packet will pass unaffected by its presence. Since the round trip time of the laser wave packet will be $2L_{cav}/c$ seconds, the same as the time period of the signal driving the modulator, the packet will be amplified by the laser material, unaffected by the modulator. Any parts of the wave packet which arrive out of phase with the modulation of the modulator, will be attenuated. Hence the modulator will ensure that all modes have a fixed phase relationship of $\Delta\phi=2\pi$, which is the condition for mode-locking. The bandwidth of the optical modulator, can be much less than the bandwidth of the optical output pulses, which can be very wide for a material with a broad laser transition. For example in the Ti:Sapphire mode-locked laser used during this project, the bandwidth of the optical pulses was approximately $B=1/T=1/(2\text{ ps})=5\times10^{11}$ Hz=500 GHz. However, the modulator used would not have such a wide bandwidth.

2.3.2   Passive mode-locking.

Certain dyes, known as saturable absorbers, exist which have a high attenuation at low light intensities, but which can be bleached, at high intensity, as is shown in figure 2.3. If such dyes are placed in the laser cavity, the action of pumping can give a high population inversion, since the spontaneous emission from the laser material is insufficient to initially bleach the dye. The irradiance is not constant, but will show short-term fluctuations. The laser medium can increase the irradiance of one of the peaks to such an extent that it passes with only a small

amount of attenuation, which allows laser action to take place. The dye becomes bleached, resulting a short duration output pulse.

2.3.3   Mode-locking of semiconductor diode lasers.

Mode-locked diode lasers are feasible, but there are other, simpler methods of generating short pulses from diode lasers. Since the diode lasers cavity is so short, typically 250 μm to 400 μm, the cavity losses would need to be modulated at a frequency given by equation 2.7, which would be between 375 GHz and 600 GHz. This would be very difficult to achieve. For this reason, mode-locked diode lasers use an external cavity to increase the overall cavity length, as shown in figure 2.5. This allows the mode-locking to be performed at a lower, more convenient frequency.
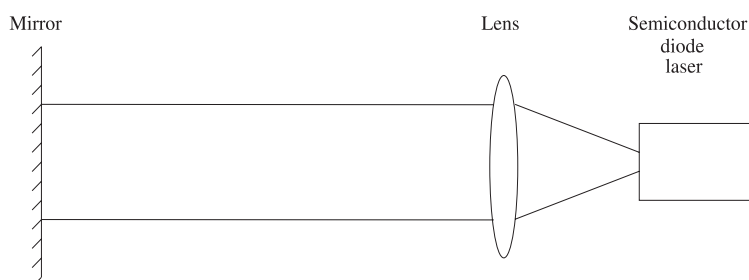
Figure 2.5 Method of coupling an external cavity to a semiconductor diode laser. The diode laser can then be mode-locked by modulating the loss of the external cavity.

Since the laser now has a composite cavity - the short diode cavity and the longer external cavity, the output spectrum is more complex than predicted by equation 2.5. Many devices have been fabricated since 1981 that produce pulse widths of less than 1 ps[98,99, 100], using a variety of different wavelengths, methods of mode-locking and cavity configurations. Such narrow pulses are not possible in other ways such as Q-switching or gain-modulation, although these narrow pulse widths are not necessary for this work.

2.4   Gain-switching of semiconductor lasers.

If the drive current to a semiconductor laser is driven with a pulsed current source (see figure 2.6), the optical output power will also be pulsed. However, the duration of the optical pulse can be very much shorter than that of the drive current. Since the repetition rate of the input current can be easily varied, so can the repetition rate of the laser pulses - something that can not be done so easily with mode-locked lasers. An advantage which gain-switching and Q-switching offer over mode-locking is the ease with which the repetition rate can be changed, although this is not necessary for this work.

2.5　　　　　Laser used for this project.

The laser system used during this Ph.D. changed slightly in form during the work. Originally a Spectra-Physics (Spectra-Physics Ltd, Boundary Way, Hemel Hempstead, Herts, HP2 7SH) model 2040-20 argon-ion laser developing about 9 W at 514.5 nm was used to pump a mode-locked Spectra Physics Tsunami titanium sapphire laser, which was tunable over the range 720 to 840 nm, although it

Figure 2.6 Typical time development of a gain-switched semiconductor laser diode.

was possible to go to shorter wavelengths by changing the mirror set. This laser produced pulses with a repetition frequency of 82 MHz (one pulse every 12.2 ns). This system was always very temperamental, and while 1 W average power could be obtained for a short while after the laser had been professionally aligned, the power and stability soon deteriorated. This was due to the pump beam wandering in position with temperature, which required constant adjustments on the Ti:sapphire laser to compensate for this. Although an active mode-locker was used, it was sometimes a black art to get the system mode-locked.

During 1998 the pump laser was changed from the argon ion laser to a Spectra-Physics Millennia V Advantage laser. This used high power laser diodes fibre coupled to pump a yttrium vanadate (NdYVO$_4$) crystal, which produced an output at 1064 nm. This was frequency doubled using a lithium triborate (LBO) crystal to produce 5 W of green light at 532 nm. The Millennia V Advantage pumped the original Ti:sapphire laser, which is tunable over the wavelength range 720 to 840 nm, although usually left at 800 nm. The newer pump laser had a reduced output of 5.0 W, so the output of the main Ti:sapphire was obviously lower (about 600 mW typical), but the stability was much improved. The slight change in wavelength of the pump laser (from the original 514.5 nm to the new 532 nm) resulted in some slight loss of efficiency using the original Ti:Sapphire crystal, although the extra stability more than compensated for the slight loss of output power. A photograph of the laser table in our laboratory is shown in figure 6.40, where the pump and main laser can be clearly seen.

The laser mode-locker had an output synchronised to the laser pulses, which was used to drive the pulse generator used in the system.
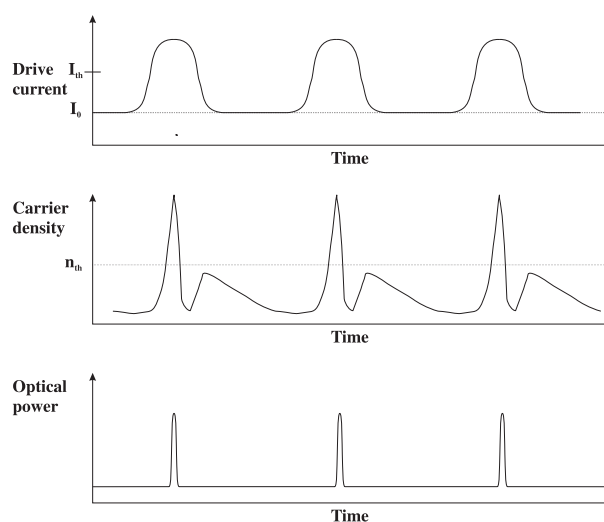
Chapter 3.

Fast Semiconductor Detectors.

3.1   PN/PIN photodiode.

A reverse biased pn diode can used be as an optical detector. If reversed biased, radiation with an energy greater than the band-gap ε_g of the material (0.66 eV for germanium, 1.12 eV for silicon and 1.424 eV for gallium arsenide), can supply sufficient energy to excite a valance band electron into the conduction band, causing a current to flow in an external circuit. There is a threshold wavelength λ_{th} for any detector which is given by:

$$\lambda_{th} = \frac{hc}{\varepsilon_g} \qquad\qquad \textbf{3.1}$$

where h is Plancks constant (6.62617×10^{-34} Js) and c is the velocity of light (3×10^8 m/s). For the three most common materials, the threshold wavelengths are: germanium 1.88 μm, silicon 1.11 μm and gallium arsenide 870 nm.

Most photodiodes are usually made of a pin structure, with a layer of intrinsic (or i-layer) material sandwiched between the heavily doped p and n type materials. This has the advantage over the pn structure, that the depletion width is fixed during the manufacture of the device and will remain constant. As long as the reverse bias is sufficient to cause the depletion layer to reach through to the underlying layer, the depletion width will be equal to the width of the intrinsic region.

Figure 3.1 shows the voltage current characteristics of a pin photodiode, both with and without optical radiation. The photodiode also has a small leakage current that flows when it is reversed biased. This is usually called the dark current I_D in photodiodes and is lowest in materials with the largest band gap. The extra current caused by the light is called the photocurrent I_{ph}. The magnitude of the photocurrent density J_{ph} (in Amps per square metre) is easily calculated, given the wavelength of light and the optical power density reaching the photodiode. Not every photon causes an electron to flow in the external circuit, due to a number of reasons:

1) The hole and electron may recombine in the intrinsic region.

2) Reflection at the front surface, means the photon never reaches the depletion region.

3) The photon is absorbed outside the i-region. The fraction η of photons that cause an electron to flow in the external circuit is called the quantum efficiency. The diode is usually reversed biased sufficiently to cause a wide depletion layer in the device, as hole-electron pairs created outside of the depletion layer will soon recombine, without producing a current in the external circuit.

3.1.1 Equivalent circuit.
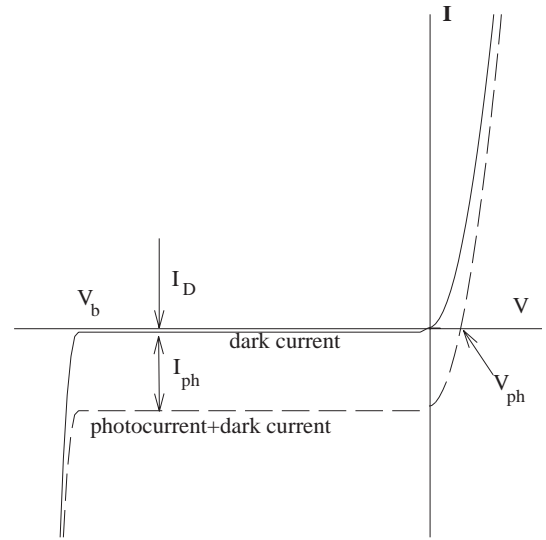
Figure 3.1 Voltage-current characterises of a pin junction photodiode. The solid light shows the dark current, similar to that for any pn junction diode. The dotted line shows the extra photocurrent, due to light.
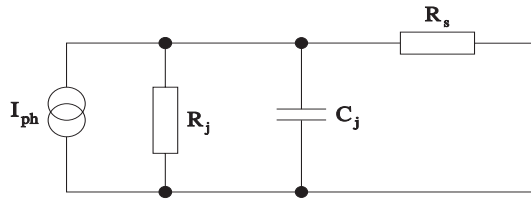
Figure 3.2 Diagram showing a simple equivalent circuit for a reversed biased photodiode. The thermal noise of the two resistors is not shown.

Figure 3.2 shows a simplified equivalent circuit of a reverse biased photodiode, where C_j is the junction capacitance and R_j is the junction resistance. The series resistance R_s is the resistance of the bulk semiconductor and the contact resistance and is usually less than 10 Ohms. I_{ph} is the photocurrent and will depend on the optical power falling on the photodiode.

3.1.2 Small signal response of a pin photodiode.

Let N_c be the number of carriers generated per square metre per second, P_d be the power density of the incoming radiation (Watts/m^2), λ the wavelength of the radiation and ν the frequency of the radiation ($\nu = c/\lambda$), then

$$N_c = \frac{\eta \; Power \; density}{energy \; per \; photon} = \frac{\eta P_d}{h\nu} = \frac{\eta P_d \lambda}{hc} \qquad \text{3.2}$$

The photocurrent density J_{ph} is:

$$J_{ph} = \frac{\eta P_d q \lambda}{hc} \qquad \text{3.3}$$

The photocurrent density J_{ph} (Amps/m^2) is simply N_c multiplied by q, the charge on an electron. Hence the photocurrent I_{ph} (in Amps) due to an optical power P_o (in Watts) is

$$I_{ph} = \frac{\eta P_0 \lambda q}{hc}$$
<div align="right">**3.4**</div>

The responsivity $\Re$ is defined as the ratio of photocurrent to optical power and is measured in Amps per Watt.

$$\Re = \frac{I_{ph}}{P_O} = \frac{\eta \lambda q}{hc}$$
<div align="right">**3.5**</div>

For a typical photodiode used in the near infra-red, the responsivity $\Re$ is approximately 0.5 A/W.

3.1.3   Large signal response of a pin photodiode.

Equation 3.4 showing the relation between photocurrent and optical power, assumes that there are sufficient carriers in the depletion layer to produce carriers for the incoming photons. If the number of photons in the active region is too great, then the current will start to saturate.

We can calculate an order of magnitude estimate of the maximum saturated current that a photodiode can supply. Assume a pin photodiode, with an active area of A (m^2), depletion width of w (m). Assume the doping density in the i-region to be N_d (m^{-3}) and the saturated velocity of the carriers to be v_s. Assuming that the carriers are generated uniformly throughout the i-region, the carriers will travel an average distance of w/2 before leaving the i-region. The average life of the carrier in the i region τ_c, will be:

$$\tau_c = \frac{w}{2v_s}$$
<div align="right">**3.6**</div>

The number of doping sites in the active region will be N_d A w, so the number of carriers that can be generated per second from the i-region will be:

$$\text{carriers per unit time} = \frac{N_d A w}{\frac{w}{2v_s}} = 2v_s N_d A$$
<div align="right">**3.7**</div>

Since each carrier will carry a charge q, the charge carried per unit time (current) will be:

$$I = 2v_s N_d A q \qquad\qquad \textbf{3.8}$$

This gives an order of magnitude estimate for the maximum current density J possible in a photodiode.

$$J_{ph} = 2v_s N_d q \qquad\qquad (Amps/m^2) \qquad\qquad \textbf{3.9}$$

Assuming typical values for the saturated carrier velocity and doping levels in a silicon photodiode, $v_s = 10^5$ m/s, $N_d = 10^{19}$ m^{-3} and q = 1.6x10^{-19} C, we can calculate a maximum current density of 3.2x10^5 A/m^2 = 320 mA/mm^2. Note the device will show considerable non-linearity well before the saturated current is reached and at least one photodiode manufacturer recommends that the current density does not exceed 50 µA/mm^2, before the responsivity drops by 1%.

3.1.4   Noise of pin photodiode.

Since the pin photodiode has no built in amplification, it is usually necessary to amplify the signal. This is often done by feeding the photodiode current through a load resistor to produce a voltage, which is subsequently amplified. The other commonly used method, is to use the photodiode as a current source with an operational amplifier configured as a current to voltage converter[101]. In order to study the noise properties of a photodiode it is necessary to consider the interaction between the photodiode and the amplifier as shown in figure 3.3.
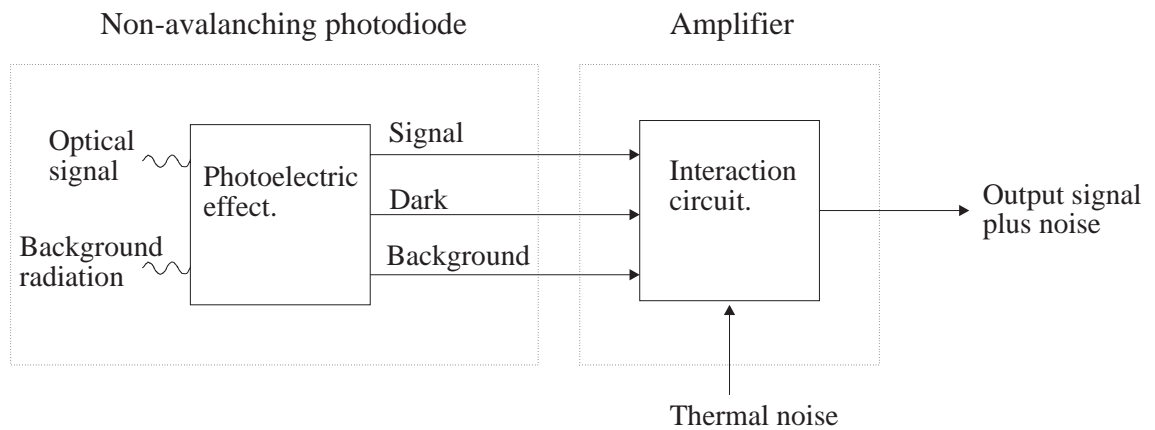
Figure 3.3 Diagram showing the photodetection process in a non-avalanching photodiode. Three current sources are assumed - the photocurrent, dark current and a current due to background radiation.

An equivalent circuit suitable for noise analysis, is more complex than shown earlier, as it requires the current sources due to dark current I_D and background radiation I_B be added. The

load resistance R_L and amplifier input resistance R_i are also required. The three current sources will all have shot noise associated with them so a current source for the total shot noise will be necessary. Since the resistors will have thermal noise, this must also be added. Figure 3.4 shows the equivalent circuit of the photodiode[102]. Here C_j, R_j and R_s are the equivalent components of the photodiode, R_L is the load resistor and R_i is the input resistance of the following amplifier. The input capacitance of the amplifier C_i is not necessary for noise analysis, but has been included for completeness.
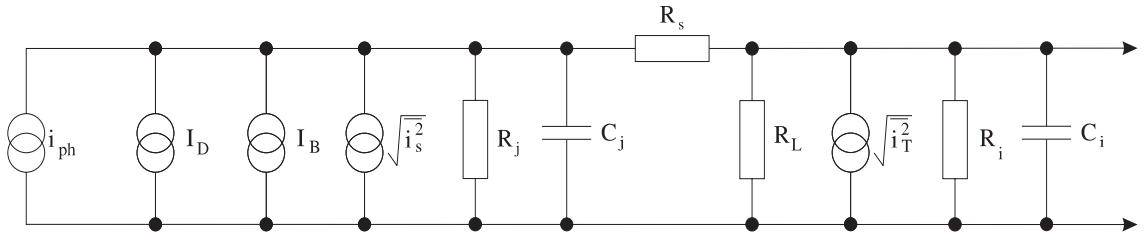
Figure 3.4 Equivalent circuit of a photodiode followed by an amplifier.

Assume an intensity modulated optical signal given by:

$$P(\omega) = P_0(1 + me^{j\omega t})$$

3.10

where m is the modulation index, P_0 is the average power and ω the modulation frequency. The average photocurrent due to the optical signal is:

$$I_{ph} = \frac{\eta q P_o \lambda}{hc}$$

3.11

For the modulated optical signal, the peak signal power is $m\,P_0$, so the rms power is $m\,P_0/\sqrt{2}$ and the rms signal current i_{ph} can be written as:

$$i_{ph} = \frac{\eta q m P_o \lambda}{hc\sqrt{2}}$$

3.12

Because of the random nature of the generation of the dark (I_D), background (I_B) and signal (I_{ph}) currents, they contribute shot noise, given by:

$$<i_s^2> = 2q(I_{ph} + I_B + I_D)B$$

3.13

where B is the bandwidth. The thermal noise current i_T is given by:

$$\langle i_T^2 \rangle = \frac{4kTB}{R_{eq}} \qquad\qquad \textbf{3.14}$$

where R_{eq} is:

$$\frac{1}{R_{eq}} = \frac{1}{R_j} + \frac{1}{R_L} + \frac{1}{R_i} \qquad\qquad \textbf{3.15}$$

k is Boltzmanns contant (1.38×10^{-23} J K^{-1}) and T is the absolute temperature. The thermal noise is also called *white noise* or *Johnson noise*. Thermal noise was first investigated experimentally by Johnson[103] and theoretically by Nyquist[104]. The signal to noise ratio can be written as:

$$(S/N)_{power} = \frac{i_{ph}^2 R_{eq}}{(\langle i_s^2 \rangle + \langle i_T^2 \rangle) R_{eq}} = \frac{\left[\dfrac{mq\eta P_O \lambda}{\sqrt{2}\,hc} \right]^2}{2q(I_{ph} + I_B + I_D)B + 4kTB/R_{eq}} \qquad\qquad \textbf{3.16}$$

This allows us to calculate the minimum optical power $P_{min}(S/N)$ necessary for any given signal to noise ratio.

$$P_{min}(S/N) = \frac{2hcB}{m\eta\lambda}(S/N)\left[1 + \sqrt{1 + \frac{I_{eq}}{qB(S/N)}} \right] \qquad\qquad \textbf{3.17}$$

where

$$I_{eq} = I_B + I_D + \frac{2kT}{qR_{eq}} \qquad\qquad \textbf{3.18}$$

In the limit when $I_{eq}/qB(S/N) \ll 1$, the minimum optical power necessary for a given signal to noise ratio is limited by the quantum noise associated with the optical radiation.

$$(P_O)_{min} = \frac{4hcB}{m\eta\lambda}(S/N) \qquad\qquad \textbf{3.19}$$

One parameter often used to indicate the noise performance of photodiodes is the noise equivalent power (NEP). This is the optical power required to detect a 100% modulated signal with a signal to noise ratio of unity in a bandwidth of 1 Hz.

$$NEP = \frac{4hc}{\lambda\eta} \qquad \textbf{3.20}$$

Unfortunately the minimum optical power required for a given signal to noise ratio is limited not by the quantum noise of the light, but by the noise of the following amplifier. In this case, $I_{eq}/qB(S/N)$ is usually much greater than 1, so the minimum optical power for a given signal to noise ratio is:

$$P_{min}(S/N) = \frac{2hc}{m\eta\lambda}\sqrt{\frac{I_{eq}B(S/N)}{q}} \qquad \textbf{3.21}$$

and the noise equivalent power is:

$$NEP = \frac{2hc}{\eta\lambda}\sqrt{\frac{I_{eq}}{q}} \qquad \textbf{3.22}$$

If the signal to noise ratio at the input of an amplifier $(S/N)_{in}$ is calculated as in equation 3.16, the signal to noise ratio at the output of a practical amplifier $(S/N)_{out}$ will always be lower due to the noise created internally by the amplifier. If the noise factor of the amplifier is F (where $1 \leq F \infty)^{[*1]}$, then the signal to noise ratio at the output of the amplifier, is given by:

$$(S/N)_{out} = \frac{(S/N)_{in}}{F} \qquad \textbf{3.23}$$

3.1.5   Frequency response of the pin photodiode.

The frequency response of photodiodes can be limited due to one of four reasons:

1) The RC circuit time constant of the photodiode and amplifier combination.

2) The carrier transit time across the device.

3) The carrier diffusion time, if electron-hole pairs are created outside the depletion layer.

4) Limited bandwidth of the active element in the amplifier following the photodiode.

These will be considered in turn.

[*1] The terms *noise factor* and *noise figure* are used interchangeably by some authors, whereas others use a convention that the *noise figure* is $10 \log_{10}$ (*noise factor*). Throughout this thesis, this convention is used.

3.1.5.1   Circuit time constant.

For calculating the bandwidth of a photodiode and amplifier combination, the equivalent circuit of figure 3.4 can be used. Since R_s is small, it can usually be ignored, so the equivalent circuit is simply the current source in parallel with the resistors and capacitors. The noise sources need not be considered. A simplified equivalent circuit is shown in figure 3.5. Since $R_s \ll R_A$, the photodiode can be simplified to a current source in parallel with a resistor R_{eq} given by equation 3.15 and capacitor C_j. Then the output voltage $V_{out}(f)$ will be given by

$$V_{out}(f) = \frac{I_{ph}(f)\, R_{eq}}{(1 + j2\pi f\,(C_j + C_{in})R_{eq})} \qquad \textbf{3.24}$$
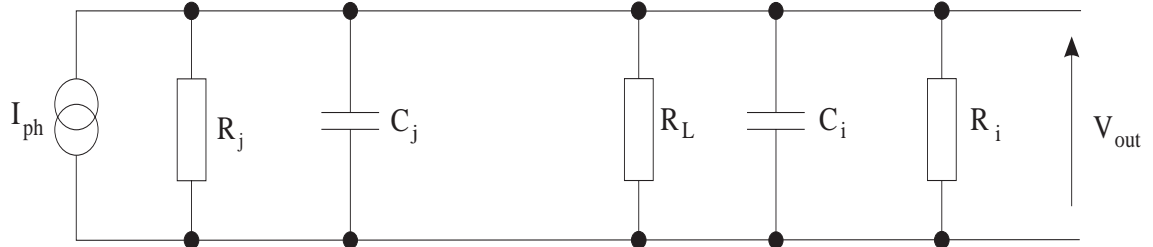
Figure 3.5 Diagram showing a simplified photodiode and amplifier combination, that can be used for analysing the high frequency performance.

Hence for good high frequency response, it is essential to reduce the photodiode capacitance, which usually means having a small area.

3.1.5.2   Carrier transit time of a photodiode.

Following absorption of a photon in the high field i-region and creation of a hole-electron pair, both carriers will soon (in fractions of a ps) acquire a constant velocity. At low fields this velocity is dependent on the field, but at higher fields, it saturates. In a correctly biased pin photodiode, the field should be sufficient to ensure all carriers move at saturated velocity, as this will give the fastest response. In this case, the response time is limited by the carrier transit time. If a pulse of light is incident on the p edge of the depletion layer of width w and all photons are absorbed there, then the time for the electrons to cross the depletion region is approximately

$$t_{tr} = \frac{w}{v_{se}} \qquad \textbf{3.25}$$

where v_{se} is the saturated velocity of electrons. If the carriers are absorbed anywhere other than

at the p edge of the depletion layer, then this equation must be modified to take into account the different electron and hole saturation velocities and the exact location where the photon was absorbed. In the worst case, if all photons were absorbed at the n edge (not very likely) then the transit time will be

$$t_{tr} = \frac{w}{v_{sh}}$$

3.26

where v_{sh} is the saturated velocity of holes. In this case, t_{tr} is longer since the saturated velocity of holes is slower than that of electrons. It is shown in Sze[102], that the 3 dB frequency is given by:

$$f_{3dB} = \frac{2.4}{2\pi t_{tr}} \approx \frac{0.4 v_s}{w}$$

3.27

where v_s is the average carrier velocity.

3.1.5.3   Limited bandwidth of the amplifier.

Unless the amplifier following the photodiode has a bandwidth much greater than that of the photodiode, then this will limit the frequency response. Very wide bandwidth amplifiers (many GHz) are now starting to become quite reasonably priced, although a few years ago they were still very expensive.

3.2   The avalanche photodiode (APD).

With a large reverse bias on a photodiode, the field in the depletion layer can be made to exceed a critical value (about 200 kV/cm in silicon) when avalanche multiplication can occur. Any free carriers which are created in the depletion layer, by for example the absorption of a photon, will be accelerated by the electric field. The kinetic energy they gain under the force of the electric field can *kick* an electron from the valance band to the conduction band. The hole and electron created can go on to repeat this process, so creating an avalanche of carriers and hence current. The process of avalanche multiplication in a semiconductor was first reported by McKay et al[105,106].

This is shown in schematically for a pin avalanche photodiode in figure 3.6. The incoming photon of energy hv excites an electron from the valance band to the conduction band at position A. The two carriers produced move in opposite directions under the influence of the applied electric field. The hole produced is accelerated to the left, but is not shown taking part

in any interactions. The free electron produced is accelerated to the right, until at position B, it has gained sufficient energy to excite an electron from the valance band to the conduction band. This creates another electron-hole pair. Both the original electron and the newly created one are in the conduction band and move to the right under the force of the electric field in the depletion layer. The hole created by the interaction at position B moves to position C, where it too creates another electron-hole pair. The hole moves to the left, but the electron moves to D, where it creates another electron-hole pair. Note that in the diagram the electron-hole

Figure 3.6 Diagram showing avalanche gain in a pin avalanche photodiode. The gain is 6 and the *K* factor (α_h/α_e) is 0.66.

pairs are created by holes at positions C and F, but by electrons at positions B, D and E. For this diagram the *K* factor (α_h/α_e is 2/3=0.66). The field strength required for avalanche multiplication is such that the carriers will usually move at saturated velocity, except for a short time (sub ps) following a collision.

3.2.1   Device configurations.

Although the pin structure will work for an avalanche photodiode, in practice avalanche photodiodes are usually made using more complex structures. According to Anderson et al[107], the avalanche multiplication should be spatially uniform across the entire light sensitive area of the diode. Small areas of the device where the breakdown voltage is less than the junction as a whole, are known as microplasmas and should be eliminated. The probability of microplasmas is reduced by making the high-field avalanche area (i-region) as small as possible. However, if the i-region is too thin, insufficient radiation will be absorbed there, making the quantum efficiency of the device poor. If the area of the device is reduced, the collection area is reduced. One solution is to make a separate absorption and avalanche region, as shown in figure 3.7. The absorption layer of width w_i is thick (say 100 µm) but the avalanche region of width w_a is thin (say 0.5 µm). Although manufacturers data was unavailable on the physical construction of

the avalanche photodiodes used in this project, it is suspected that they are not simple pin structures. The typical breakdown voltage of 150 V suggests an i-region thickness of 5 μm, but this would be too thin to serve as a useful absorption layer.

To reduce leakage (dark) current, which as shown in equation 3.16 will deteriorate noise performance, guard rings are often employed.

3.2.2 Gain of an avalanche photodiode.

If we denote the electron ionisation rate, defined as the number of ionising collisions per unit length by α_e, the hole ionisation coefficient by α_h and the width of the depletion layer by w then

Figure 3.7 Diagram showing construction and electric field profile of a typical silicon reach through avalanche photodiode (RAPD). Note the separate absorption and avalanche regions.

the gain M in a pin avalanche photodiode can be shown[102] to be:

$$M = \frac{(1-K)e^{\alpha_e w(1-K)}}{1 - Ke^{\alpha_e w(1-K)}} \qquad \textbf{3.28}$$

where K is defined as

$$K = \frac{\alpha_h}{\alpha_e} \qquad \textbf{3.29}$$

For most semiconductor materials, the electron and hole ionisation rates are different, but in GaP, they are identical, which leads to a much simpler expression for gain.

$$M = \frac{1}{1 - \alpha_e w} = \frac{1}{1 - \alpha_h w} \qquad \textbf{3.30}$$

The ionisation coefficients are strongly dependent on the electric field within the semiconductor. Figure 3.8 shows the variation of ionisation rates versus electric field for two common

semiconductors used in APD's.
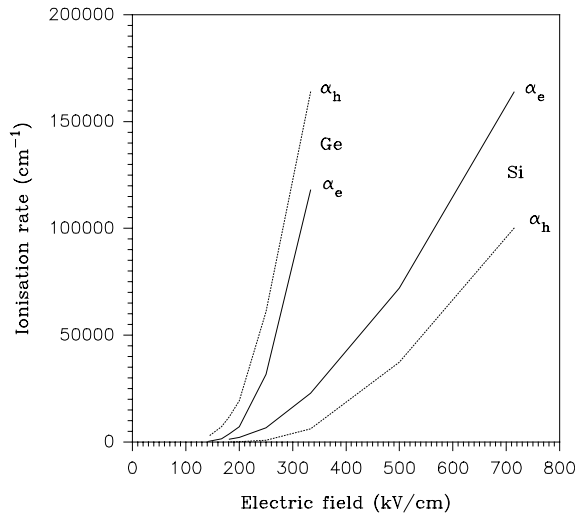
Figure 3.8 Ionisation rates of both carriers in silicon (Si) and germanium (Ge). Data taken from Sze[102].
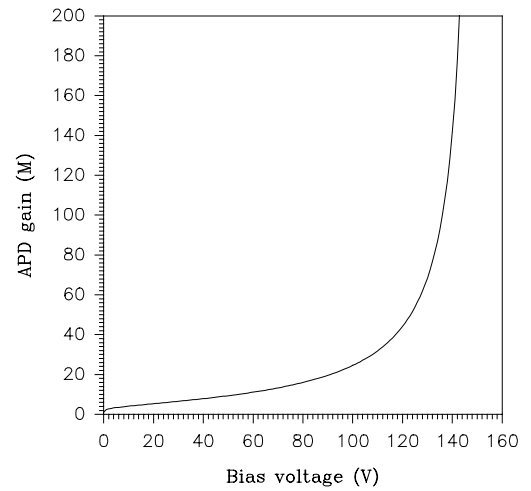
Figure 3.9 Graph showing the variation of gain of a typical APD with bias voltage.

This explains the very strong dependence of gain of an avalanche photodiode with voltage. Figure 3.9 is a typical voltage versus gain characteristic for a silicon avalanche photodiode.

In practice, the gains of avalanche photodiodes are difficult to predict accurately from purely theoretical methods, due to the presence of traps in the semiconductor. An empirical relationship has been developed that models the variation of gain with voltage for an avalanche photodiode[108].

$$M = \frac{1}{\left[1 - \left[\dfrac{V_R - I R_s}{V_B}\right]^{\gamma}\right]}$$

3.31

where V_R is the applied reverse voltage to the diode, V_B is the breakdown voltage of the diode, I is the diode current, R_s is the series resistance of the diode and the exponent γ varies with the semiconductor material, temperature and wavelength. APD manufacturers usually give both the breakdown voltage and the voltage necessary for some fixed gain, typically 100, although for germanium devices its usually much less (10 or so). There are available some high gain APD's, running at very high voltages. Assuming we know the breakdown voltage V_B, the current flow is negligible, and we know the voltage necessary for some gain M, say V_M, then the exponent γ is easily calculated as:

$$\gamma = \frac{\log_e((M-1)/M)}{\log_e(V_M/V_B)} \qquad \textbf{3.32}$$

3.2.3 Equivalent circuit for the APD.

As for the pin diode, C_j is the capacitance of the diode, R_s the series resistance and R_j the junction resistance.

3.2.4 Small signal response.

The calculation of the photocurrent for the APD, follows exactly the same method as for a normal pn or pin diode, except that the current is increased by the gain M of the avalanche device. Hence the photocurrent is given by:

$$I_{ph} = \frac{M\eta P_0\lambda q}{hc} \qquad \textbf{3.33}$$

The responsivity $\Re$ is defined in the same way as for the pin diode, ignoring any gain. Hence the photocurrent of an APD at a gain M is:

$$I_{ph} = M\Re P_0 \qquad \textbf{3.34}$$

3.2.5 Noise in the APD.
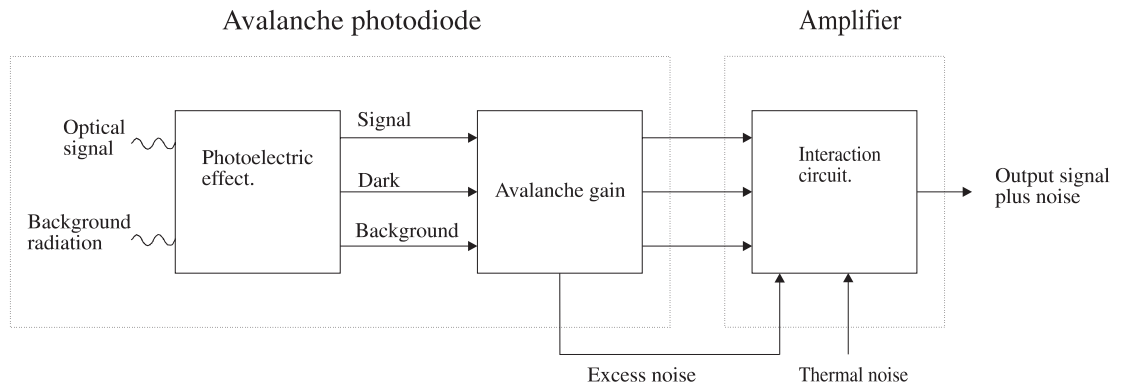
Figure 3.10 Diagram showing the photodetection process of an APD.

The noise analysis of an APD is slightly more complex than for a normal pin diode, as

the multiplication process itself introduces some noise. Analysis follows the same procedure as for the pin diode in section 3.1.4. This noise on the multiplication process is taken into account by a use of a larger shot noise current source than in the pin photodiode. The rms signal photocurrent i_{ph} is identical to that for the pin diode (equation 3.12) except that it is multiplied by the internal gain mechanism:

$$i_{ph} = \frac{M\eta qm\lambda P_O}{hc\sqrt{2}} \qquad \textbf{3.35}$$

where M is the APD gain, η is the quantum efficiency, m is the modulation index of the intensity modulated light source, λ the wavelength (m), P_0 is the average power of the light source (W), h is Planks constant (6.62617×10^{-34} Js) and c is the velocity of light in vacuum (3×10^8 m/s). The mean-square shot noise current after multiplication is given by

$$<i_s^2> = 2q(I_{ph} + I_B + I_D)<M^2>B \qquad \textbf{3.36}$$

or

$$<i_s^2> = 2q(I_{ph} + I_B + I_D)M^2F(M)B \qquad \textbf{3.37}$$

where $<M^2>$ is the mean-square value of the internal gain and the noise factor $F(M) = <M^2>/M^2$ is a measure of the increase in shot noise compared to an ideal noiseless multiplier.

As with the pin photodiode, the resistors R_s, R_{in}, R_L, R_j can be combined into a single equivalent resistor R_{eq}. If we again assume $R_s \ll R_j$, then R_{eq} is given by equation 3.15 and the thermal noise current is given by equation 3.14.

The signal to noise ratio is calculated as follows:

$$(S/N)_{power} = \frac{i_{ph}^2 R_{eq}}{(<i_s^2> + <i_T^2>)R_{eq}} = \frac{\frac{1}{2}\left(\frac{mq\eta\lambda P_O}{hc}\right)^2 M^2}{2q(I_P + I_B + I_D)F(M)M^2B + 4kTB/R_{eq}M^2} \qquad \textbf{3.38}$$

A comparison of the above equation with that for the pin diode (equation 3.16), shows that the internal multiplication in the APD, increases the signal to noise ratio, by reducing the effect of the noise in the amplifier following the detector. Since F(M) is always greater than unity in an APD, the multiplication process introduces noise itself. This noise increases with increasing gain. Hence there is an optimum value of gain in an APD, above which the signal to noise ratio deteriorates. Usually an APD is not operated above this gain, although for this project, the APDs

were operated at higher voltages. The reference APD (see figure 6.3) was operated above this to achieve a large output pulse, which was necessary. In this case there was lots of light, so the the reduction in signal to noise ratio in this instance was of no consequence. The correlating APD (see figure 6.3) was operated above this voltage to achieve a large gain modulation.

Equation 3.38 can be solved for the minimum optical power required to produce a given signal to noise ratio. The power is

$$(P_O)_{min} = \frac{2hcB}{m\eta\lambda}(S/N)\left[1 + \sqrt{1 + \frac{I_{eq}}{qBF(M)^2(S/N)}}\right] \qquad \textbf{3.39}$$

where

$$I_{eq} = (I_B + I_D)F(M) + \frac{2kT}{qR_{eq}M^2} \qquad \textbf{3.40}$$

If $I_{eq}/qBF(M)^2(S/N) \ll 1$, the minimum detectable power is the same as for the pin photodiode, which shows no advantage in the use of the APD.

$$(P_O)_{min} = \frac{4hcB}{m\eta\lambda}(S/N) \qquad \textbf{3.41}$$

However, in wide bandwidth optical detection, the minimum detectable power is sometimes limited by the thermal noise of the load resistance and the noise figure of the following amplifier. In this case $I_{eq}/qBF(M)^2(S/N) \gg 1$ and minimum optical power necessary for a given (S/N) ratio is:

$$(P_O)_{min} = \frac{2h\nu}{m\eta}\sqrt{\frac{I_{eq}B(S/N)}{q}} \qquad \textbf{3.42}$$

For the case when the detection sensitivity is limited by the thermal noise of the load resistance and the noise figure of the following amplifier, the expression for the noise equivalent power simplifies to:

$$NEP = \frac{2h\nu}{\eta}\sqrt{\frac{(I_B + I_D)F(M) + 2kT/qR_{eq}M^2}{q}} \qquad \textbf{3.43}$$

3.2.6 Avalanche multiplication noise.

We have shown that the signal to noise ratio of an APD and amplifier depends on the gain and the APD noise factor $F(M)$, which was said to increase with gain. We will now consider the noise factor in greater detail.

McIntyre[109] showed that for optimum noise performance the injected current should be from one type of carrier only (ie. holes or electrons but not both) and ideally it would be from the carrier with the highest ionisation coefficient (holes for germanium, electrons for silicon). Hence silicon APD's should be constructed as in figure 3.11. When electron-hole pairs are created in the i-region, the holes move to the right but do not take part in the avalanche. However, electrons move to the left into the avalanche region, so starting the avalanche.
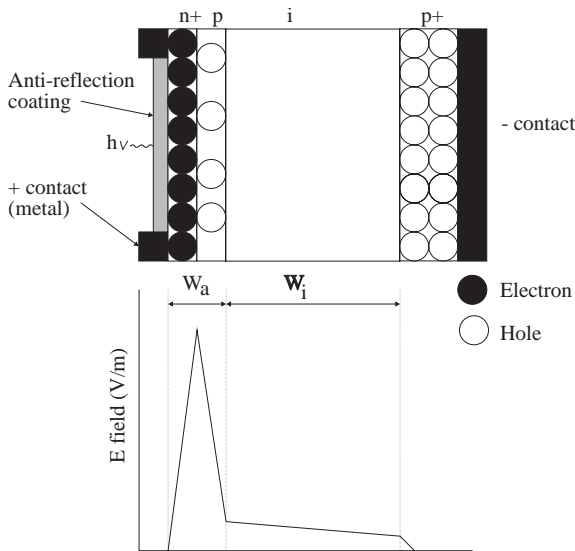
Figure 3.11 Diagram showing how a typical silicon APD might be constructed. Free electrons created in the i-region will move under the E-field to the avalanche region.
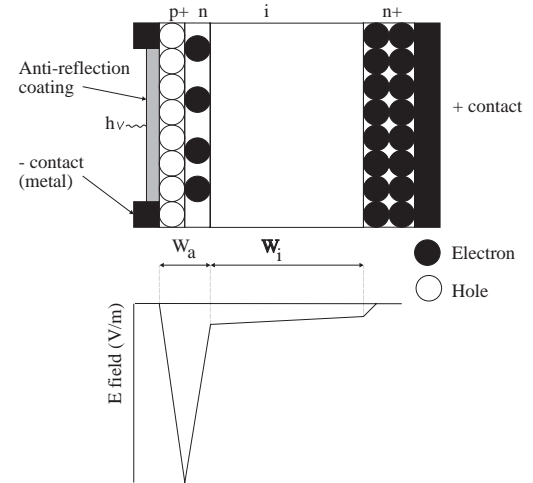
Figure 3.12 Diagram showing how a typical germanium APD might be constructed. Note that unlike silicon APD's, the device is constructed so holes are more likely to start an avalanche.

The converse is true for germanium APD's as shown in figure 3.12. Here the free electrons created in the i-region move to the right and so do not take part in the avalanche process. However free holes generated in the i-region move to the left and start avalanche multiplication in the avalanche region.

McIntyres proof will not be reproduced, but one can easily see that this is true without resort to an extensive mathematical treatment. The gain M of the device is obviously related to the number of carriers produced in the avalanche region, for each incident photon. Since the avalanche process is a statistical one, there will always be fluctuations in the number of carriers present in the avalanche region and hence some fluctuation of gain, even if photons arrived at

a constant rate. If $\alpha_e = \alpha_h$ and the diode is illuminated evenly across the whole of the depletion width, then there will on average be only three carriers present in the multiplying region; the primary and its secondary hole and electron. A fluctuation of one electron or hole then represents a significant perturbation from the mean number and hence high multiplication noise. On the other hand, if one of the coefficients approaches zero, far more carriers will be present in the multiplying region for every incident photon and so a fluctuation of one will be a very small perturbation. The APD noise factor *F(M)*, which is ratio of the mean-square value of gain to the mean-squared will therefore be less.

For electron injection alone, the noise factor F(M) can be written as

$$F(M) = KM + (2 - \frac{1}{M})(1 - K) \qquad \textbf{3.44}$$

where K is defined in equation 3.29. If $\alpha_e = \alpha_h$, then $K = 1$ and the noise factor *F(M)* equals *M*, which is very poor. If K is zero (ie $\alpha_h = 0$), then *F(M)* tends to 2 at large *M*. We can see that there is a dramatic reduction in noise factor if hole ionisation coefficient is zero with electron injection.

The expression for *F(M)* for hole injection is:

$$F(M) = K'M + (2 - \frac{1}{M})(1 - K') \qquad \textbf{3.45}$$

where K' is defined as the inverse of K, ie

$$K' = \frac{1}{K} = \frac{\alpha_e}{\alpha_h} \qquad \textbf{3.46}$$

Figure 3.13 shows theoretical curves of noise factor *F(M)* against multiplication *M*, assuming electron injection, for a number of different values of *K*. To find data for hole injection, simply substitute the value of *K'*, for *K* on the graphs. Frequently for system evaluation, the simple approximation

$$F \approx M^{\chi} \qquad \textbf{3.47}$$

given in equation 3.47 is often used. Typically χ takes on values between 0.2 and 1, and is called the *excess noise index*. For the diodes used here, the 200 µm diameter Hamamatsu S2381, the excess noise index is said to be between 0.30 and 0.42. A plot of $F = M^{0.36}$ (the mean of 0.3 and 0.42) is also shown on figure 3.13.

For a semiconductor, the ionisation rates rise rapidly with electric field, as was shown on figure 3.8. The graphs of ionisation coefficient were taken from Sze[102], although the original data comes from different sources. For some materials such as gallium arsenide and germanium, the ionisation rate of holes is greatest, the current injection should be due to predominantly holes.
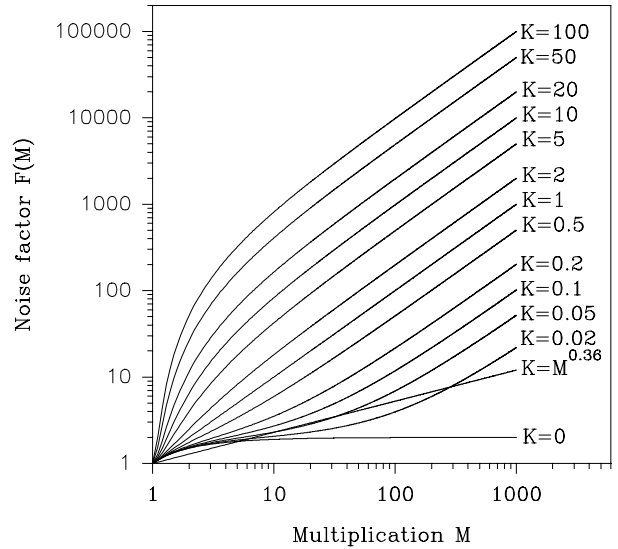
Figure 3.13 Theoretical curves relating the noise factor F(M) to the multiplication M, for various values of K. These curves assume electron injection. The straight line is the equation $F(M)=M^{0.36}$.

3.2.7   Frequency response of an APD.

The gain mechanism in an APD has been shown to increase the signal to noise ratio. However it also decreases the bandwidth of the APD, compared to a non-avalanching device. For reasons stated in section 3.2.1, APDs are usually not simple pin structures. Consider the p^+-i-p-n^+ shown in figure 3.11. The regenerative avalanche process results in carriers being present in the high-field region, long after the primary electrons have traversed that region. Unlike the pin diode, where the response time is dependent only on the carrier transit time, there are three components to the response time of an APD.

1) The electron transit time t_{re} across the absorption region

$$t_{re} = \frac{w}{v_{se}}$$

3.48

2) The time taken for the avalanche to develop, t_a, which depends on K^{70}. If $K=0$, the avalanche develops within the electron transit time, whereas for $K>0$, it requires multiple passes. At high M, with $0<K<1$,

$$t_a = \frac{MKw_a}{v_{se}}$$

3.49

3) The transit time of the last holes produced across the absorption region t_{rh}

$$t_{rh} = \frac{w_a + w_i}{v_{sh}}$$ **3.50**

Hence the overall response time t_r is given by:

$$
\begin{aligned}
t_r &= t_a + t_{re} + t_{rh} \\
&= \frac{MKw_a}{v_{se}} + \frac{w_i}{v_{se}} + \frac{w_a + w_i}{v_{sh}} \\
&= \frac{(MKw_a + w_i)}{v_{se}} + \frac{(w_a + w_i)}{v_{sh}}
\end{aligned}
$$ **3.51**

The 3 dB bandwidth of the photodiode f_{3dB} is given approximately by[70]:

$$f_{3dB} \approx \frac{0.44}{t_r}$$ **3.52**

If we calculate the response time of two similar photodiodes, one non-avalanching, the other an avalanche photodiode, we can see the effect on the frequency response. The example in table 3.1 is taken from Gower[70].

Table 3.1 Comparison of the frequecny response of a pin photodiode and an avalanche photodiode

Parameter	PIN diode	p^+-pi-n-p^+ APD
Saturated electron velocity v_{se}	10^5 ms^{-1}	10^5 ms^{-1}
Saturated hole velocity v_{sh}	5×10^4 ms^{-1}	5×10^4 ms^{-1}
K	-	0.1
Width of absorbing region w_i	50 μm	50 μm
Width of avalanche region w_a	-	0.5 μm
Gain M of device	1	100
Electron transit time t_{re}.	500 ps	500 ps

Time for avalanche to develop t_a.	-	50 ps
Transit time of last holes t_{rh}.	-	1.002 ns
Total transit time t_r	500 ps	1.552 ns
3 dB bandwidth.	880 MHz	284 MHz.

It has been shown, that an APD having similar characteristics to a non-avalanching photodiode, has achieved a gain of 100, by sacrificing the bandwidth by a factor of approximately 3. The recent availability of very wide bandwidth low noise amplifiers at reasonable price, has somewhat reduced the commercial interest in fast APDs.

3.2.8   Gain-modulation of an APD.
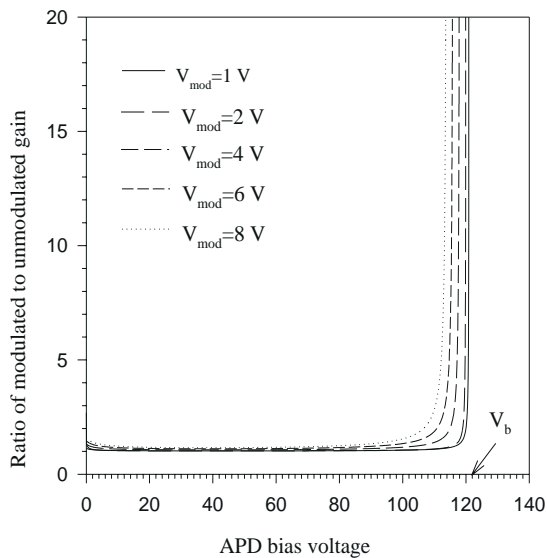
Figure 3.14 This series of curves shows the ratio of the gain of an APD, when biased with a modulation voltage in addition to the normal DC bias, to when just biased with DC.
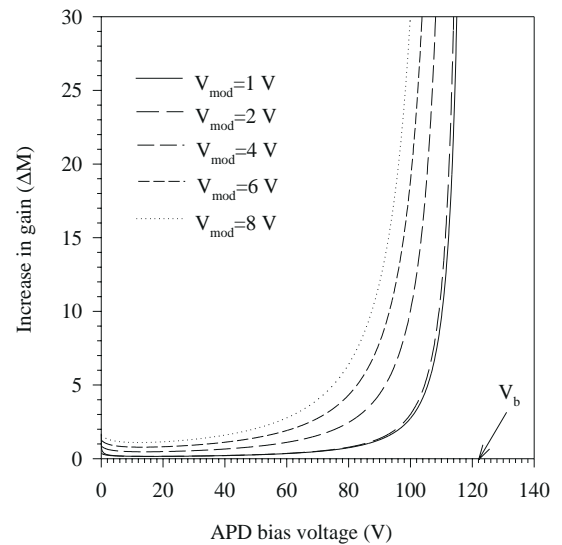
Figure 3.15 Graphs showing the change in gain as a function of bias voltage for 5 different modulation voltages.

Since the bias voltage of an APD controls the gain, modulation of the bias voltage can be used to modulate the gain. Figure 3.14 shows the ratio of the gain of an avalanche photodiode when biased with a pulse of height V_{mod} in addition to the DC bias, to that with just the DC bias alone. A number of curves are shown, for different modulation voltages. The results for this graph, were obtained from an approximate gain characteristic of an APD given by equation 3.31

where $V_B=122$ V, $\gamma=0.0968$, which closely approximates the gain characteristic of one of the 200 μm diameter Hamamatsu S2381 APDs used in this project. It is apparent that to achieve a large on/off gain ratio, it is necessary to use as large a modulation voltage as possible. If the modulation voltage is V_{mod}, then the diode must be biased with a DC voltage of $\leq V_b-V_{mod}$, as otherwise the modulation will cause the device to break down. Figure 3.15 shows the increase in gain ΔM, that the various modulation voltages theoretically give rise to.

It will be shown later how a gain-modulated APD forms a very important part of the instrumentation developed.

3.3   Photon counting APD's.

APDs are available that are designed for single photon counting applications, like that discussed in section 1.7.5. APDs designed for single photon counting applications are called single photon avalanche photodiodes (SPADs). The bias voltage V_A of a SPAD is held just above the breakdown voltage V_B, by an amount called the excess bias voltage V_E, where $V_E=V_A-V_B$. The excess bias voltage is around 10% of the bias voltage. At the bias voltage V_A, the electric field is so high that a single photon that creates a photoelectron, will create an avalanche, after which the current rises quickly (nanosecond or sub nanosecond) to a steady value of a few milliamps. This current pulse is then used to determine the arrive of a photon. If the bias voltage of the SPAD was held constantly above the breakdown voltage the dissipation would probably destroy the device, and in any case, it would be of little use, since the system would not be able to detect any further photons, as it would be avalanching. Photons arriving while the device is avalanching would not be detected. Hence the avalanche must be quenched by reducing the bias voltage to a quenching voltage V_q, sufficient to stop the avalanche, then increasing it again. This stops the avalanche and so allows the SPAD to detect further photons. The voltage necessary to quench the avalanche is slightly in excess of the breakdown voltage.

The excess bias voltage has a number of effects on performance, so it may be changed to optimise a particular parameter. The effects of the excess bias voltage are:

1) Increasing the excess bias voltage improves the detection efficiency, as a photon is more likely to create an avalanche if the electric field is high.

2) The resolution improves at higher electric field.

3) The dark count, which is the effect of thermally generated carries, also increases at higher bias voltages, since the probability of a thermally generated carrier producing an avalanche increases.

3.3.1   Passive quenching circuits (PIQ's) for SPAD's.

The simplest method to quench the avalanche is to make use of a fairly high value (few hundred kΩ) resistor in series with the APD bias. The increase in current at avalanche results in a sufficiently high voltage drop to bring the bias voltage below breakdown. Unfortunately, this does not work too well, with there being a number of after-pulses. These mean the time before further photons may be detected with reasonable probability is fairly long, meaning count rates of less than about 10,000 - 50,000 counts per second are necessary.

3.3.2   Active quenching circuits (AIQ's) for SPAD's.

The effect of after pulses can be considerably reduced, and the count rate considerably increased, by the use of an active, rather than a passive quenching circuit. The active circuits typically measure the current, and drop the voltage below breakdown if the current rises. Active quenching circuits may be found in several publications of Professor Cova[110,111,112] and on the world wide web at **http://www.ioi.co.uk/tech/dera/p0677.htm**

Chapter 4.

Variable Time Delay Mechanisms.

In order to measure the TPSF, using the cross-correlation method proposed, we must generate a variable delay. Methods of achieving this will now be considered. There are two types of time delays available - those that are continuously variable and those that introduce a fixed step. Ideally we require a continuously variable delay, but a number of fixed steps with sufficient resolution would also be suitable. Since the pulse repetition period of the laser is 12.2 ns (82 MHz) it is necessary to have a delay variable up to 12.2 ns to ensure the pulse generator modulates the APD at the same time as a signal is received on the APD. However, only a small fraction of this, set by the maximum TPSF duration, needs to be continuously variable or variable in small steps - the rest can use fairly large steps. Assuming the maximum TPSF duration that is expected is 2 ns, a continuously variable delay of 2 ns would be acceptable with a coarsely variable delay of the remaining 10.2 ns. This would allow any desired 2 ns section to be analysed with continuous resolution, without requiring a long continuously variable delay.

4.1 Continuously variable time delay mechanisms.

There are several continuously variable time delay mechanisms, which will now be considered.

4.1.1 Relative movement between source and detector.

If a variable time delay is required in an optical system, the simplest method of achieving this is shown in figure 4.1,

hν /\/\/\ ⬅————————— w —————————➤hν /\/\/\
Incoming   Exiting
light   light
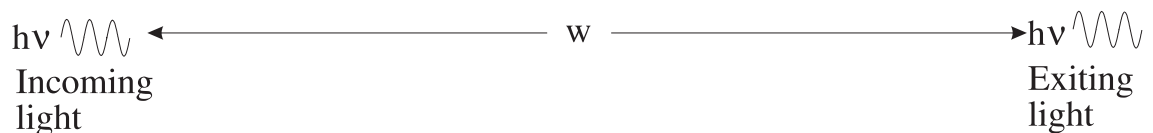
Figure 4.1 Diagram showing how varying the space between the incident and detected light, alters the time delay.

By changing the separation w between the light source and the detector, a variable delay can be obtained. The time delay τ of the signal is simply given by:

$$\tau = \frac{w}{c} \qquad\qquad \textbf{4.1}$$

A variable delay of 12.2 ns would require a separation between source and detector that is variable over at least 3.7 m, but this could be reduced to 600 mm if we accept a smaller (2 ns) continuously variable delay and add a course delay in large steps.

If the phase shifted signal was required in an electrical, as opposed to optical form, then it will be necessary to convert the electrical signal to light (with an LED or laser), then convert
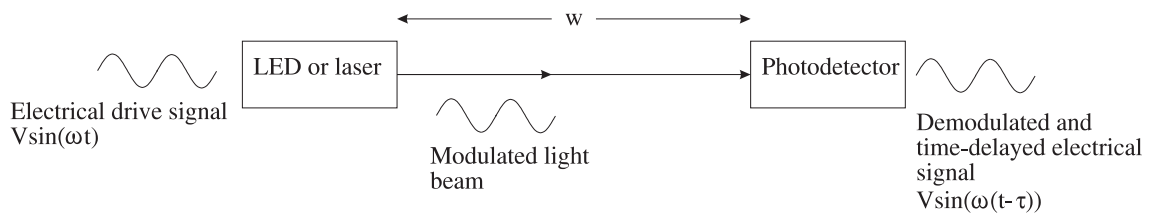
Figure 4.2 Method of introducing a time delay in an electrical circuit, by modulating a light source, then varying the distance between the source and a detector, which demodulates the light.

back to electrical with a photodetector of some variety, as shown in figure 4.2. Even if we introduce two delays, one continuously variable and one variable in steps, a variable distance of more than 600 mm may be inconveniently long, so the path may be folded, by use of a corner-
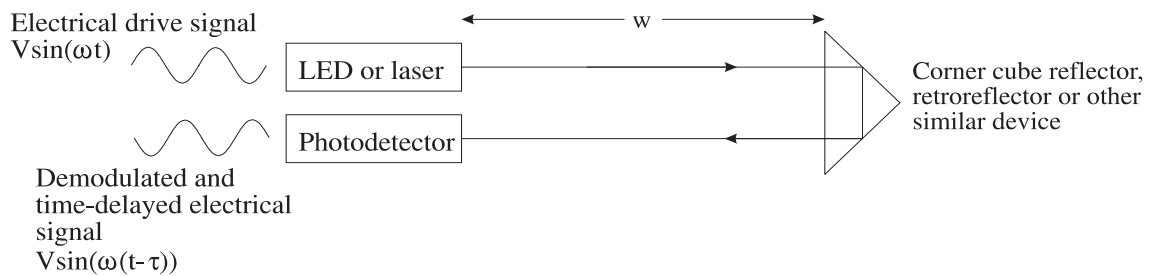
Figure 4.3 Method of decreasing the physical dimensions of a time-delay mechanism, by folding the optical path.

cube reflector or retro-reflector as shown in figure 4.3. The path may be folded a number of times to reduce further the physical size of the delay line as shown in figure 4.4, but as the number of reflections is increased, the difficulties of alignment also increase. The advantages of this method include:

1) Simple in principle and design.

2) True time delay.
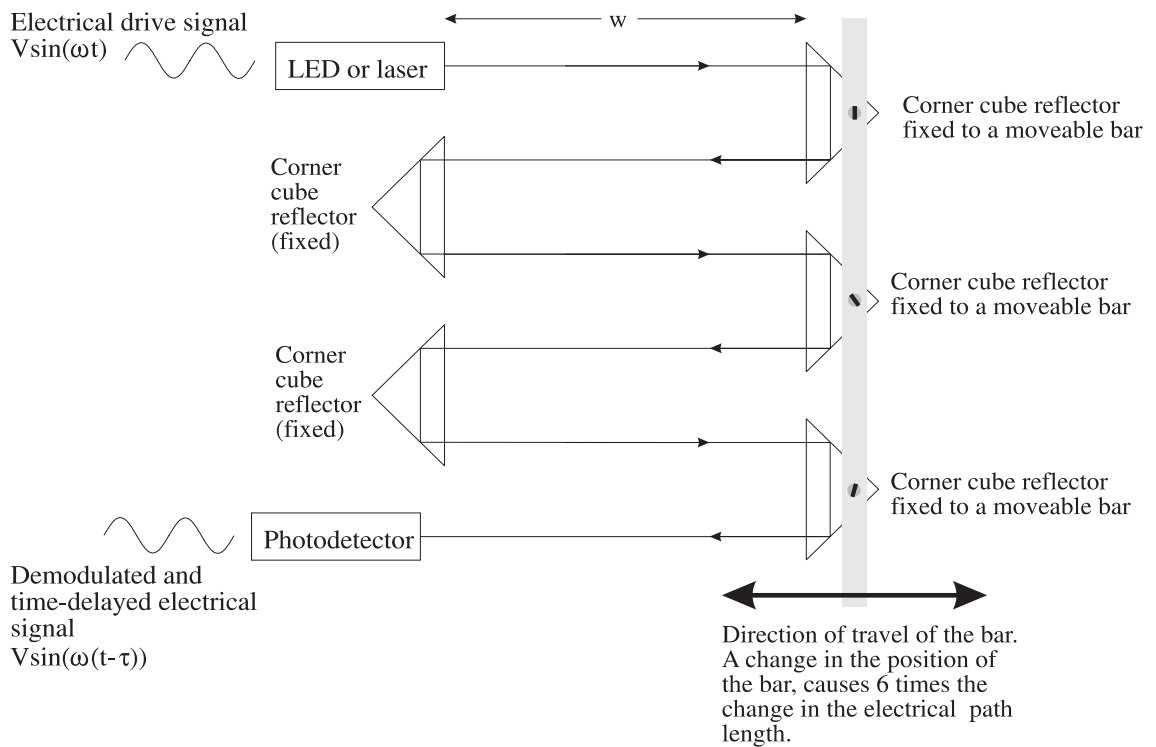
Figure 4.4 Folding the optical path a number of times decreases the overall width w, but leads to increased alignment problems, especially for a portable instrument.

The disadvantages include:

1) As the light beam diverges, the optical power density reaching the photodetector decreases as the phase shift increases. The area of the detecting photodiode must therefore be larger than the beam at all times. This limits the use at high modulation rates, due to high capacitance of large area photodiodes.

2) Mechanical difficulties of maintaining accurate alignment at all positions of the movable element.

3) Reliability is not as good in principle as an electronic phase shifter with no moving parts.

4) Speed of response is slower than with an all electronic phase shifter.

4.1.2 Trombone.

It is possible to use a section of transmission line in a trombone. A section of transmission line can be changed in physical length as needed, by sliding one section in and out of the other. This changes the electrical length and hence the phase of the signal. This is almost an electrical equivalent of the optical phase delay system discussed in section 4.1.1, but for long delay times is similarly bulky.

4.2   Discretely variable time delay mechanisms.

There are several discretely variable time delay mechanisms, which will now be considered.

4.2.1   Switched line phase shifter.

Another conceptually simple phase shifter uses a double pole double throw (DPDT) switch to route an electrical signal along one of two sections of transmission line of length l_1 and l_2. The phases difference between the two paths is then

$$\Delta\phi = \beta(l_2 - l_1) \qquad\qquad \textbf{4.2}$$

where β is the propagation constant at the operating frequency. The propagation constant in a lossless transmission line is given by:

$$\beta = \frac{2\pi}{\lambda} \qquad\qquad \textbf{4.3}$$

where λ is the wavelength of the signal in the cable, which is always longer than the free-space wavelength, as the velocity of propagation along a transmission line s, is less than the velocity of propagation in free space. Most cable manufacturers quote the velocity factor of the cable, which is defined as:

$$\text{velocity factor} = \frac{s}{c} \qquad\qquad \textbf{4.4}$$

Hence

$$\Delta\phi = \beta(l_2 - l_1) = \frac{2\pi(l_2 - l_1)}{\lambda} = \frac{2\pi f(l_2 - l_1)\,\text{velocity factor}}{c} \qquad\qquad \textbf{4.5}$$

The form of the switch might be a pin diode switch or coaxial relay.

The advantages of the switched line phase shifter include:

1) Simple in both principle and design.

2) Only practical means of achieving true time-delay operation

The disadvantages include:

1) Lossy, as switching diodes are in signal path.

2) Because of the unequal transmission line paths, the loss depends on the phase shift. This will be negligible if the cable loss is small compared to the switch loss. This results in phase

modulation to amplitude modulation conversion.

3) At higher microwave frequencies, the loss in the switch may make the device impractical.

4) 4 pin diodes are needed per section, as compared to two for some other types of phase shifters.

Digital phase shifters giving any required degree of precision can in principle be made as shown in figure 4.5. If the number of stages is n for a phase shifter giving up to ϕ_{max} degrees of phase shift, then the resolution $\Delta\phi$ is given by:

$$\Delta\phi = \frac{\phi_{max}}{2^n}$$
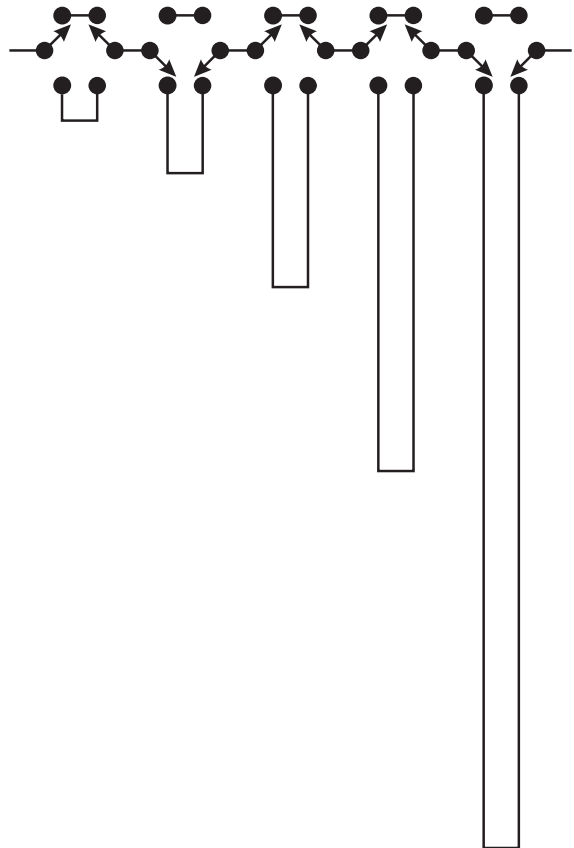
4.2.2 Reflection phase shifter.
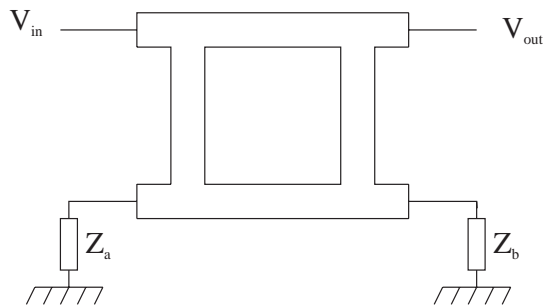
Figure 4.6 Reflection phase shifter - also known as a hybrid coupler phase shifter, since it uses a 3 dB quadrature hybrid coupler.

Figure 4.5 Digital phase shifter consisting of lengths of cable that may be switched in or out to change the phase delay. The switches might typically be implemented using pin diodes.

The reflection phase shifter[113] shown in figure 4.6 uses a 3 dB quadrature hybrid. The phase shift that the signal receives depends on the terminating impedances Z_a and Z_b. The output voltage V_{out} is then:

$$V_{out} = \frac{jV_{in}}{2}(\Gamma_a + \Gamma_b)$$

4.7

where Γ_a is the reflection coefficient of Z_a and Γ_b is the reflection coefficient of Z_b. By selecting different lengths of short or open circuited line for Z_a and Z_b, any phase shift can be achieved.

4.2.3   High-pass/low-pass phase shifter.
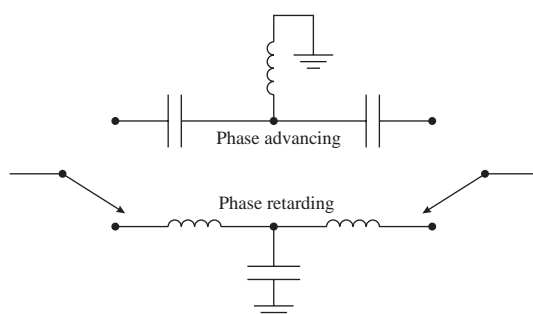
Figure 4.7 Low-pass/high-pass T-network phase shifter.
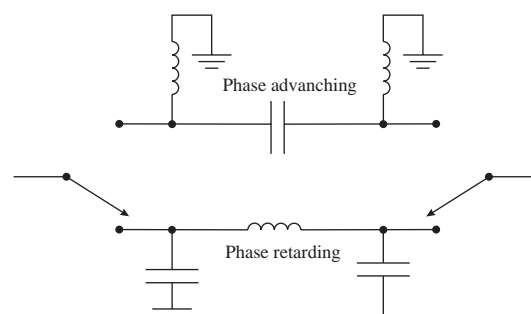
Figure 4.8 Low-pass/high-pass π-network phase shifter.

A high-pass filter can be shown to act as a phase advancing network. Similarly, a low-pass filter acts as a phase retarding network. Both of these networks can be implemented as a T (figure 4.7) or π network (figure 4.8).

4.2.4   Analogue phase shifter.

Analogue phase shifting networks give continuously variable phase as a function of a control voltage. Typically they use a reflection type phase shifter like that in figure 4.6, with varactor diodes used to control the phase.

4.3   Preferred time delay mechanism.

This project requires a method of generating a time delay which is ideally continuously variable, or at least variable in small steps. Experience shows that equipment used in a hospital environment needs to be very solidly constructed, as it is often transported along uneven surfaces. It also needs to be compact. As such, the movable source and detector method, discussed in section 4.1.1 is not really suitable, as it would be difficult to maintain optical alignment under these conditions, especially given the size of the detector, which would need to be very small to allow the generation of fast pulses. Although a trombone transmission line would be suitable, the physical size is undesirable. A switched line phase shifter would be suitable, if the steps were sufficiently small. However, this still requires a lot of cable although this could be folded.

The actual delay mechanism used was an R&K model PS-3 Phase shifter (R & K Japan, distributed by Pascal Microwave Ltd., Saxon House, Downside,Sunbury-on-Thames, Middlesex, TW16 6RY). This consisted of three 3 dB quadrature hybrids with varactor diodes varying the

reflection at the ports. This gave a continuously variable 0 to 360° phase shift, controlled by an analogue control voltage. The reasons for the final adoption of this phase shifter are explained more fully later, as they are related to the other items used in the system.

Chapter 5.

Pulse Generators.

In order to gain-modulate the correlating APD it was necessary to modulate its bias voltage with a short pulse. It has been shown earlier, that the greatest increase in gain occurs for large modulation voltages. For the greatest time resolution, it is necessary to generate as short a pulse as possible. Methods will be described for pulse generation.

5.1 Use of a photodiode for pulse generation.

If a fraction of the pulsed laser power is coupled to a photodiode, then it is possible to generate a short pulse of voltage. This method was investigated, but found to be unsuitable. The method of connection is shown in figure 5.1. Note the centre connection to the high voltage input connector is positive - a convention used throughout this thesis, but which is not standard practice. It is standard to use a negative centre, as this is safer when using PMTs, allowing the PMT photocathode to be at ground potential.

Small fraction of laser output power is coupled to photodiode, to generate reference pulses.

DC bias for photodiode.   R   Output of electrical pulses   C   50Ω
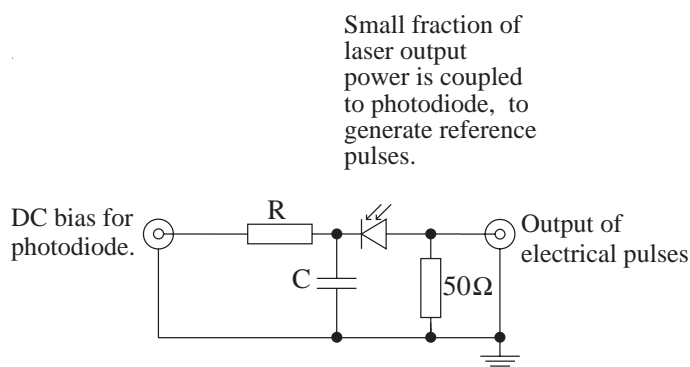
Figure 5.1 Diagram showing the reference APD receiver. The capacitor and 50 Ω resistor are soldered to the APD with negligible lead length.

Unfortunately, the voltages that can be generated by this method are very small. They can of course be amplified, but the bandwidth of the amplifier needs to be considerable to pass the pulse without causing significant broadening.

If the photodiode is replaced by an avalanche photodiode, then the level of the signal can be increased by the carrier multiplication in the APD.

5.1.1 Disadvantages of using an APD as the pulse generator.

There are a number of problems with using the APD as the pulse generator. These include:

1) Some fraction of the laser power must be used for pulse generation, which limits the maximum range and/or signal to noise ratio of the final instrument, due to decreased optical power injected into tissue.

2) It was only possible to achieve a voltage of about 2 V peak, when using 90% of the available laser power to generate the pulse.

3) The pulse width was greater than desirable.

4) The phase delay, must be a mechanical moving mirror, as used by Berndt, since the pulse risetime needs to be very fast (say <100 ps) and the bandwidth of any alternative phase shifting network needs to be at least 3.5 GHz.

5) The chopper used to generate a reference for the lock-in amplifier needs to have a wide bandwidth. Hence this too must be a conventional mechanical chopper or a very expensive high-speed electro-optic device or LCD shutter. There are no cheap wide-bandwidth, non-mechanical choppers available.

5.2   Use of a step recovery diode (SRD) for short pulse generation.

An electrical signal was available on the laser mode-locker, which is synchronised to the laser pulses. This is not a standard waveform (sine, square, triangle etc), but when passed through a low-pass filter with a cutoff frequency of approximately 100 MHz, the output becomes approximately sinusoidal at a frequency of 82 MHz. A pulse generator, that could generate ultra short electrical pulses synchronised to this sine wave drive would offer an ideal solution. The advantages would be:

1) It is no longer necessary to use any of the laser power for pulse generation. All the power could be injected into the tissue. Whilst lack of power is not a problem when using the Ti:sapphire laser, it is expected to be more of a problem when using solid state diode lasers, with their much lower output powers.

2) The time delay necessary for calculating the cross-correlation could be achieved with a phase shifting network operating at a single frequency of 82 MHz, avoiding the need for bandwidths of many GHz. All electronic phase shifters for doing this are readily available.

3) The chopper necessary for the lock-in amplifier reference can consist of a switch which switches on/off the 82 MHz sine wave to the pulse generator. The switch would no longer need a bandwidth of many GHz. Such a switch is again readily available.

A step recovery diode can be used to generate the required short electrical pulses from a sine

wave drive. For the device used for this project, there are also other advantages compared to the APD method of pulse generation.

1) The amplitude of the voltage pulse is greater (~7 V compared to 2 V of APD).

2) The pulse is narrower (~130 ps, compared to the 800 ps of APD).

A step recovery diode[114,115,116,117,118,119] is a pin junction diode that has a special doping profile, to give it an unusual characteristic curve. The idealised current characteristic of a step recovery diode are shown in figure 5.2, along with the output spike produced by this rapidly changing current. Like a normal pn diode, the SRD conducts when forward biased. Unlike the normal pn diode, it also shows significant conduction when reversed biased. However, this only continues for a short time, after which the diode no longer conducts when reversed biased. The transition from the conducting to non-conducting states of the reversed biased diode occur in a time of typically less than 100 ps. This means the current waveform has a high harmonic content and can be used for generating short pulses.

Figure 5.2 Idealised current waveform of an SRD, driven with a sine wave, and the resulting output spike produced.

Step recovery diodes are available as individual diodes, or as a module containing the SRD and a few passive components. SRD's have 2 main applications - (1) and (2) below and a third (3), less frequently used.

1) A multiplier[114,115,116,117,119]. This multiples an input frequency by a fixed integer. A typical x7 multiplier module, driven at 1 GHz would produce an output at only 7 GHz.

2) Comb generators[114,115,116,117]. These generate a number of harmonics simultaneously, which look like a comb if viewed on a spectrum analyser.

3) Occasionally SRDs are used to sharpen the rise time and reduce the pulse width[120] of pulses generated by other SRDs or other components.

5.2.1   Theory of a step recovery diode.

An SRD is made so that majority carriers[102] (electrons on the n side, holes on the p side) injected across the junction under forward bias create a number of minority carriers (electrons on the p side, holes on the n side). Unlike a normal diode, where these soon recombine, the minority carrier lifetime in an SRD is sufficiently long to ensure that carriers do not immediately recombine, so that charge is stored. When the applied voltage changes to reverse bias the diode, these minority carriers (charge) are withdrawn across the junction due to the applied reverse bias, where they become majority carriers. Hence a current flows in reverse bias, until all the carriers are back to their original side of the junction. In order for this to work, two conditions must be met:

1) The carriers must not recombine before they are withdrawn.

2) The minority carriers must not diffuse too far from the junction, that they can not be withdrawn under reverse bias.

It should be evident, that the minority carrier lifetime of the diode determines the frequency of operation. Silicon can be made to have a minority carrier lifetime of approximately 10^{-6} s to 10^{-8} s, which limits the fundamental frequency of operation of silicon devices from approximately 10 MHz to 1000 MHz. Gallium arsenide, which has a minority carrier lifetime of less than 10^{-10} s, would only be useful if driven at mm-wavelength microwave signals (100 GHz). As such, all SRDs are currently made with silicon, although one can expect GaAs devices to be available in the future.

The SRD is also known as a charge storage diode, since its operation depends on the ability to store charge. Sometimes it is known as a snap diode, for obvious reasons.

5.2.2   Ideal dynamic characteristics.

The charge stored can be easily obtained from the charge continuity equation:

rate of increase of charge=rate of injected charge - rate of recombination

hence

$$i = \frac{q}{\tau_{mc}} + \frac{dq}{dt} \qquad \textbf{5.1}$$

where i = instantaneous diode current.

q = charge stored at junction.

τ_{mc} = minority carrier lifetime of diode.

This is easily solved by any standard means - Laplace transforms[121] will be used here.

$$LT\left[i\right]\equiv LT\left[\frac{q}{\tau_{mc}}\right]+LT\left[\frac{dq}{dt}\right] \tag{5.2}$$

$$\frac{I(s)}{s}=\frac{Q(s)}{\tau_{mc}}+sQ(s)-Q_0 \tag{5.3}$$

where the convention of using upper case characteres to denote parameters in the s domain and lower case characters for the time domian is used (i.e. $f(t) \leftrightarrow F(s)$). Q_0 is the initial charge.

$$Q(s)=\frac{I}{s\left(s+\frac{1}{\tau_{mc}}\right)}+\frac{Q_0}{s+\frac{1}{\tau_{mc}}} \tag{5.4}$$

if there is previously no stored charge ($Q_0=0$) and a constant forward current I_F

$$Q(s)=\frac{I_F}{s(s+\frac{1}{\tau_{mc}})} \tag{5.5}$$

taking an inverse Laplace transform of equation 5.5, we get a solution for the stored charge as a function of time.

$$q_F=i_F\tau_{mc}\left[1-\exp\left(-\frac{t_F}{\tau_{mc}}\right)\right] \tag{5.6}$$

where q_F is the charge stored from the forward current i_F and t_F is the length of time for which the forward current i_F is applied.

If we now apply a reverse current i_R to withdraw the charge and switch the diode, the time required to do so is easily found from equation 5.4 with the initial charge Q_0 not equal to zero, but some finite value due to the forward current injection. Taking an inverse Laplace transform of equation 5.4 gives

$$q(t)=-i_R\tau_{mc}\left[1-\exp\left(-\frac{t_R}{\tau_{mc}}\right)\right]+q_0\exp\left(-\frac{t_R}{\tau_{mc}}\right) \tag{5.7}$$

The negative sign in front of i_R is due to the assumed positive value of the reverse current, even

though it flows in the opposite direction to the forward current. When the reverse current has been flowing for some time t_R, such that all the charge is removed, the diode will snap. This time can easily be found by solving equation 5.7 for q(t)=0.

$$t_R = \tau_{mc} \ln\left[\frac{q_0 + \tau_{mc} i_r}{\tau_{mc} i_R}\right] \qquad \textbf{5.8}$$

Substituting for the initial charge q_0 due to forward current injection

$$q_0 = i_F \tau_{mc}\left[1 - \exp\left(-\frac{t_F}{\tau_{mc}}\right)\right] \qquad \textbf{5.9}$$

gives

$$t_R = \tau_{mc} \ln\left[1 + \frac{i_F\left(1 - \exp\left(-\frac{t_F}{\tau_{mc}}\right)\right)}{i_R}\right] \qquad \textbf{5.10}$$

This is usually called the snap time, t_s. If t_F is long compared to τ_{mc} and $i_F \ll i_R$, then

$$t_R = t_s \approx \frac{\tau_{mc} I_F}{I_R} \qquad \textbf{5.11}$$

5.2.3   Equivalent circuit of an SRD.

Figure 5.3 is an equivalent circuit for an SRD[120].

where

C_p = Package capacitance

L_p = Package inductance

R_s = Dynamic series resistance

C_j = Reverse bias (depleted) junction capacitance

ϕ = Contact potential (0.7 V)

5.2.4   Actual dynamic characteristics.

A number of assumptions have been made in the section 5.2.2, that are not strictly true and so modify the actual device characteristics. These include:
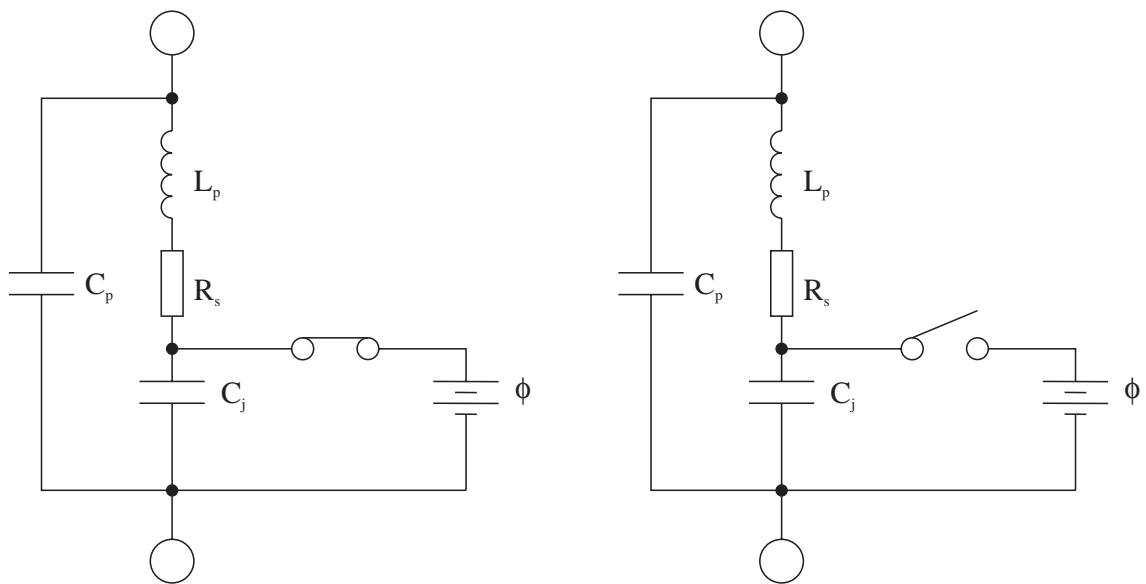
Forward bias   Reverse bias

Figure 5.3 Equivalent circuit of an SRD in both forward and reverse bias.

1) Zero voltage drop under forward bias.

2) Infinitely fast transition time.

3) Zero package inductance.

The effects of these will now be considered in turn.

5.2.5   Forward bias voltage drop of an SRD.

When forward biased, a voltage drop V_F will occur given by:

$$V_F = \phi + i_F R_s \qquad\qquad \textbf{5.12}$$

Most practical circuits employing an SRD use capacitive coupling, so this does not appear at the load.

5.2.6   Transition time of an SRD multiplier.

In practice, the SRD does not instantaneously change from being in a conducting to a non-conducting state. There is always a transition time, usually called the snap time and denoted T_s, defined as the time the current waveform takes to change from 80% to 20%, as shown below. Some manufacturers call it the transition time t_t. The diode snap time T_s should be less than the output period of the highest frequency component in the output of an SRD multiplier.

$$T_s < \frac{1}{f_{max}}$$

5.13

For this project, where a pulse of as short as duration as possible is required, one might expect that a diode be selected with as short a snap time as possible. Unfortunately, diodes with short snap times, always have short minority carriers lifetimes, as the following table giving characteristics[122] for a number of SRD's formerly made by Hewlett-Packard, who were one of the largest manufacturers of SRD's, but have now stopped production completely.

Table 5.1 Characteristics of some Hewlett-Packard Step Recovery Diodes

Part number	Junction capacitance C_j (pF)	Reverse breakdown voltage V_B (V)	Minority carrier lifetime τ_{mc} (ns)	Snap time T_s (ps).
5082-0032	4.00	65	150	250
5082-0112	1.4	35	50	175
5082-0151	0.5	15	10	100
5082-0835	0.4	15	10	75
5082-0018	0.5	25	20	70
5082-0008	0.38	15	10	60

It was shown earlier, that the minority carrier lifetime needs to be adequately long, as otherwise there will be significant recombination before the diode will snap, resulting in a much reduced amplitude for the pulse. This means that SRD's driven at high frequencies will be able to produce higher harmonics (faster pulses) than those driven at lower frequencies.

In practice, the output current of an SRD does not change as fast as the transition time, due to package capacitance and load resistance. The risetime from 10% to 90% of the current waveform t_r is given by[120]

$$t_r = \sqrt{T_s^2 + (2.2 R_{eq} C_j)^2}$$

5.14

where R_{eq} is the equivalent resistance consisting of the parallel combination of the source and

load resistance.

5.2.7   SRD Package inductance.

The package inductance L_p causes a voltage spike, V_L due to the rapid change of current i_d through the diode.

$$V_L(\text{max}) = L_p \left[\frac{di_d}{dt} \right] \qquad \textbf{5.15}$$

When used for fast pulse generation, this can be a significant voltage spike. For this reason, the package inductance is generally kept as small as possible.

5.2.8   Reducing risetime of pulses with an SRD.

SRDs can be used to reduce the rise time of a pulse generated by other means. This is discussed fully in the HP application note[120], where equations are given for designing practical circuits. It is said that a 10 ns risetime pulse can be sharpened to <300 ps with one SRD, or even less if two or three circuits are coupled together. The application note also discusses how to reduce the pulsewidth of a pulse, using an SRD.

5.2.9   High order frequency multiplier.

This circuit takes a sine wave at one frequency, say 1 GHz and produces another at some high multiple, for example 20 GHz. It is of no use for this project, but is discussed here as the comb generator, which is a minor modification of the multiplier, is important.

5.2.10   Input matching circuit.

The function of this is to match the output of the sine wave drive, which is conventionally 50 Ω to the SRD. Unfortunately, the SRD has two impedance states, one during forward drive, which is about 1-20 Ω and the other in reverse bias which is 50-500 Ω, so the match is never very good.

5.2.11   Impulse circuit.

This functions as a voltage pulse generator. This produces the voltage spikes from the rapid change in current when the diode snaps. The impulse voltage is given by

$$V = L_i \frac{di_d}{dt} \qquad\qquad \textbf{5.16}$$

The impulse circuit will be most efficient in pulse generation if:

1) The snap time of the diode is less than the time period of the output frequency.

2) The impulse circuit resonates at or just above the output frequency required.

3) The diode cutoff frequency f_c, given by

$$f_c = \frac{1}{2\pi RC} \qquad\qquad \textbf{5.17}$$

is at least 10 times the output frequency required.

5.2.12 Output matching circuit.

The purpose of this is to match the diode to the 50 Ω transmission line. A band-pass filter is included in the output matching circuit to select the required harmonic.

5.2.13 Comb generators.

By far the most common application of SRDs is in making comb generators. These are harmonic generators, used to produce simultaneously a large number of harmonics from a single frequency source. A harmonic generator, driven at say 1 GHz, will produce outputs at 1 GHz, 2 GHz, 3 GHz ... to some maximum value, determined by the SRDs capacitance, transition time and the output circuit of the comb generator.

SRD comb generators consist of an input matching circuit, as with the harmonic generator. The impulse circuit containing the SRD must not resonate at or below the maximum frequency required for the comb generator. The output matching circuit, must match the SRD to 50 Ω over a wide band. The match is inevitably poorer than for a harmonic generator, where match is required only at one frequency. As such, the output circuit should be as simple as possible, to
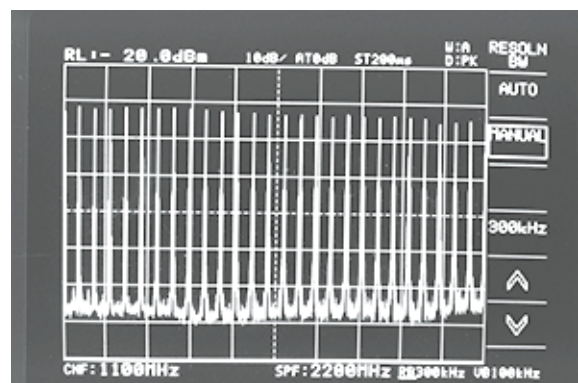
Figure 5.4 Spectrum analyser display showing the output from DC to 2200 MHz of a Herotek GC082-112 SRD comb generator driven at 82 MHz. This is the SRD module used in this project.

prevent unwanted resonances.

Considered in the time domain, comb generators produce a series of narrow pulses, each separated by a time equal to the time period of the fundamental drive voltage. Figure 5.5 shows the output of a Herotek GC082-142 SRD comb generator module as measured on a fast sampling oscilloscope described on page 138. The output is seen to consist a series of pulses, which if shown on an expanded scale, are around 130 ps FWHM with a repetition time of 12.2 ns. These pulses can be used for changing the gain of
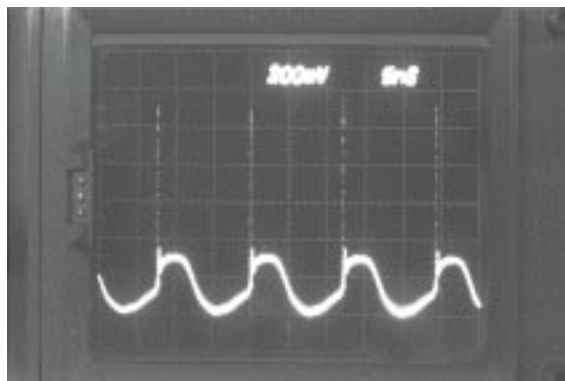
Figure 5.5 Diagram showing the output of the Herotek GC082-112 SRD comb generator as measured on a fast sampling oscilloscope. A 20 dB attenuator is fitted between the SRD and the oscilloscope.

the APD. However, the residual sine wave with an amplitude of approximately 2 V peak is undesirable. It is shown later that this caused problems in the cross-correlator and steps were taken to remove the residual sine wave, whilst leaving the pulses intact.

5.2.14 Temperature dependence of SRDs.

The minority carrier lifetime of a semiconductor is strongly temperature dependent. The rate of change is from approximately +0.5% to + 0.7%/°C[123]. This can be a problem when diodes are used for generating harmonics for microwave signal generation, as the optimum bias level is dependent on the minority carrier lifetime.

Unfortunately, the problem is more serious when, as in this application, we are interested in the position in time of the pulses. As can be seen from equation 5.11, a change of minority carrier lifetime, causes a change in snap time t_s, which is a change in the position in time of the pulses relative to the driving waveform. Since the driving waveform is synchronised to the laser pules, a change of temperature causes a timing drift. If the minority carrier lifetime increases due to an increase in temperature, without a change in bias level, we would expect the following to happen:

1) The charge injected at the end of the forward bias cycle will be increased, as indicated by equation 5.6.

2) This will increase the duration of the reverse cycle until the diode snaps, as indicated in equation 5.10. Hence an increase of temperature and therefore minority carrier lifetime will cause

the pulses to occur at a later stage of the RF cycle.

3) If the assumption that $I_F \ll I_R$ is valid, the simplified equation 5.11 can be used. This indicates a linear relationship between the snap time t_s and τ. Clearly by altering the forward and reverse currents, it is in principle possible to temperature compensate.

Some liberties will now be taken in equating results obtained for a square wave earlier, with a sine wave, so the results of this analysis can only be approximate. If we assume that the snap time t_s is a quarter of the fundamental AC cycle, then the snap time will be given by:

$$t_s = \frac{1}{4f}$$
                                                                    5.18

From equation 5.11, we know that for a constant forward and reverse currents, the snap time is directly proportional to the minority carrier recombination time τ_{mc}. Assuming a +0.6%/°C change in τ with temperature, t_s will also change at +0.6%/°C. If the drive signal is 82 MHz, the snap time t_s should be approximately 3 ns, so a 1°C increase in temperature, will cause a shift of the pulses of approximately 0.6% of 3 ns, or 18.3 ps/°C. This would suggest that temperature changes can not be ignored for this application, but are not likely to prevent the technique from working, if simple precautions are taken.

5.2.15   Radio frequency drive levels for SRD diodes.

The bias currents, are frequently generated by the diode alone. The comb generator modules require no DC supplies at all. The bias currents are generated by the RF waveform, and so depend on the drive level. The snap time and hence the accuracy of synchronisation to the laser pulses, depends too on the drive level.

5.2.16   Practical aspects of pulse generation by SRD.

The step recovery diode module used for this project required a drive of +27 dBm (500 mW) at 82 MHz. The power output from the laser mode-locker after filtering to remove harmonics was -12.6 dBm. Two amplifiers connected in series were needed to increase the -12.6 dBm to +27 dBm. The first stage, providing up to 30 dB of gain, used a commercial wide band amplifier, to which an AGC voltage could be applied. The final stage, which increased the power level to +27 dBm, was used to directly drive the SRD comb generator.

5.2.17   Specification and measurement of SRD parameters.

Apart from specifications which are common to all diode types, such as junction

capacitance, series resistance, reverse breakdown voltage, maximum current, maximum junction temperature etc, the SRD has a number of specifications which are unique to it. It should be noted that not all manufacturers define these in the same way, so care is needed in comparing specifications.

Transition time t_r

HP define this as the time between the 20% band 80% amplitude points of the overall step amplitude of the SRD. Since it is measured in a test fixture, the true time of the diodes t_r is less than the measured value, due to the diode junction capacitance. It is typically 60 ps to 400 ps.

Carrier lifetime τ

This is the minority carrier lifetime of the semiconductor. This parameter is particularly important, as it determines what frequency the SRD can be used at. It is typically 10 ns to 150 ns.

5.2.18   Conclusions about the SRD method of reference pulse generation.

The step recovery diode does in principle offer an ideal way of generating the reference pulses needed for this project. It provides large amplitude, short pulses, with a phase that is easily controllable by a simple, fairly inexpensive, rugged, all electronic phase shifter. This has been implemented by purchase of an SRD comb generator, model GC82RC from Herotek Inc, Sunnyvale, CA 94086, USA.

It has been shown that by the use of a number of step-recovery diodes, it is possible to further reduce the risetime of pulses and probably reach a pulse width of 50 ps (risetime approximately 20 ps). Unfortunately, such methods require a custom designed pulse sharpening circuit, since the component values depend on the amplitude, repetition rate, risetime, pulsewidth and falltime of the original pulse as well as the risetime, pulsewidth and falltime of the required pulse. Suitable pulse sharpening circuits are not available commercially - there are too many combinations for any manufacturer to make. Although such circuits can be designed, this has not been done here, since the problems of making a wide band circuit, that show no unwanted resonances to a frequency of approximately 10 GHz is very considerable.

The approach which was taken here was to use the Herotek GC082-142 SRD harmonic

generator, as it offers a method of generating very short pulses, of large amplitude, with a phase that is controllable by altering the phase of the RF drive signal.

Chapter 6.

Instrument Design and Development.

This chapter describes the instrument design and development process. The final design was very much determined from both experiments performed on small subsections of the complete instrument and experiments performed on complete earlier prototypes of the instrument.

6.1 Lock-in amplifier filter response.
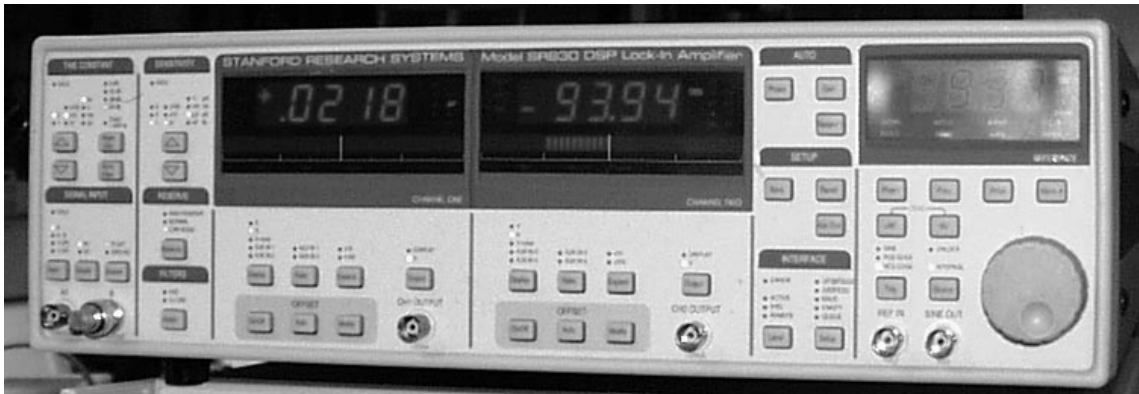
Figure 6.1 Photograph of the Stanford Research Systems SR-830 DSP based lock-in amplifier used throughout this research.

The system proposed in figure 1.26, required a lock-in amplifier. A Stanford model SR830 digital signal processor (DSP) based lock-in amplifier (Stanford Research Systems Inc, 1290-D Reamwood Avenue, Sunnyvale, California, 94089, USA) was purchased. A photograph of this instrument is shown in figure 6.1. Before attempting to make measurements with the system using the lock-in amplifier, it was essential to know the settling time of the instrument's low-pass filters. If insufficient time were left for the amplifier to settle, the results would be inaccurate, but if left too long, time would be wasted, which would result in a poorer signal to noise ratio than necessary. Unfortunately, the settling time was not specified in the lock-in amplifier's manual. To find this out, the internal oscillator was set to 1 kHz and the input amplitude varied from 0 to 5.000 V in a step. The time constant T was set to 1 s, and the voltage measured every 125 ms, up to 16 seconds. All measurements were made via a computer connected to the General Purpose Interface Bus (GPIB), also known as the Hewlett-Packard Interface Bus (HPIB). Figure 6.2 shows the data collected, which has been normalised. The
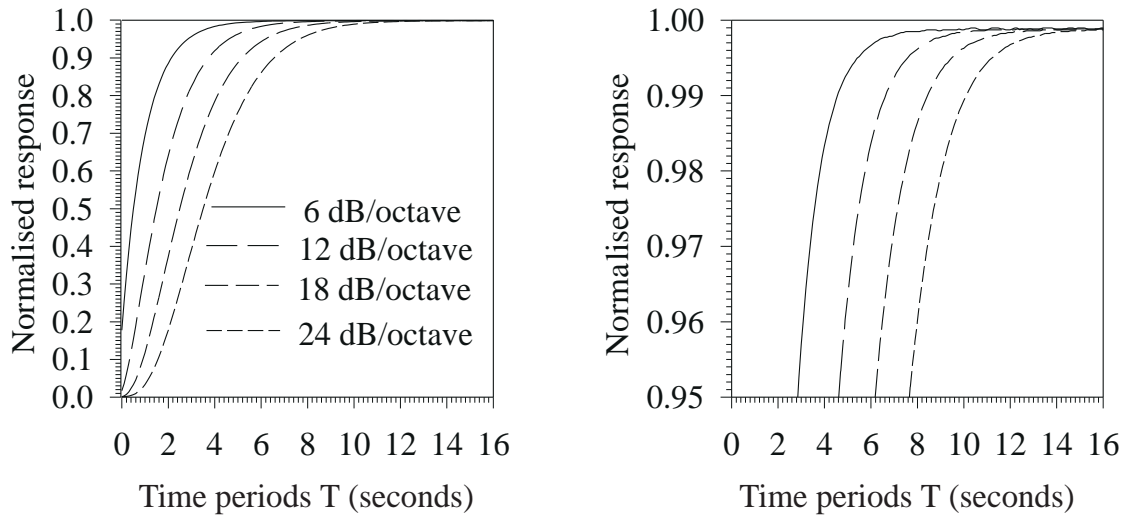
Figure 6.2 Measured step response of the Stanford Research Systems SR-830 DSP lock-in amplifier at filter slopes of 6, 12, 18 and 24 dB/octave.

settling times to 99 % are 4.6 T, 6.7 T, 8.6 T and 10.2 T for the 6, 12, 18 and 24 dB/octave filter settings respectively. All subsequent data was collected with this in mind.

The -3 dB bandwidth of the filters f_{3dB} is simply given by

$$f_{3dB} = \frac{1}{2\pi T}$$ **6.1**

so a 1 second time constant has a -3 dB bandwidth of 0.16 Hz. The equivalent noise bandwidth (ENBW) of the filters is **not** the same as the -3 dB bandwidth. For Gaussian noise, the equivalent noise bandwidth (ENBW) of a low-pass filter is the bandwidth of the perfect rectangular filter which passes the same amount of noise as the real filter. The equivalent noise bandwidth of the filters is given in the lock-in manual and is reprinted in table 6.1.

Table 6.1 Effective noise bandwidth of lock-in filter

Slope	ENBW (Hz) for time constant T (s)
6 dB/oct	1/(4 T)
12 dB/oct	1/(8 T)
18 dB/oct	3/(32 T)
24 dB/oct	5/(64 T)

6.2   Hybrid electrical and mechanical design.

Initially, no way of developing a fully electronic cross-correlator, with no moving parts was known. Hence development was started on the hybrid electronic and mechanical system, shown in figure 6.3. This was similar to Berndt's system shown in figure 1.25. Light was first focused from the laser onto a fibre using a fibre-optic coupler. The light was then split into two
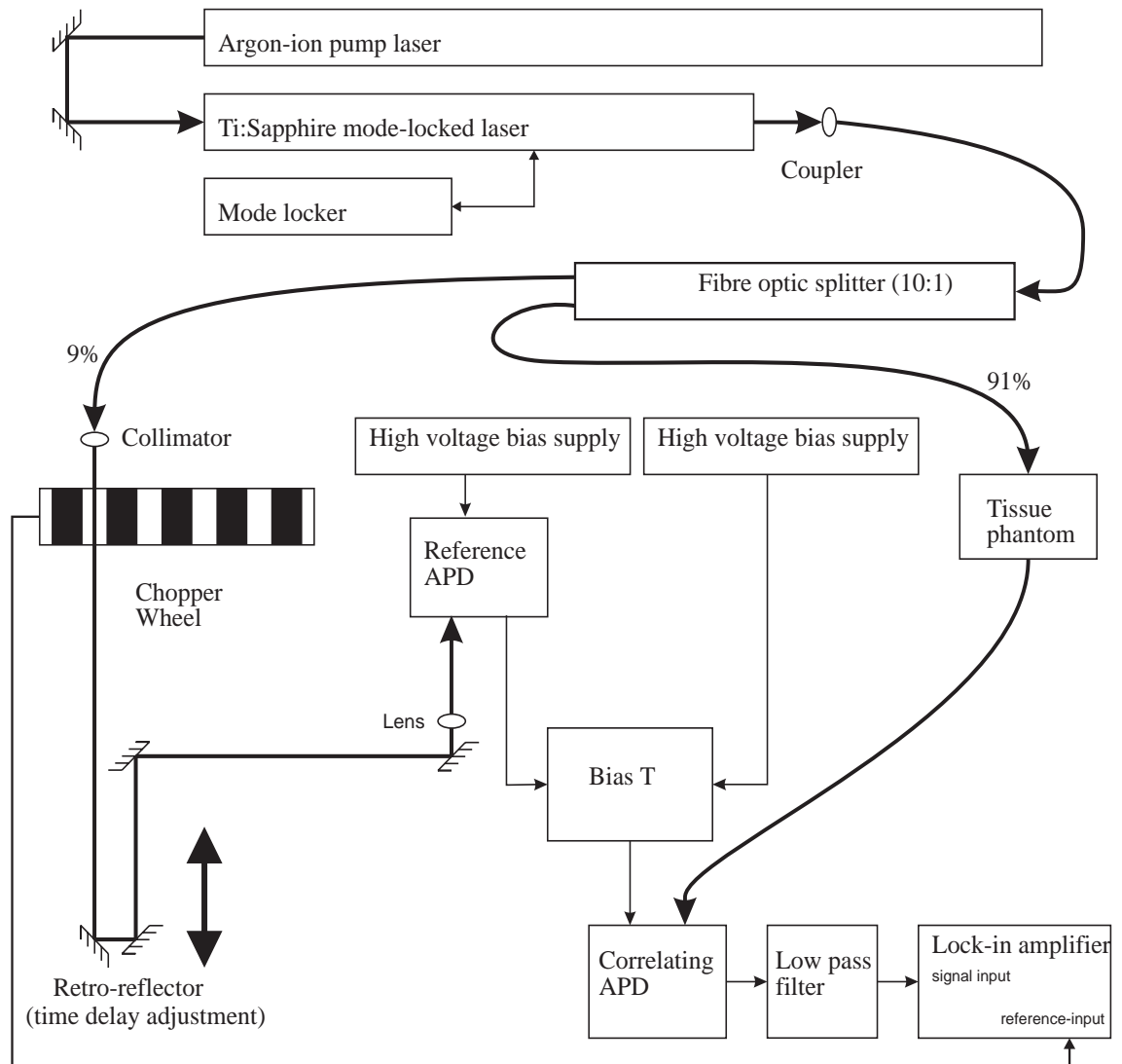
Figure 6.3 The design a cross-correlator, on which development commenced during the early stages of this project. Optical connections are shown as thick lines, electrical connections as thin lines.

components using a commercial fibre-optic splitter, which was specified as having a 10:1 split ratio. The smaller fraction of the lasers output was then sent to an APD, via a mechanical chopper wheel and a time-delay mechanism for the generation of the reference pulses (see section 6.2.1), with the larger fraction going to a solid tissue phantom[124,125] and the

correlating APD. The time-delay adjustment consisted of a collimator, and an Aerotech retro-reflector (Aerotech, 101 Zeta Drive, Pittsburg, PA 15238, USA) mounted on an Aerotech PR-12/M 300 mm slide rail. A number of precision adjustable optical components were added, so that the beam could be precisely aligned on the APD. To enable a time delay of 2 ns to be achieved, the beam was folded once (see figure 4.3), so that 300 mm movement of the slide, caused a 600 mm change in the optical path length. This was found to be very difficult to align, as when the optical slide was moved, the highly collimated beam wandered off the small ADP, resulting in a large (factor of 2) change in the size of the optical pulses. It soon became apparent that the fibre splitter did not provide an accurate 10:1 split ratio, but that the split depended on exactly how the light was focused onto the input fibre. Under some conditions, it was possible to get more laser power from the low power arm of the splitter than from the high power arm! Attempts at mode scrambling by passing the light first through 20 m of fibre wound around a convoluted path did not help the problem. This resulted in it being impossible to keep the power falling on either the solid tissue phantom or the reference APD constant. In short, the fibre splitter could not form part of a stable system.

6.2.1 Reference pulse generator.

The reference pulse generator consisted of an Hamamatsu S2381 avalanche photodiode, decoupling capacitor and series resistor (see figure 5.1). An abbreviated specification of the APD is given in Appendix 1. The manufacturer tested each APD individually and gave the following test data for the devices purchased. These devices were not all purchased together - some were bought at a later date, but are included here as a complete list of those used throughout the project. Each APD has an active diameter of 200 μm (0.03 mm^2).

Table 6.2 Characteristics of APD's used.

Serial number	Colour.	V_b	γ	V_{100}	I_r (@ M=100)
A0887	Mauve	122.0 V	0.0922	109.4 V	21 pA
A0940	Black	125.0 V	0.0915	112.0 V	22 pA
NA04649	Green	135.0 V	0.1026	122.4 V	47 pA
A0726	Orange.	150.0 V	0.0981	135.4 V	33 pA

where V_b is the breakdown voltage, V_{100} is the voltage necessary to produce a gain of 100 and

I_r is the leakage current at a gain of 100. The *colour* is simply a colour marked on the devices for identification purposes. The factor γ, is the constant calculated according to equation 3.32. All the manufacturers measurements were taken at 25°C and 803 nm.

Since initially only two devices (*mauve* and *orange*) were purchased, it was necessary to choose which one to use for the APD reference pulse generator and which to use for the detector. Since the detector was needed to operate at very low light levels, where photocurrents would be small, it was decided to use the *mauve* device as the detector and the *orange* device as the pulse generator, as the *mauve* device had a smaller dark current. It was shown earlier (equation 3.16 in chapter 3) that the dark current reduces the S/N ratio.

The APD was connected in a standard configuration, as recommended by the device manufacturer and shown earlier in figure 5.1. This particular configuration, has the advantage that the load is connected to earth on one side, whereas other biasing arrangements have a floating load. Note the inclusion of the 50 Ω resistor which provides an output impedance of 50 Ω, enabling connection to the correlating APD using a high quality 50 Ω microwave coaxial cable to standard 50 Ω test equipment. Since the cable was also terminated at the load end by 50 Ω, the photodiode current generated a voltage across an impedance of 25 Ω. This limited the output voltage somewhat, but was expected theoretically to give the best high frequency response. The APD manufacturer Hamamatsu states that the resistor at the source end can be increased to 1000 Ω, which results in an almost 100 % increase in output voltage. However, this would degrade the pulse shape slightly and so was not done. The reference pulse generator was constructed in a die-cast box, completely separate from any other components. The pulse generator had just three connections, these being a bayonet fitting high voltage DC connector, FC-PC optical connector and a screw type SMA connector for the electrical output pulses. The 50 Ω resistor was a microwave chip resistor, rated for low voltage standing wave ratio (VSWR) to 12 GHz. This was soldered on the back of the APD, with effectively zero lead length.

The high voltage connector used was a bayonet fitting Shuner MHV/H4 type. This was connected by a length of RG59 73 Ω coaxial cable to the high voltage power supply which was housed separately. The MHV connector was used to prevent accidental mating with the more common connectors, such as BNC found in the laboratory and because the centre pin (positive) of this connector is well shielded. The outer of the RG59 coaxial cable is at earth potential.

6.2.1.1   APD pulse generator measurements.

For assessment of the performance of the photodiode as an impulse generator, measurements were made with a Tektronix 7604 oscilloscope, fitted with a 7T11 sampling sweep unit, a 7S11 sampling unit and an S-4 sampling head (Tektronix Inc, P. O. Box 500, Beaverton, OR 97077, USA). This gave a sampling oscilloscope with a risetime of ≤25 ps (3 dB bandwidth ≥ 14 GHz). This was connected via a short, high quality microwave coaxial cable to the output from the photodiode.

The circuit in figure 5.1 was used for the first set of measurements, with the series resistor equal to 100 kΩ and the decoupling capacitor 10 nF. These are typical values given in manufacturer's data sheets. The bias voltage was 135.4 V, which gave a gain of 100. First, the laser, with a beam diameter of approximately 1.5 mm was attenuated to an average power of up to 5 mW. The response from the photodiode was a repetitive train of pulses, with a peak amplitude of 150 mV (at 5 mW) and a fast risetime (~200 ps), but a much longer fall time (~3 ns). At lower input powers, the output signal was smaller. It was difficult to accurately determine the actual power falling on the photodiode which had a much smaller area (0.03 mm^2) than the laser beam (1.8 mm^2), since neither the laser beam diameter, nor its profile, were accurately known. This also meant that the response from the photodiode changed depending on its exact position in the beam. When moved to the approximate centre of the beam, very small changes in the position of the photodiode, caused large (sometimes a factor of two) changes in its output signal. Clearly the beam profile is not constant, due to the mode structure of the laser beam. As such, all results were obtained by putting the photodiode approximately central to the laser beam, then adjusting the exact position to maximise the photodiode response.

The 150 mV signal recorded with 5 mW of optical power was not initially considered sufficient to adequately modulate the correlating APD, so the laser power was increased to 13 mW. Unfortunately, this caused a reduction in the amplitude of the signal to 50 mV peak, but an increase in the falltime to approximately 10 ns! This effect is probably due to saturation of the pn junction of the diode, so that the carriers cannot be swept away by the applied electric field. Although the data sheet does not give the maximum optical power for the APD, which is designed for very low light levels, Hamamatsu suggested no more than 100 μW be used at a gain of 100. Assuming a responsivity of 0.5 A/W (equation 3.5), 100 μW would give a current of approximately 5 mA. It is useful to consider what the peak optical power on the photodiode was when exposed to the laser beam, as this might indicate a figure well in excess of the recommended maximum 100 μW.

Assuming the laser beam has a top-hat profile, which was a reasonable assumption for a short pulse laser, then if the mean laser power as measured on a power meter with a large area photodiode is P_{avg} (W), the beam diameter of D_{laser} (m), a time between pulses of T (s), a pulse duration of t_p (s) and a photodiode diameter of $D_{photodiode}$ (m), then the peak power reaching the photodiode will be:

$$P_{peak} = \frac{P_{avg} T D_{photodiode}^2}{t_p D_{laser}^2}$$
6.2

Assuming a laser pulse width of 2 ps (which was typical for the Tsunami laser), 12.2 ns between pulses, a photodiode diameter of 200 μm, a laser beam diameter of 1.5 mm and 5 mW mean power, we get 1.38 W as an estimate for the peak laser power on the photodiode! It was thus not surprising that the APD's characteristics are deviating from the ideal.

It was clear that it not possible to increase the output signal from the reference photodiode by simply increasing the input optical power. Also, the large power required to get even 150 mV of output, would cause problems in using an economical low-power laser diode. Using a larger area photodiode would allow a larger output, but the pulses would be longer due to the increased device capacitance. At the start of this project, the cost of a very wide bandwidth amplifier for the APDs output was too expensive. However their cost has fallen rapidly in the last few years, so this is now not a limiting factor although in the final system, the reference APD was dispensed with completely.

Although the decoupling capacitor used in the above experiments was 10 nF, as recommended in the data sheet for the APD used, it was noted that the capacitor used (ATC 900C103KP300, made by American Technical Ceramics, New York, USA) had a parallel self resonant frequency of approximately 100 MHz. This was considered undesirable for a system which was hoped to have a bandwidth extending to many GHz, since at parallel resonance, the capacitor is effectively absent. Generally speaking, the self resonant frequency of capacitors falls as the value of the capacitor increases, as figure 6.4 shows, for the ATC 111 series capacitors, which are high frequency chip capacitors intended for microwave applications. The smallest value of capacitor, that once fully charged, would supply sufficient charge, to allow an output voltage and pulse width similar to that Berndt had achieved using the APD, without serious change in the potential across the APD, was calculated.

Assuming a peak output pulse of 1.5 V and a load resistance of 25 Ω, the peak current would need to be 60 mA. Berndt had achieved a 320 ps FWHM. Assuming for simplicity, that

the output current was a constant 60 mA for a period of 320 ps, the output charge Q per pulse would be

$$Q=i.t=60\times10^{-3}\times330\times10^{-12}=2\times10^{-11}C \qquad \textbf{6.3}$$

The decoupling capacitor will initially be charged to the bias voltage of the APD, which was 135.4 V. As it supplies charge, which will be deposited in the load resistance, the capacitor voltage must fall. We can ignore the charge that will reach the capacitor from the HV supply during the period of the output current pulse, since this will be limited by the series resistor to a negligibly small amount. Clearly the APD, which has a strong gain versus voltage characteristic, needs to have a fairly constant voltage to keep its gain stable. This is why highly stabilised supplies must usually be used for APD receivers to keep their gain constant and so they function in a linear manner. However, in this application, where we were interested in obtaining a very large signal (~1.5 V), linearity was not so important. From inspection of figure 5.1 it
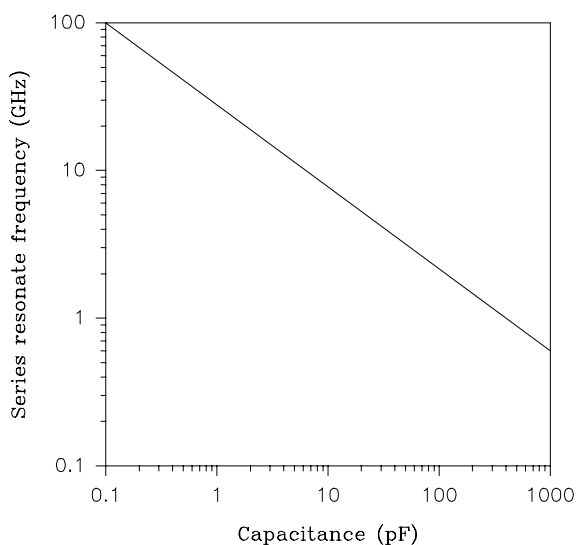
Figure 6.4 Series resonant frequency for the ATC 111 series of 100 V working capacitors for RF/microwave applications.

can be seen that the potential at the anode of the APD (relative to earth) must drop by an amount equal to the output pulse (say 1.5 V), no matter how effectively the decoupling capacitor keeps the potential at the cathode constant (relative to earth). Hence there will be little advantage in increasing the decoupling capacitor beyond a value which causes say a 10% (150 mV) fall in the cathode potential during the optical pulse.

Let the initial voltage on the decoupling capacitor be V_1, which is the bias voltage. Hence the energy initially stored in the capacitor E_1 is given by equation 6.4:

$$E_1=\frac{1}{2}CV_1^2 \qquad \textbf{6.4}$$

At the end of the output current pulse, the voltage across the decoupling capacitor will have dropped to a new voltage V_2 ($V_2<V_1$), and the energy will have fallen to a new value E_2 ($E_2<E_1$) given by equation 6.5.

$$E_2 = \frac{1}{2}CV_2^2 \qquad\qquad\qquad \textbf{6.5}$$

The change in energy in the capacitor must be equal to the energy supplied to the load,

$$E_1 - E_2 = \frac{1}{2}C(V_1^2 - V_2^2) = V_{out}it \qquad\qquad\qquad \textbf{6.6}$$

hence the minimum value of the decoupling capacitor necessary to keep the voltage on the cathode constant within a voltage $\Delta V = V_1 - V_2$ is given by equation 6.7:

$$C \geq \frac{2itV_{out}}{V_1^2 - V_2^2} = \frac{2itV_{out}}{2V_1\Delta V - \Delta V^2} \approx \frac{itV_{out}}{V_1\Delta V} \qquad\qquad \textbf{6.7}$$

Solving this for an initial bias voltage of $V_1=135.4$ V, allowing a drop of 150 mV ($V_2=135.4-0.15V=135.25$ V), an output pulse of 1.5 V peak ($V=1.5$) at 60 mA for 320 ps, gives a minimum capacitance of 1.5 pF. Hence the original choice of 10 nF was considerably greater than necessary for a short impulse generator. If we were to allow the voltage drop across the capacitor to be made equal to the load voltage, we only need a capacitance of 0.15 pF, a value which will certainly be exceeded by stray capacitance to earth.

The value of decoupling capacitor was changed from the original 10 nF to 8.2 pF (since one was readily available), then it was removed completely. The effect on the output pulse amplitude or width was negligible. Hence the initial concern about this being critical was unfounded, probably since there were no resonances at the 82 MHz drive or a harmonic of this.

The value of the series resistor was reduced from 100 kΩ to 10 kΩ, which made it possible to achieve an output signal of approximately 800 mV peak, at approximately 750 ps FWHM, as measured on the Tektronix sampling oscilloscope (see page 138), but since the fibre coupler was not stable together with the difficulty of aligning the laser on the APD, this was often unstable over time. No measurements were made with the resistor below 10 kΩ, for fear of destroying the APD. The system shown in figure 6.3 was clearly unsatisfactory, but since an idea for a fully electronic system had subsequently evolved, it was decided to concentrate efforts on a fully electronic system. Ultimately, a system requiring a number of critical optical alignments (which the system in figure 6.3 would always have), would not be satisfactory for the proposed purpose.

6.3   An electronic prototype - basic thoughts.

A step recovery diode, which is generally used to generate a considerable number of harmonics of an RF signal, must hence provide a short pulse in the time domain. This was essentially what was needed to gain-modulate the APD, since the gain-modulation pulse must be highly correlated with the laser output, which was a short pulse. It should be possible to obtain a suitable RF drive signal from the laser mode-locker, although a method of connecting the SRD to the laser mode-locker was not immediately apparent.

6.3.1   Method of connection to the mode-locker.
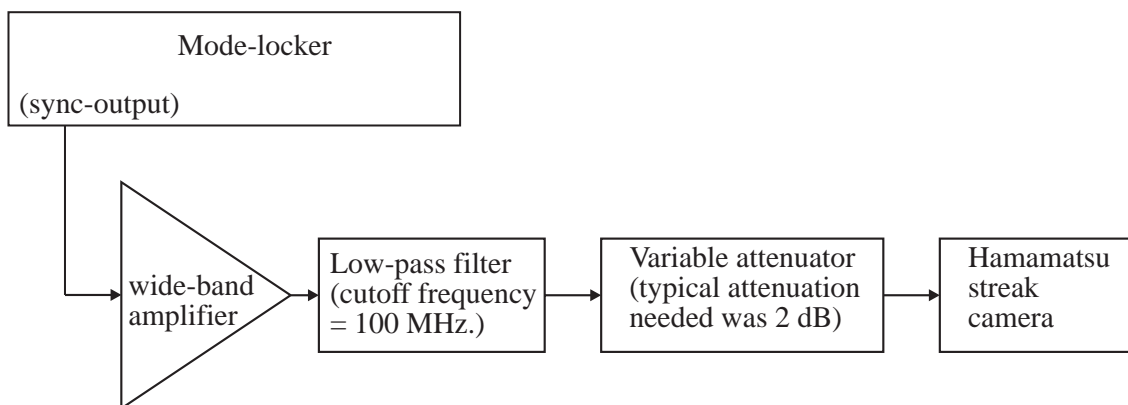
Figure 6.5 Diagram showing the original connection of the streak camera to the laser mode-locker. It was desirable to disturb this as little as possible.

Some thought was needed on how best to obtain the RF drive for the SRD from the laser mode-locker, because the *sync* output of the mode-locker was being used to trigger the streak camera. The original method of connection between the mode-locker and streak camera is shown in figure 6.5. Since optimum performance of the streak camera was very dependent on the signal from the laser mode-locker, it was decided to alter this as little as possible. Also, since the cross-correlator was expected to be used simultaneously with the streak camera, a switch was not possible.

For the original streak camera drive, the output of the mode-locker was first amplified with a wide band amplifier, with a gain of approximately 30 dB, then filtered with a low-pass filter with a cutoff frequency of 100 MHz, to remove the harmonics above 82 MHz. The signal was then passed through a variable attenuator, to allow the drive level for the streak camera to be set correctly. Typically 1 or 2 dB of attenuation was required on the attenuator. This circuit

```
┌─────────────────────────────────────┐
│           Mode-locker                │
│                                      │
│ (sync-output)                        │
└─────────────────────────────────────┘
```

wide-band amplifier

Low-pass filter cutoff frequency = 100MHz.

20dB directional coupler

(in)   (out)

(coupled)

Variable attenuator (typical attenuation needed was 2dB)

Hamamatsu streak camera

Approximately sinusoidal waveform at 82MHz, -12.6dBm, synchronised with the Ti:sapphire laser output. This signal is used by the cross correlator.
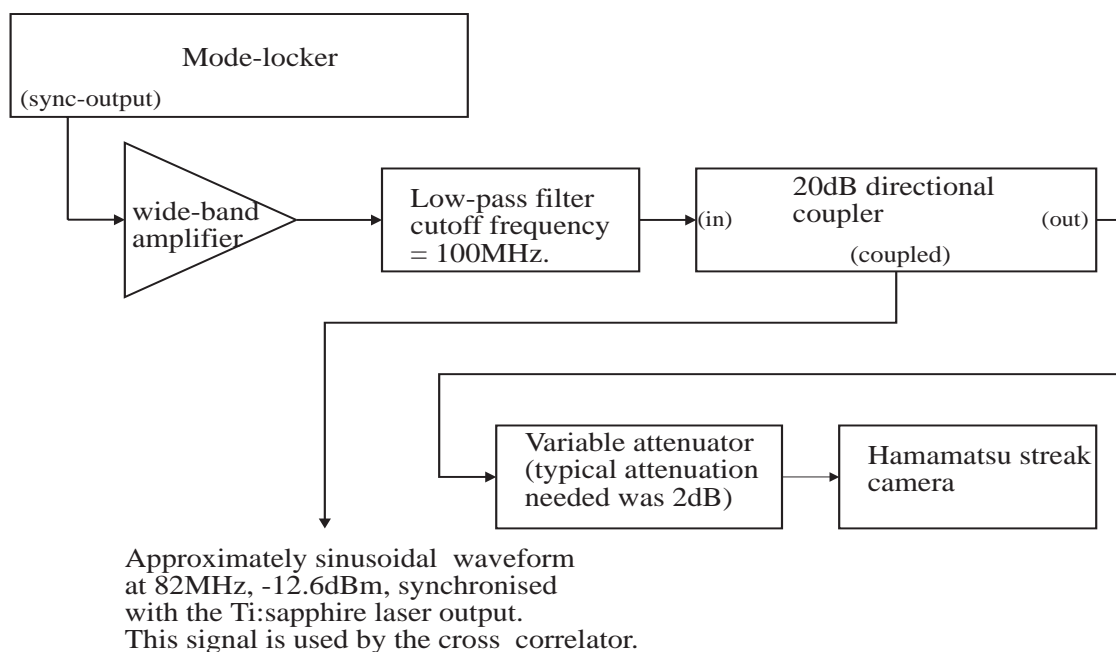
Figure 6.6 Diagram showing the method of connecting both the cross-correlator and the streak camera simultaneously to the laser mode-locker.

was developed by Dr. Mark Cope.

The revised circuit providing both streak camera and SRD drive is shown in figure 6.6. The output of the mode-locker was first amplified by the wide band amplifier. It was then passed through the original low-pass filter. The sine wave was then passed through a 20 dB directional coupler, which has an insertion loss of less than 0.3 dB. The main output then passes to the variable attenuator and finally the streak camera. The level of the signal at the input of the attenuator was very similar to that at the same position in figure 6.5, since the only extra loss is approximately 0.3 dB for the directional coupler. The level of signal from the coupled port of the directional coupler was -12.6 dBm, which was used in all subsequent experiments to drive the SRD. Around 7 dBm was fed to the attenuator for the streak camera.

6.4   Test oscillator design.

For the purposes of bench testing, a small 82 MHz oscillator was built to provide an output power of -12.6 dBm, since at the start of this project, there was no RF signal generator in the laboratory - fortunately this has now changed! An output level of -12.6 dBm was chosen as this was the output power available from the directional coupler attached to the laser mode-locker. The circuit diagram of the test oscillator is shown in figure 6.7. The first BFY90 NPN
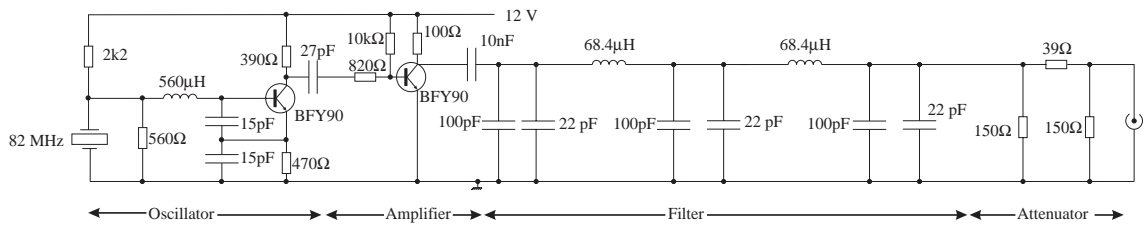
Figure 6.7 Circuit diagram of the 82 MHz test oscillator. This has an output of -12.6 dBm at 82 MHz and the second harmonic at -48 dBc (-54 dBm).

transistor was used as an oscillator, using a 82 MHz 5th overtone crystal. The design for the oscillator was adapted from one described in an amateur radio publication[126]. The second BFY90 acts as a simple common-emitter amplifier. A 5-pole Chebychef low-pass filter with a 90 MHz cutoff frequency and a 3 dB ripple follows this to reduce any harmonics. Finally an attenuator of around 6.4 dB was added to reduce the output to the -12.6 dBm required and to ensure the output impedance was close to 50 Ω. The 2nd harmonic was then the most significant at 48 dB below the fundamental. All higher harmonics were further down, so the design provided an ideal source for test purposes. Since the oscillator did not provide any voltage stabilisation, it was powered from a well regulated supply to prevent frequency drift.

6.5   Details of an early prototype of the electronic cross-correlator.

The RF signal leaving the directional coupler at -12.6 dBm, was amplified to +27 dBm to drive the SRD - a gain of approximately 40 dB. Since the time of the pulses generated by a step recovery diode are fixed relative to the timing of the drive signal (occurring at the most negative part of the RF drive cycle - see figure 5.2), it was thought that the timing of the pulses could be altered by varying the phase of the RF drive signal. A system as shown in figure 6.8 was considered as the basis for a totally electronic cross-correlator. This was subsequently constructed. Figure 6.8 also includes some small diagrams showing the shape of the signals at various points in the system. A mode-locked laser (see photograph in figure 6.40) was used to generate light for interrogating the tissue phantom. The tissue phantom, consisted of a block of epoxy based material loaded with glass micro-spheres to cause scattering. On opposite sides of the phantom were attached two FC-PC optical connectors, by drilling and tapping with an M2 thread the phantom material. The mode-locker, which has an RF output signal synchronised to the laser pulses, was suitably filtered and amplified to provide a drive signal to the Herotek GC82RC step recovery diode (Herotek Inc., 222N Wolfe Road, Sunnyvale, CA 94086, USA). The phase of the sine wave reaching the SRD was altered by an electronic phase shifter model PS-3-82 MHz (R&K, Japan). The pulses generated by the SRD were then used to gain-modulate
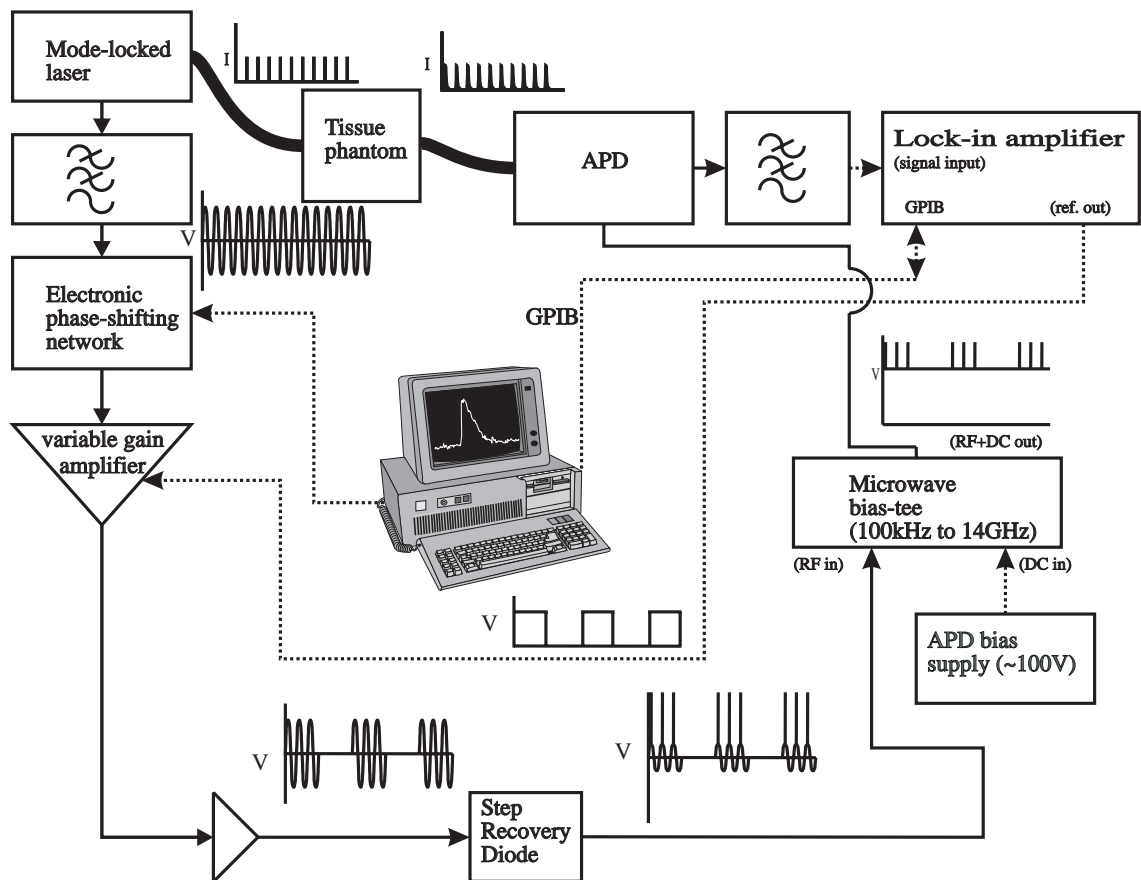
Figure 6.8 An early version of the cross-correlator that had an excessive temporal response of approximately 2.5 ns FWHM.

the APD detecting the light from the tissue phantom, using a Picosecond Pulse Labs 5530A bias–tee (Picosecond Pulse Labs Inc., P. O. Box 44, Boulder, Colorado, 80306, USA) to add the pulses to the high voltage generated by a low current (8 mA maximum) 0-500 V DC power supply, model PMT-05CP-1 (Bertan High Voltage, 121 New South Road, Hicksville, NY 11801, USA). The internal oscillator of the lock-in amplifier provided a signal to gain-modulate a variable gain amplifier type ZFL-1000GH (Minicircuits, P.O. Box 350166, Brooklyn, New York, USA), used to provide RF drive to the SRD. By switching the RF drive to the SRD from zero to 500 mW, the pulses from the SRD could be turned on and off. A brief specification of the amplifier is given in Appendix 1. Some readers may question the use of a variable gain amplifier, rather than an RF switch, using pin diodes or similar, for the purpose. It was found that some automatic gain control (AGC) was needed to stabilise the RF drive to the SRD. A variable gain amplifier was therefore used, rather than a switch.

The APD was fibre coupled in this instance. The *green* APD (see Table 6.1 on page 136) was fitted to an FC-PC optical connector by Access Pacific Ltd (Bedford, UK). A bull lens

was used to focus the light from the connector onto the APD. The responsivity of the photodiode at a bias voltage of 4.44 V, was measured by Access Pacific as 0.54 A/W at 810 nm with an optical power of -20 dBm (10 µW) leaving the FC-PC connectorised 100 µm core optical fibre.

A 16 MHz IBM compatible 80386 computer, with 8 Mb of RAM, had two cards inserted into its expansion slots to aid fast data acquisition. The first, a model NI-488.2™ GPIB card (National Instruments, 6504 Bridge Point, Parkway, Austin, TX 78730-5039, USA) was used to control the lock-in amplifier and collect data. The second card, an Amplicon Liveline model PC24AT (Amplicon Liveline Ltd, Centenary Industrial Estate, Brighton, East Sussex, BN2 4AW, England) contained a digital to analogue converter whose output voltage was fed to the phase shifter hence controlling the phase shift produced. The maximum output voltage of the DA card was specified at 10.0 V, but the R&K phase shifter required 10.2 V. Rather than build a DC amplifier with a gain of 1.02, the maximum output capacity of the DA card was increased to 10.25 V by changing a resistor on the board.

The computer's mother-board was later changed from one using an Intel 16 MHz 80386 to one using an Intel 33 MHz i486, as such a board became available. The extra speed made no difference to data acquisition, but made the computer boot faster, and made it possible to use a graph plotting package (Sigmaplot for DOS), without an excessive delay.

The first fully electronic prototype of the cross-correlator is shown in figure 6.8. Note that the full voltage output of the SRD, some 8 V peak, was applied to the bias-tee and so 8 V modulation was put onto the APD. This seemed most logical, since it has already been shown (figures 3.14 and 3.15) that the greater the modulation voltage, the greater the increase in gain and so the greater the response of the system. The theory used in obtaining the graphs in figures 3.14 and 3.15 did not predict any upper limit to the modulation voltage, other than that when the sum of the DC bias and the modulation voltage caused avalanche breakdown in the APD.

Note also the following differences between figures 6.8 and 6.3:

1) The 9:1 fibre optic splitter is no longer needed, as none of the laser output is required to generate the reference pulses.

2) The chopper wheel has been removed and replaced by an amplifier whose gain can be varied in a switch-like manner.

3) Instead of the chopper wheel providing a reference input to the lock-in amplifier, the internal oscillator of the lock-in amplifier is used to switch the gain of the variable gain amplifier.

4) The reference APD and its associated high voltage bias supply has been removed and replaced by an SRD for the generation of short pulses.

5) The mode-locker is now electrically connected to both the cross-correlator and the laser.

6) The mechanical delay line has been removed and replaced by an electronic phase shifting network.
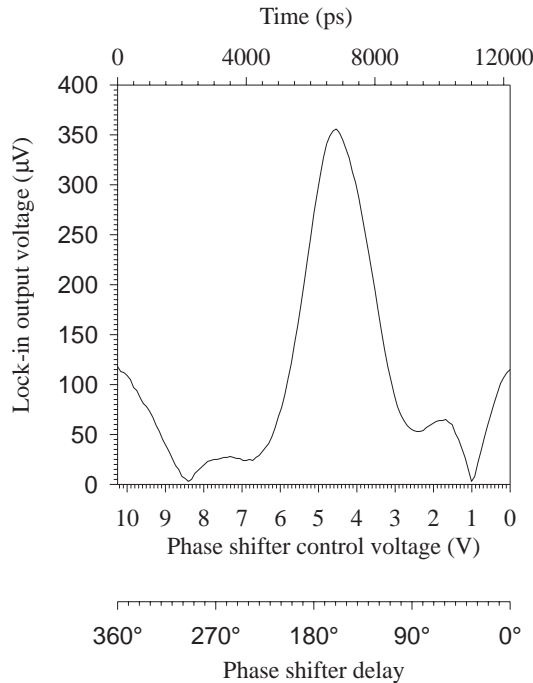
Figure 6.9 Data produced with the cross-correlator shown in figure 6.8. APD gain = 100. Full SRD modulation voltage on the APD.

To test the performance of the complete cross-correlator, the system was set up as shown in figure 6.8, with the laser output being fed directly to the correlating APD rather than through a tissue phantom. The bias voltage on the *mauve* APD (see table 6.2) was 109.4 V, giving a gain of 100. The laser was attenuated to a level of 1.5 mW, although the actual power intercepted by the APD would have been much less than this. Since no scattering medium was placed between the laser and the APD, the temporal profile of the light at the detector was very narrow (around 2 ps FWHM) and so the system response alone was measured. The results are shown in figure 6.9. Data collected at a lower DC bias voltage for the APD, giving a lower gain, is shown in figure 6.10, where the signal is of a lower level and more noisy. Note the temporal response is 3 ns FWHM. This is about the same value of the longest TPSFs likely to be encountered and so is completely unsuitable for the task.

It was soon found that the temporal profile could be improved dramatically by fitting a 20 dB attenuator to the output of the SRD. This has two effects.

1) The signal driving the APD is reduced by 20 dB.

2) The impedances of the system are much better matched, so reflections are reduced.

Figure 6.10 Impulse response of the system in figure 6.8, using a DC bias voltage of 80.1 V for the APD.

The system impulse response of the revised system is shown in figure 6.11, where we see there is a short pulsed section, but the system impulse response is dominated by a residual component that looks like a rectified sine wave, although the two halves of the rectified sine wave are of slightly different lengths.
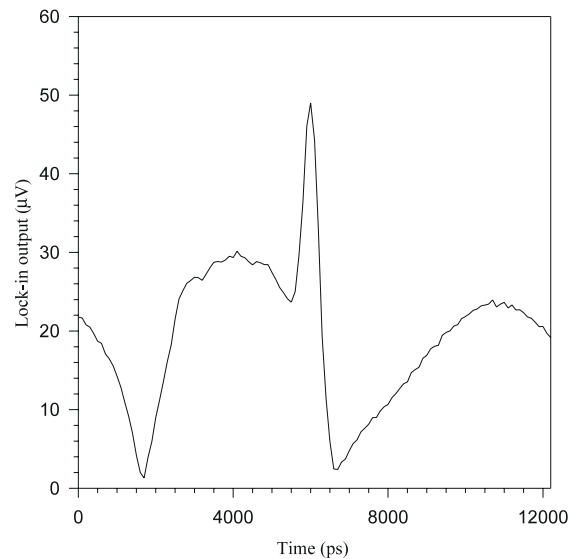
Figure 6.11 System response obtained when the SRD amplitude to the APD was by reduced by 20 dB. System connected as in figure 6.8, but with an attenuator after the SRD.

6.6 The reasons for the appearance of a rectified sine wave in the system response.

An oscilloscope view of the output of the SRD, shown in figure 6.12, shows the reasons for the appearance of a rectified sine wave signal in figure 6.11. This has been attenuated by 20 dB, to avoid damage to the input of the oscilloscope, which is 5 V maximum. Clearly visible, but fairly faint due to the risetime, is the short pulse. Also visible is also a large residual component at 82 MHz due to sine wave feed-through. Since the lock-in amplifier's magnitude output was
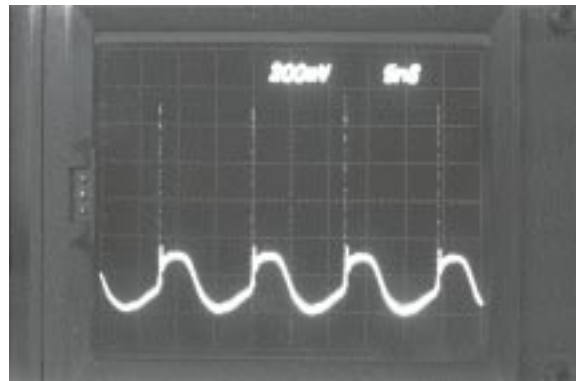
Figure 6.12 An oscilloscope photograph of the SRD output, attenuated by 10 dB. Note both the wanted pulse and the unwanted components. 5 ns/div, 200 mV/div.

used it is always positive and is so effectively rectifying its input. Hence the approximately sinusoidal component at 82 MHz on the SRD output causes an approximately $(sine)^2$ component at the output of the R component of the lock-in amplifier. The reason why the two sides of the $(sine)^2$ waveform were of unequal length (time), was believed to be because the phase shifter has a slight (specified up to 15 %) non-linearity of the phase shift with respect to the control voltage.
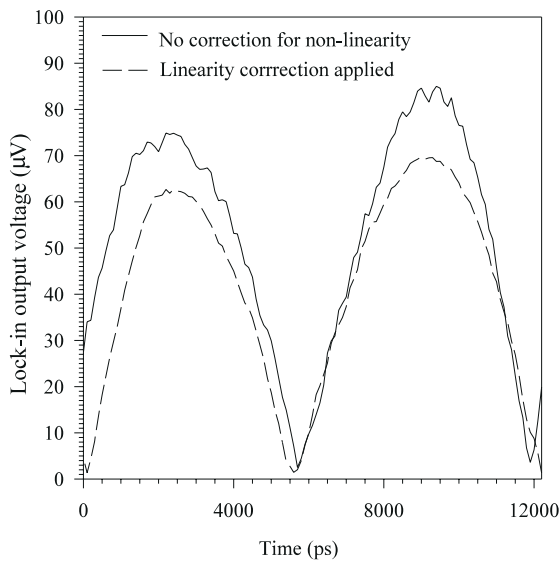
Figure 6.13 System response with the SRD removed. The drive to the APD was a sine wave of -5 dBm. This shows the effect of timebase linearity corrections.

To check this, the SRD was removed and the sine wave drive that usually went to the input of the SRD was fed directly to the APD. The data collected is shown as the solid line of figure 6.13. By measuring the time between the troughs of the cycle, we can check the timebase linearity. The first half started at 5.25 ns, ending at 11.9 ns, a half-period of 6.65 ns. The second half is then 12.2 - 6.65 = 5.55 ns long. Each half, should of course, be 12.2/2=6.1 ns in length.

The manufacturer of the phase shifter (R&K of Japan) supplied a graph of the variation of phase shift as a function of the control voltage. This is shown in figure 6.14. This was fitted to a 4th order polynomial, using Sigmaplot^{TM} a commercial scientific graph plotting programme, so that an expression for the voltage V (Volts) necessary for a particular phase shift ϕ (degrees) could be determined. The expression obtained is given in equation 6.8 below. Other degrees of polynomial fit (from 2nd to 5th order) were tried, but the 4th order fit gave the smallest RMS error.

Figure 6.14 Variation of phase shift as a function of the control voltage on the phase shifter. This is the manufacturer's test data.

$$V=0.0314\phi \ +2.547\times10^{-5}\phi^2 \ -2.020\times10^{-7} \ \phi^3 \ +3.004\times10^{-10} \ \phi^4 \qquad \textbf{6.8}$$

Using this modification in the software driving the DAC, the data shown in the dashed line of figure 6.13 was obtained. This data has a minimum at 5.7 ns and another at 11.9 ns, so the half period is 6.2 ns, not far from the 6.1 ns we would expect. A more thorough investigation of the time-base linearity was later performed, and is discussed in section 6.13.

It can also be seen that the amplitude of the two half cycles in figure 6.13 are different.

The phase shifter was known from the manufacturer's test data to have some variation of insertion loss with control voltage, as reproduced in figure 6.15, so an automatic gain control (AGC) system was built to keep the drive to the SRD constant, irrespective of small variations in insertion loss when the phase shifter was altered.

Figure 6.15 Manufacturer's test data showing the variation of loss through the phase shifter as a function of the control voltage.

6.7   AGC system.

Since the phase shifter has a loss that depends on the phase shift selected, as shown in figure 6.15, an AGC system was developed to keep the drive to the SRD constant. The ZFL-1000GH (Minicircuits, P.O. Box 350166, Brooklyn, New York, USA) variable gain amplifier, already in use to effectively switch on and off the drive to the SRD, was used to stabilise the drive at +27 dBm during the periods when the RF was applied to the APD. It was initially hoped to use the same amplifier as both a switch and an AGC control element, but this proved difficult to achieve, so a GaAs RF switch type ZYSWA-2-50DR (Minicircuits) was therefore used to switch the RF on and off to the SRD and the variable gain amplifier used to ensure the drive was at the correct level of +27 dBm. A brief specification of the GaAs switch is given in Appendix 1. A GaAs switch was chosen in preference to a pin diode switch, since the much higher maximum frequency of operation (5 GHz) of the GaAs switch would allow the SRD's output to be switched directly, rather than modulate its RF input, with little degradation in pulse performance, although this has currently not been investigated. The readily available pin

Figure 6.16 Graph showing the variation of gain of the Minicircuits ZFL-1000GH variable gain amplifier as a function of control voltage.

diode switches had a maximum frequency of 2 GHz, which would have prevented their use for

such tests.

The manufacturer did not give very precise data on the variation of gain with control voltage, so this was first measured and is shown in figure 6.16. A small fraction of the RF drive to the SRD was measured with an HP 5082-2835 Schottky diode configured as an RF power sensor. The resulting DC voltage was amplified using a OP-07 operational amplifier and compared to a DC reference voltage. A second OP-07 operational amplifier was then used in a proportional controller. If the detected power decreased, the output of the proportional controller decreased, which increased the gain of the RF amplifier. A block diagram is shown in figure 6.19.

6.8   **Band-pass filter.**
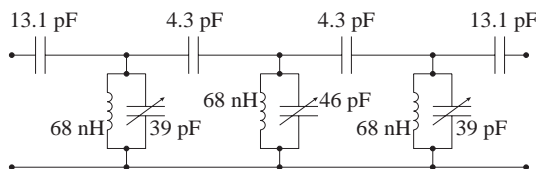
Figure 6.17 82 MHz narrow-band, band-pass filter. The component values shown are the theoretical values, not those actually used.

Since the output of the laser mode-locker was not a clean sine wave, but had components at 41 MHz, 82 MHz, .. it was decided to add a couple of band-pass filters, centred on 82 MHz, to ensure the drive to the SRD was as close as possible to the sine wave it required. The public domain programme FDS was used to design a filter. This was designed to use 68 nH inductors, since this is a standard value. The design is shown in figure 6.17.

Since the filter needed non-preferred values of capacitors, it was decided to make all the series capacitors (13.1 pF and 4.3 pF) from a two preferred values in parallel. Two 2.2 pF capacitors in parallel were approximately 4.3 pF and 4.7 pF in parallel with 8.2 pF was approximately 13.1 pF. The inductors were hand wound, since this was faster than ordering them, although 68 nH inductors are readily available commercially. The formula for the inductance was found in an amateur radio publication[126]. To allow for component tolerances, the three parallel resonant circuits were tuneable by use of 68 pF variable capacitors. These were optimised iteratively on the spectrum analyser. Figure 6.18

Figure 6.18 Plot of the theoretical performance (solid line) of the band-pass filter and the actual performance (filled dots).

shows the actual measured performance, along with the theoretical performance if the perfect components of the exact value were obtainable. The measured performance is adequate, giving more than 45 dB of attenuation to both the 41 MHz and the 164 MHz signals, while having only 0.66 dB of insertion loss. These two band–pass filters were added to the system in the two places shown in figure 6.19.

Two quarter-wave stubs were also made in an attempt to remove the unwanted components at 41 MHz. These consisted of a length of approximately 1.22 m of 50 Ω line, open at the far end, placed in parallel with the drive to the SRD. Unfortunately, the notch obtainable was only 20 dB - far worse than the 59 dB obtained with the LC filter. The poor performance of this was probably due to the loss in the rather thin (2.6 mm OD) and hence lossy RG174A/U coaxial cable used for the purpose. Using a thicker cable would have probably improved matters, but would have taken considerable space, so this was not pursued any further.
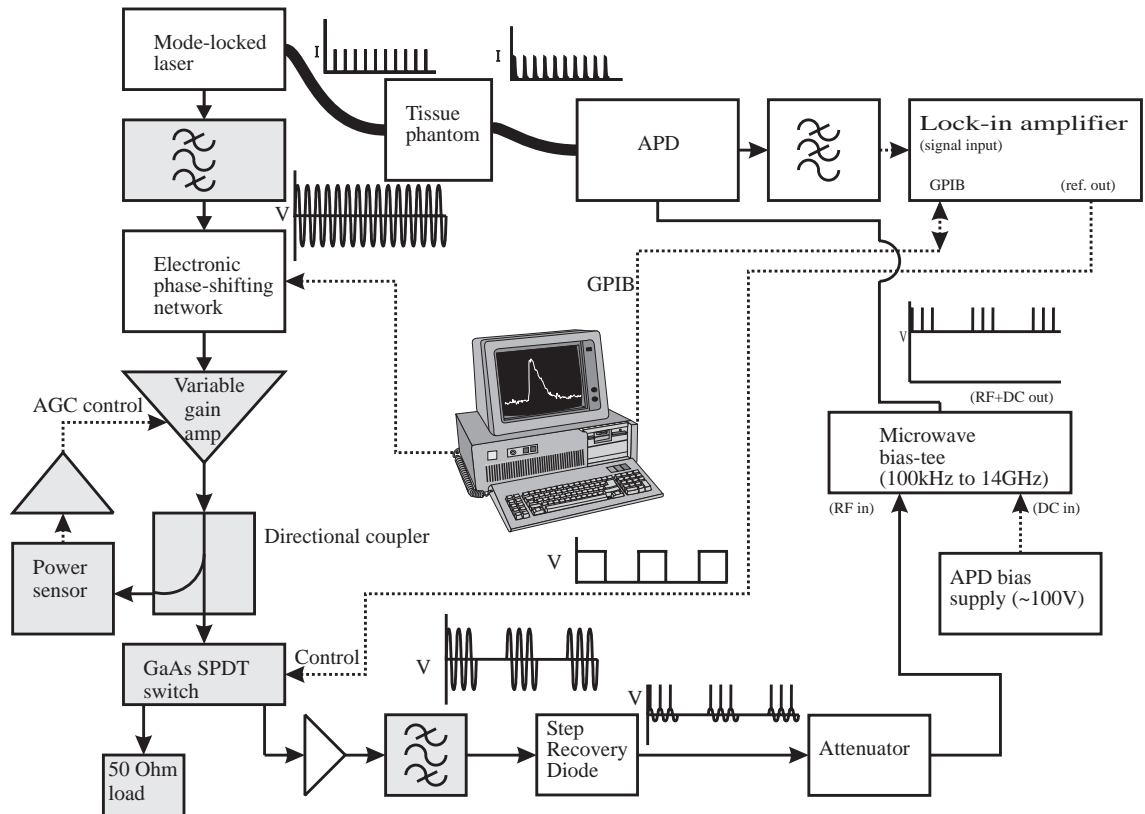
Figure 6.19 Revised system showing the addition of filters and an AGC system. New components are shown shaded.

6.9   Improving the system response.

Having obtained a clean 82 MHz drive to the SRD by use of the band–pass filters, two

ways were considered to correct for the non-ideal behaviour seen in figure 6.11. The first was to leave the SRD output unaltered, but to correct its effects in software. The second was to alter the shape of the pulse from the SRD pulse generator, so it more closely resembled the required waveform, which is a series of delta pulses. Both were investigated:

6.9.1   Improving the system response by software post-processing.

Figure 6.20 shows some data collected after the timebase had been corrected as described earlier. Note the spurious (sine)^2 components. To remove the effects, a signal of the form:

$$A\sin(2\pi.\ 82x10^{6}+\phi)^2 \qquad\qquad\textbf{6.9}$$
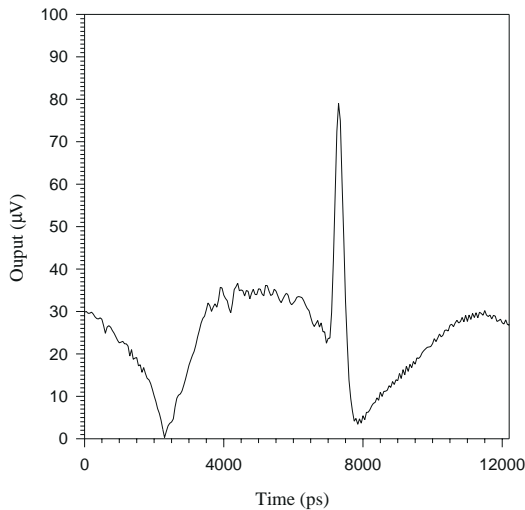
Figure 6.20 Data collected to show the system response. Timebase correction has been applied.
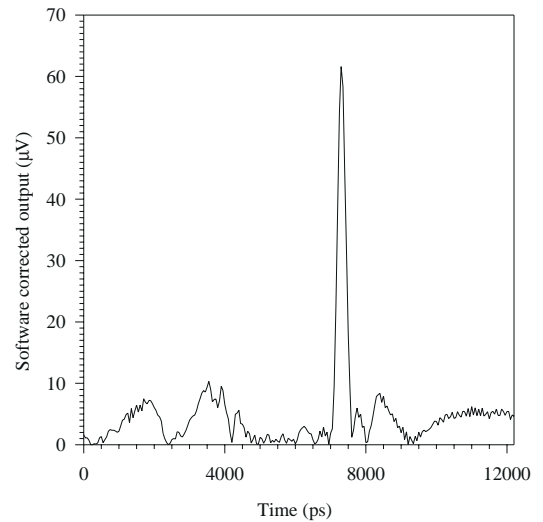
Figure 6.21 Data from figure 6.20, that has been corrected to remove the unwanted sinusoidal output of the SRD.

was added to the output of the lock-in amplifier in software. The peak amplitude of the signal A and the phase ϕ, were both adjusted using a least squares fitting algorithm, to give the minimum mean-square component outside of the range 6900 to 7600 ps. The effect of this is to remove the majority of the effects of the residual SRD output, as shown in figure 6.21. An expanded view is shown in figure 6.22. This procedure has not worked completely however, as the unwanted SRD output shown in figure 6.12, is not exactly a sine wave, so correcting by adding a pure sine wave can not be expected to work perfectly. These results were presented at the SPIE conference on in San José, 1995[127].
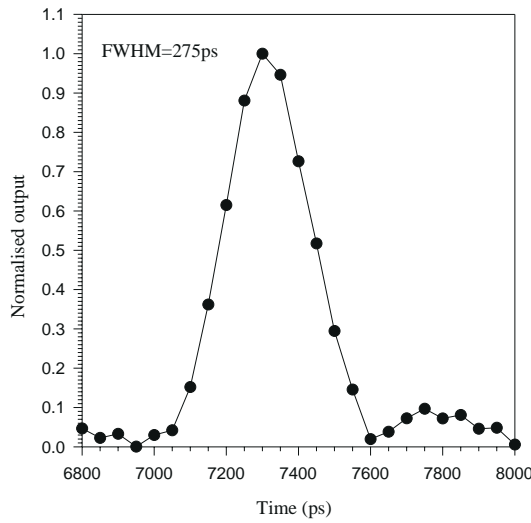
Figure 6.22 Expanded view of figure 6.21, showing the individual data points collected at 50 ps intervals. Impulse response is 275 ps FWHM.

6.9.2   Improving the system response by adding additional hardware.

Since the software method of improving the system response was not satisfactory, other methods were tried to improve the system performance. These included various attempts to improve the output of the SRD.

6.9.2.1   Removing the residual output by adding a sine wave in hardware.
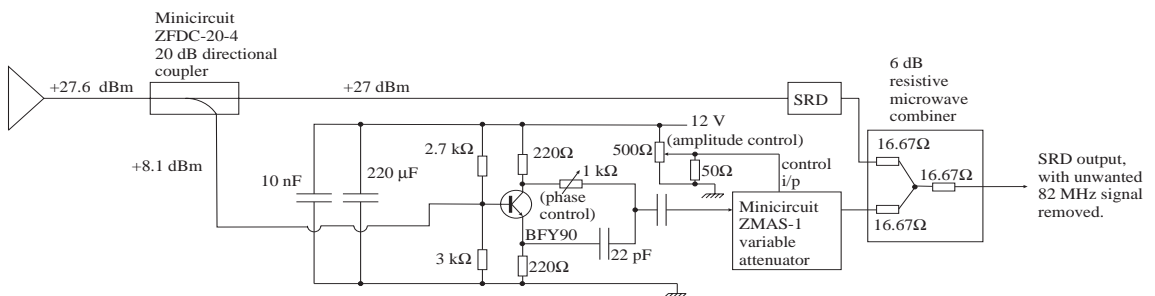
Figure 6.23 Circuit diagram of the variable phase shifting network used to combine the output of the SRD with a 82 MHz sine wave of variable amplitude and phase.

The software post-processing detailed in section 6.9.1, removed the effects of what was considered a sine wave in software. It was then realised that theoretically, the TPSF could only be recovered from the measurements if the modulation signal is an even function of time, as was shown on page 76. It was a natural progression to try to remove the unwanted 82 MHz output

from the SRD pulse generator before it reached the APD, so the SRD pulse became an even function of time. Approximately 1 % of the 82 MHz drive power to the SRD was sampled using a 20 dB directional coupler. This was put into the variable phase shifting network shown in figure 6.23, which allowed independent control of the amplitude and phase of its 82 MHz output. A Minicircuits ZMAS-1 pin diode based variable attenuator was used to adjust the output level, using the
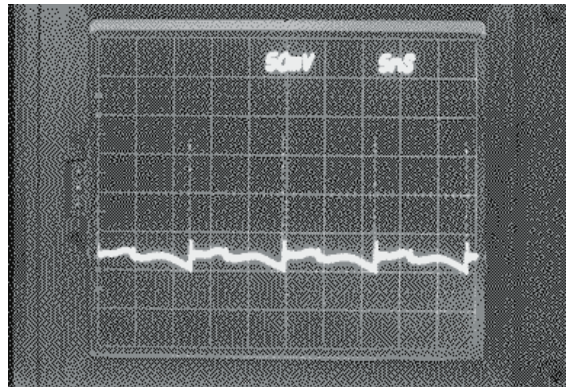
Figure 6.24 Output of the SRD pulse generator after applying a sine wave at 82 MHz in anti-phase to the unwanted output.

500 Ω variable resistor, which altered the current into the pin diodes. The 1 kΩ resistor allowed the phase to be changed by connecting the input of the attenuator towards either the collector or the emitter of the BFY90 transistor. This allowed a sine wave of the correct amplitude and phase to be added to the SRD output to cancel the unwanted 82 MHz component normally present. Since the phase shifter in figure 6.23 only provided approximately 100° of phase control, it was necessary to use appropriate lengths of coaxial cable to provide a course phase adjustment. To combine the outputs from the SRD and the 82 MHz sine wave, a high frequency (18 GHz) 6 dB power resistive splitter/combiner was used. Although only an 82 MHz signal was being combined, which is relatively low frequency, a high frequency combiner was necessary to avoid degrading the SRD output, which has components extending to at least 26 GHz. Figure 6.24 shows the SRD output when the amplitude and phase of the signal were adjusted to provide the minimum unwanted output as seen on the oscilloscope.

Using this method, it was possible to cancel the 82 MHz component such that it was below the noise floor of the spectrum analyser. This corresponds to a signal level of less than –120 dBm. Unfortunately, there were two problems with this method. Firstly, the cancellation was not very stable, due to phase changes with time and temperature. Secondly, since there were also components at 164 MHz (2 x 82 MHz), 246 MHz (3 x 82 MHz) and higher frequencies of too large an amplitude, it would have been difficult to adjust them all to the correct level, without having numerous trimmers and a potentially difficult system to align. Eventually, a different approach was found to give better results.

6.9.2.2   Dissipating the unwanted 82 MHz component into a load resistor.

One option considered was to place a series RLC circuit across the output of the SRD. This would have an impedance of R at resonance, so by correct choice of R, the fundamental sinusoidal component could be reduced by the required amount. Since the source and load are both 50 Ω, we can express the attenuation as a function of R:

$$\text{Attenuation(dB)}=20\log_{10}\left(\frac{50+R}{R}\right) \qquad \textbf{6.10}$$

However, this method proved unsuitable, as it is impossible to avoid oscillatory behaviour in the series RLC, while still providing acceptable filtering, as we can see theoretically. The roots of the equation describing the RLC circuit depend upon the magnitudes of R, L and C[128].

$$\left(\frac{R}{2L}\right)^2 > \frac{1}{LC} \qquad\qquad \text{Overdamped} \qquad \textbf{6.11}$$

$$\left(\frac{R}{2L}\right)^2 = \frac{1}{LC} \qquad\qquad \text{Critically damped} \qquad \textbf{6.12}$$

$$\left(\frac{R}{2L}\right)^2 < \frac{1}{LC} \qquad\qquad \text{Underdamped (oscillatory)} \qquad \textbf{6.13}$$

At critical damping, we can obtain the following two equations for L and C

$$L=\frac{R}{4\pi f} \qquad \textbf{6.14}$$

$$C=\frac{4L}{R^2} \qquad \textbf{6.15}$$

where f is the resonant frequency in Hz. Setting R=5.5 Ω, which provides 20 dB attenuation to the 82 MHz signal (the approximate amount of attenuation required), we find L=5.33 nH, and C=705.8 pF. A graph of the theoretical attenuation of the network is
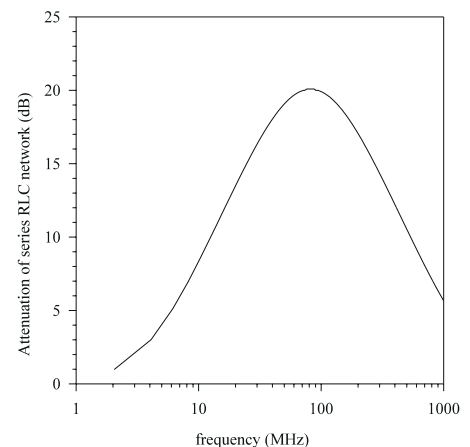
Figure 6.25 Theoretical attenuation curves of an RLC series tuned circuit, with R=5.5 Ω, L=5.33 nH and C=706 pF. This is critically damped.

shown in figure 6.25. This has the correct 20 dB attenuation at 82 MHz, but its loss at 164 MHz, where we want little attenuation, is some 18 dB. It is not possible to change this, as an increase in circuit Q would make the behaviour oscillatory. Hence it is impossible to satisfy both the time and frequency domain requirements for this circuit simultaneously.

6.9.2.3   Removing the unwanted SRD output with a high-pass filter.

Since the main unwanted component of the SRD pulse generator was at 82 MHz, with other components up to about 300 MHz having excessive amplitude, it was decided to attempt to pass the output of the SRD through a 300 MHz high-pass filter. This had been considered earlier, but was rejected since the problems of making a filter with a 300 MHz cutoff frequency and that did not have unwanted resonances somewhere below the maximum output frequency of the SRD were considerable.

6.9.2.3.1   300 MHz cutoff high-pass Butterworth filter.

A 3rd order 300 MHz high-pass Butterworth filter was designed using the Public Domain filter synthesis programme called FDS[*1]. The circuit diagram of the filter is shown in figure 6.26. The frequency response was checked on the excellent low cost ($10) RF and microwave
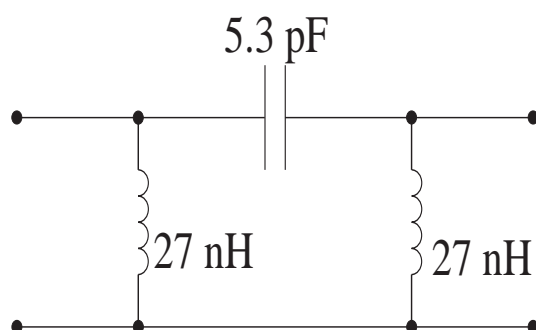
Figure 6.26 Circuit diagram of the 300 MHz high-pass Butterworth filter.
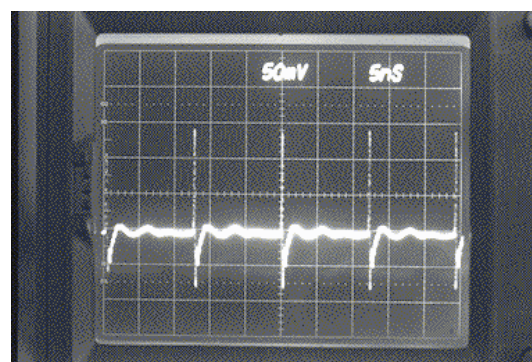
Figure 6.27 Output of the SRD comb generator after passing through a Butterworth high-pass filter with a 300 MHz cutoff frequency.

computer aided design package *Puff* (sold by the Department of Electrical Engineering, California Institute of Technology, Pasadena, California, 91125, USA). This showed the desired frequency domain behaviour, having the -3.01 dB point at 300 MHz. However, the measured time domain performance was poor, with

---

[*1] FilDes (FDS) was a public domain programme produced by B. Lombardi in 1993. Unfortunately, it no longer seems to be available at any Internet site.

unacceptable ringing, as shown in figure 6.27.

6.9.2.3.2   300 MHz cutoff high-pass Bessel filter.
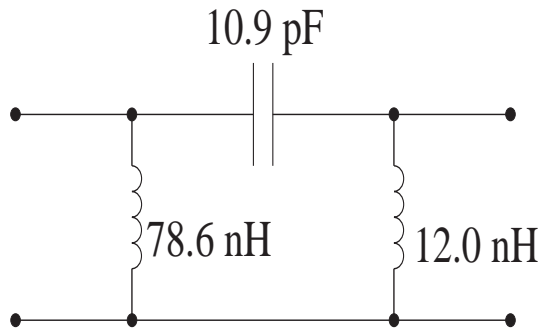
Figure 6.28 Circuit diagram of a Bessel high-pass filter with a cutoff frequency of 300 MHz.
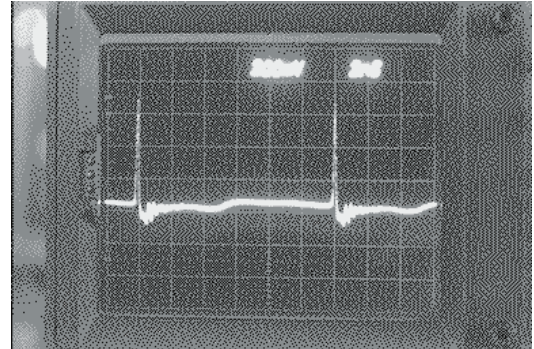
Figure 6.29 Time domain performance of a Bessel filter.

The Butterworth filter had reasonable frequency domain performance, although it is not as good as either the Chebyshev or Elliptic filter. However, its performance in the time domain is not as good as the Bessel filter, which has poorer frequency domain performance, but a better time-domain response. Hence a Bessel filter was constructed using low-pass filter data from Williams[129] and converted into a high-pass filter using the method described below. The circuit diagram of the filter is shown in figure 6.28. The time-domain performance of this filter, shown in figure 6.29 was better than the Butterworth filter (figure 6.27), but was still not acceptable, due to ringing.

6.9.2.3.3   Low-pass to high-pass filter transformations.

Although computer programmes are available that can design high-pass filters, designs for high-pass filters are rarely seen in filter books. Usually only low-pass designs are given, which have the following characteristics.

1) Low-pass filter with a cutoff frequency f_c of 1 radian/second - ie $f_c = 1/(2\pi) = 0.159$ Hz.

2) Input and output impedance is 1 Ω.

A set of rules is then used to scale the low-pass filter to have the required cutoff frequency and termination impedances. To convert to a high-pass filter, the inductors are swapped for capacitors and the capacitors for inductors, and all components scaled by the required constant, which again depends on the cutoff frequency and termination impedances. Unfortunately, this method of designing high-pass filters from low-pass filters does not work satisfactorily if one is interested in preserving the time domain performance - only the frequency domain characteristics are

preserved in the transformation. In order to make high-pass filters that have good time domain characteristics, it is necessary to follow a different method[130], but this too failed to work satisfactorily, as result of unavoidable unwanted resonances in the components. Unfortunately, the large inductor values necessary for a filter with a cutoff frequency in the VHF/UHF range, can not be made to have acceptable microwave performance.

6.9.3 Improving the system response by use of a diode in series with the SRD output.
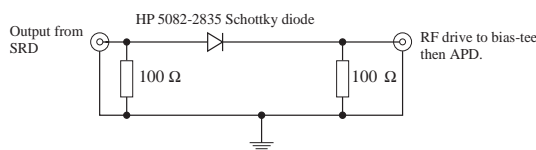
Figure 6.30 Method of reducing unwanted SRD output using a single series connected HP 5082-2835 high-speed Schottky diode.

Figure 6.31 Output of the pulse generator after insertion of a single series connected HP 5082-2835 Schottky diode in series with the SRD output. 5 ns/div, 200 mv/div.

Figure 6.32 Method of reducing the unwanted SRD output using 4 HP 5082-2835 high-speed Schottky diodes.
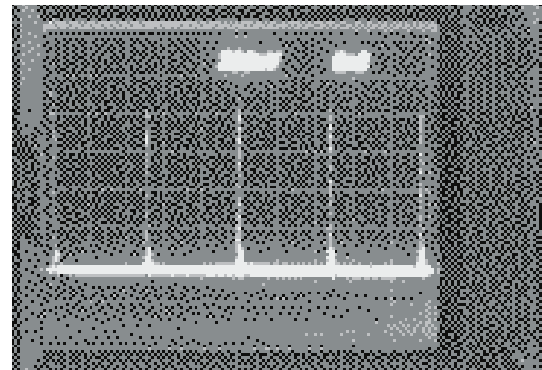
Figure 6.33 Output of the SRD pulse generator after fitting four HP 5082-2835 Schottky diodes in series with the output. 5 ns/div, 200 mv/div.

The SRD output consisted of a low amplitude unwanted component and a large amplitude wanted component. The next method tried to remove the unwanted SRD output was to pass it through a fast series connected Schottky diode, type HP 5082-2835 (Hewlett Packard, P.O. Box 10301, Pal Alto, CA 94303-0890, USA), as shown in figure 6.30. To ensure that the diode system was as well matched to 50 Ω as possible, a 100 Ω resistor was added at each end. If the load is terminated properly in 50 Ω, the input impedance varies from 25 Ω to 100 Ω, assuming the diode is a perfect switch. Hence the standing wave ratio (SWR) is 2:1. The output

is shown in figure 6.31, where it can be seen that some of the sinewave feedthrough has been removed, but a small amount still remains. The turn–on voltage of these HP diodes is approximately 500 mV, so as the unwanted output is 2 V peak, 4 diodes in series should have the required turn-on voltage, so that only the pulse component of the SRD output gets through. Hence the system shown in figure 6.32 was constructed, where four diodes rather than one were used. The output is shown in figure 6.33, where it can be seen that most of the unwanted output has been removed. Figure 6.34 shows the pulse on an expanded time-base of 200 ps/division.
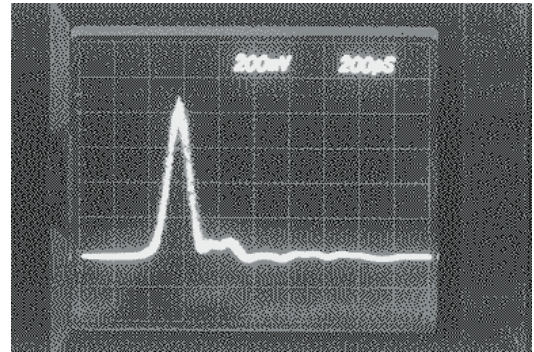
Figure 6.34 Identical data to that in figure 6.33, but viewed with a faster oscilloscope timebase. 200 ps/div, 200 mv/div.
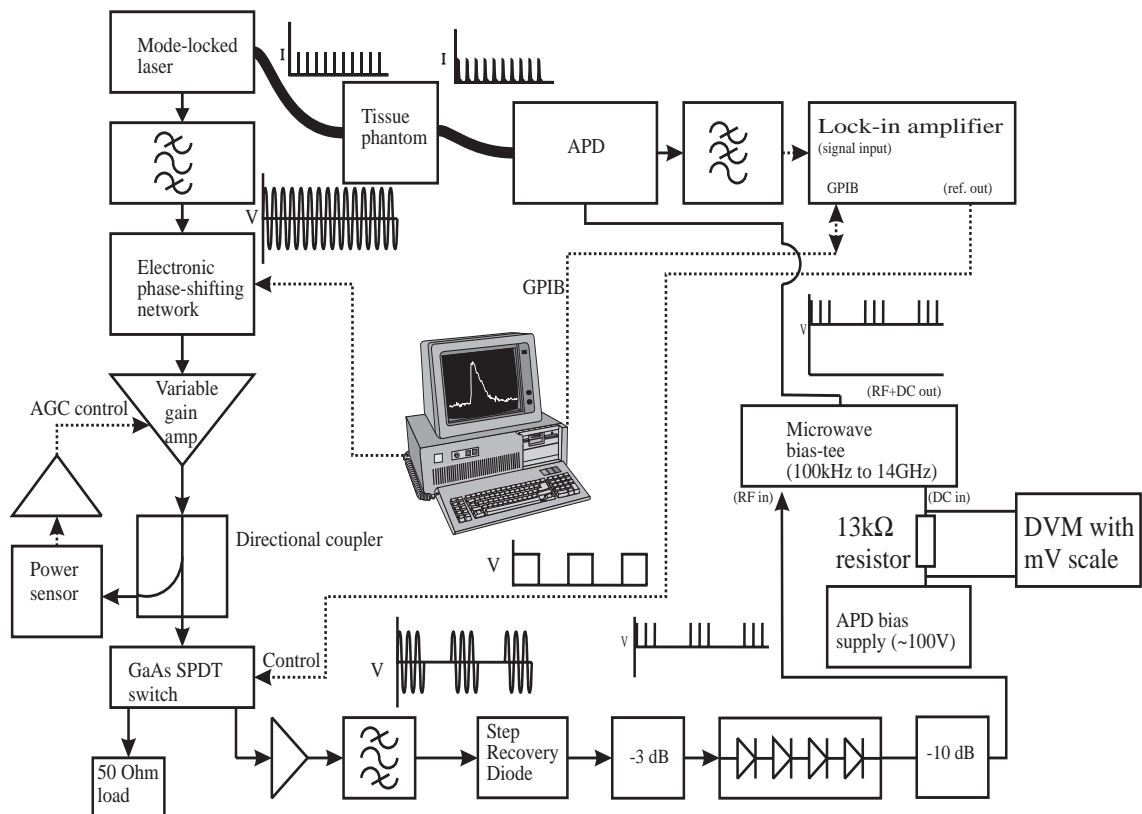
Figure 6.35 Block diagram of the cross-correlator in its final configuration.

The addition of 4 high-speed Schottky diodes in series with the output of the SRD considerably improved the output waveform from the SRD, as seen on an oscilloscope (see figures 6.33 and 6.34). However, it did little to improve the cross-correlator performance. In

particular, the impulse response of the system was still unacceptably wide. This was later found to improve considerably if the output from the SRD was first attenuated with a 3 dB attenuator placed before the diodes and finally a 10 dB attenuator was placed on the output of the diodes, before the bias-tee, as shown in figure 6.35. The improvement can be attributed to the improvement in impedance matching caused by the attenuators. Various combinations of attenuators have been tried, but those shown in figure 6.35 were found to be optimal. This has significantly reduced the amount of base-line noise evident on figure 6.11. The temporal resolution has deteriorated slightly to 380 ps FWHM, but is much cleaner overall, as shown later in figure 6.44.

A 13 kΩ resistor is shown between the APD power supply and the bias-tee, with a digital volt meter connected across the 13 kΩ resistor to enable the APD current to be calculated from the voltage drop. A Tektronix (Tektronix Inc, Measurement Division, PO Box 1520, Pittsfield, MA 01202-9952, USA) DMM916 4½ digit digital volt meter (DVM) with a 400 mV DC full scale range was used. The average APD current, as measured on the DVM gives a good indication of the likely signal to noise ratio of the measurements. A photocurrent of 79.6 nA gave a reasonably good signal to noise ratio (see figure 6.46) while the photocurrent should not exceed a few μA (see figure 6.45) to avoid non-linearity.

6.10   Final system design.

A block diagram of the final system design was shown earlier in figure 6.35. Figure 6.36 shows the front view of the main enclosure of the cross-correlator containing most of the electronics developed as part of this project. The steel enclosure is a standard 19" rack mount unit, 430 mm wide and 250 mm deep and 3U (125 mm) high. The fibre to the left is the optical

Figure 6.36 Front view of the main electronic enclosure for the cross-correlator. The detection fibre can be seen to the left. The BNC socket shown connects to the lock-in amplifier input.
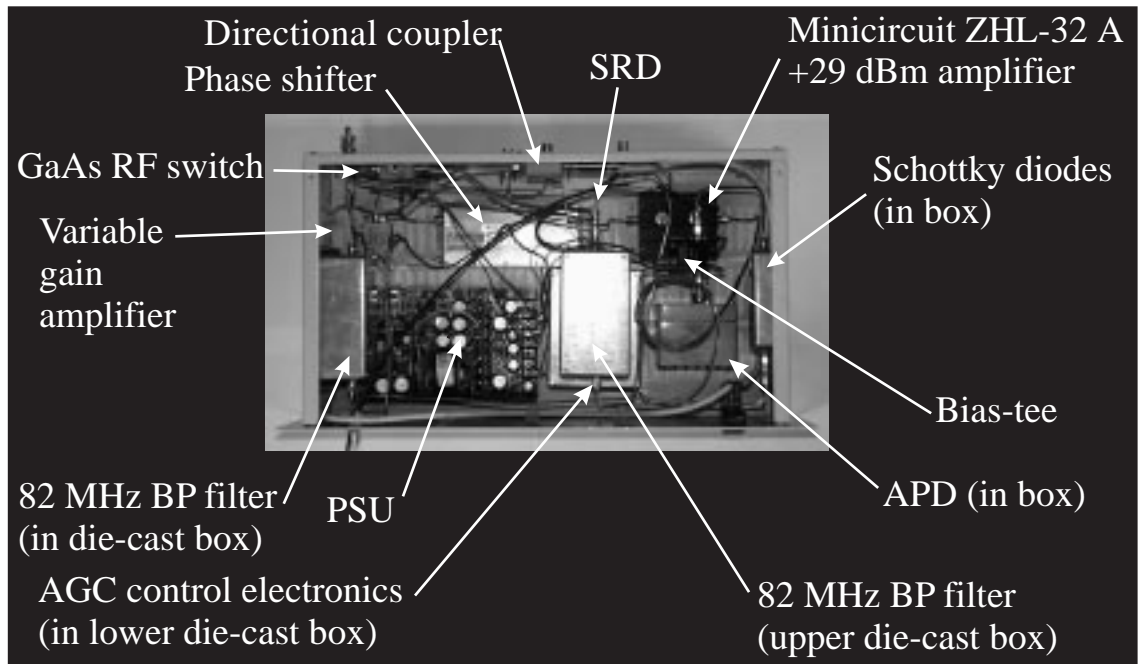
Figure 6.37 Top view of the main enclosure containing the components of the cross-correlator.

input which is connected to the APD via a FC-PC connector internally. The fibre enters the front panel through a rubber grommet. A chassis mount FC-PC to FC-PC connector was not used to reduce losses. The BNC socket is connected to the output of the filter following the APD and connects to the *A* input of the lock-in amplifier. Figure 6.37 is a top view of the enclosure containing most of the cross-correlator electronics. The 3 mm thick front panel is 3U (132 mm)
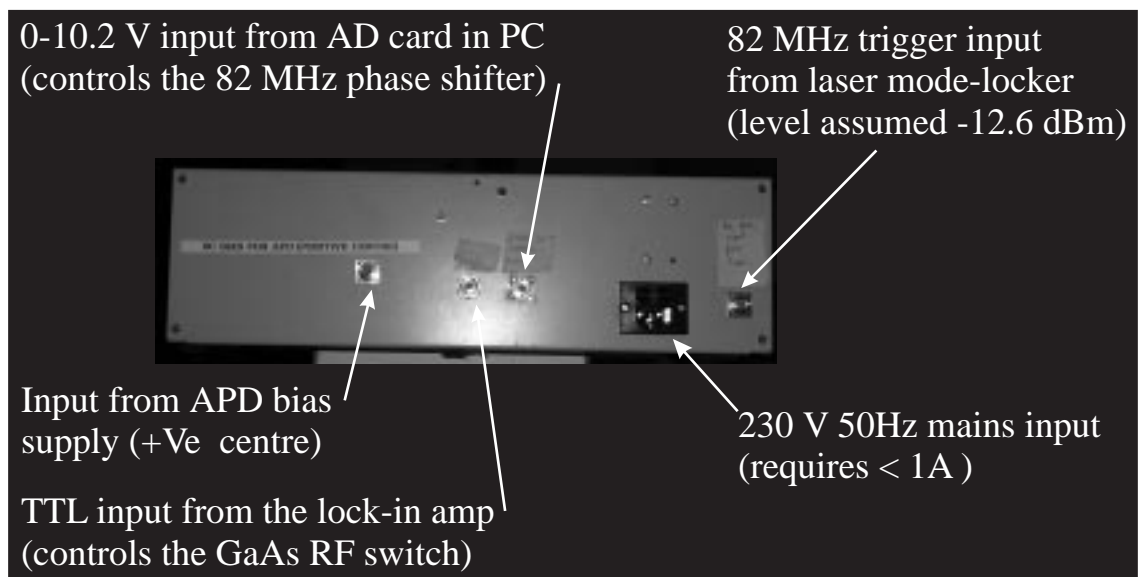
Figure 6.38 Rear view of the main 3U high electronics enclosure for the cross-correlator.

high.

On the rear panel of the main enclosure (see figure 6.38) there are only inputs (no outputs). The inputs were for mains electricity (230 V 50 Hz on a standard IEC socket), an 82 MHz synchronisation signal that came from the laser mode-locker (BNC socket), the APD bias supply (MHV/H4 high voltage socket), a TTL input from the lock-in amplifier (BNC socket) and a 0-10.2 V control signal from the A/D card that alters the phase shifter (another BNC socket).

Figure 6.39 Photograph showing the APD power supply. The holes in the front panel are for two digital voltage meters, which were never fitted.

Figure 6.39 is a photograph of the APD power supply. The 3 mm thick aluminium front panel is 2U (88 mm) high x 483 mm wide. This is a 2U high (19" rack cabinet) and is 350 mm deep. An internal view is not shown, but it consists of one commercial 24 V power supply and a commercial 500 V power supply that was energised from the 24 V supply. The variable resistor on the right-hand side of the front panel sets the APD bias voltage.

The laser table in our laboratory is shown in figure 6.40. The Millennia pump laser which can produce 5 W of green light at 532 nm is shown to the right and the main Tsunami infra-red laser to the right. The bottle of nitrogen, visible towards the upper left hand side of the photograph is used to purge the Tsunami cavity of air, since the absorption of laser light in the near infra-red by water is not insignificant (see figure 7.3) and so would result in less stable laser performance.

The APD used in the final design was the *green* one (see table 6.1 on page 136). It was found to run optimally at 125.8 V, which gave it a gain of approximately 128 (calculated using the APD characteristics in table 6.1 and equation 3.31). The 200 mV modulation voltage increased the voltage to approximately 126.0 V, which gave a peak estimated gain of 137, or an increase of approximately 8.7%. It is interesting to note that Berndt's system ran at a gain increase of up to 800%, even though the APD used as a pulse generator had a much smaller output than the SRD used here. However, using the full SRD output was found to be impossible in this system without the previously described degradation of the system response. The attenuators are believed to be necessary, since they prevent multiple reflections propagating

Figure 6.40 Photograph showing the laser table in our laboratory. The right hand laser is the Millennia pump laser producing 5 W at 532 nm. The laser to the left is the infra-red Tsunami laser, tunable from 720 to 840 nm.

between the RF input port of the bias-tee and the diodes at the output of the SRD. Neither of these are matched properly to 50 Ω, so some resistive padding is necessary. While increasing the DC bias voltage above 125.8 V is possible, which would give increased gain-modulation and increased output, the system's temporal response would deteriorate, as discussed in section 6.15.

6.11 Noise analysis.

Originally the lock-in amplifier was configured to operate as a voltage measuring device (input impedance = 10 MΩ in parallel with 25 pF) and it measured the voltage developed across a 50 Ω resistor placed after the low-pass filter. This worked well at high light intensities but less well at low light intensities, as the signal was buried in the noise. An example of this is shown in figure 6.41 where the average photocurrent was 1.5 μA, the lock-in filter was set for a time constant of 100 ms, the filter slope to 6 dB/octave and the full scale deflection (FSD) to 10 μV.

The lock-in amplifier was then tried in current input mode, where its input impedance
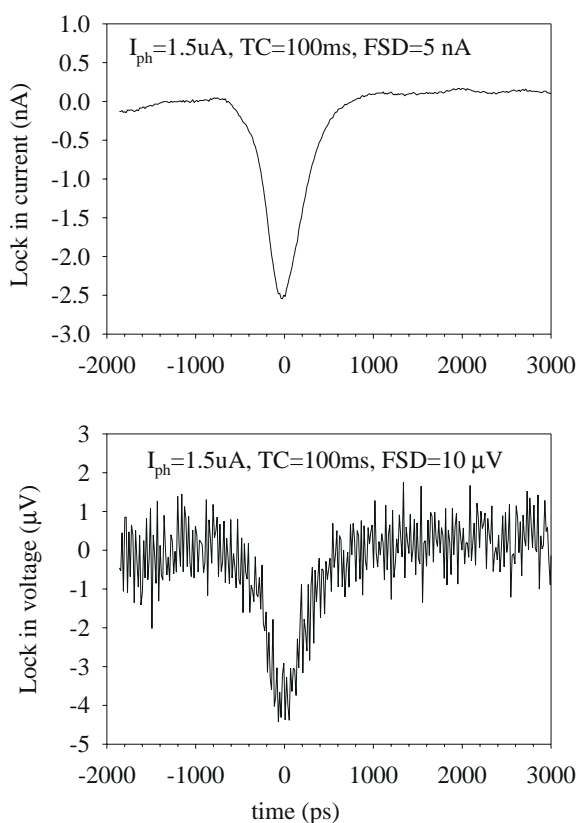
Figure 6.41 Comparison of the lock-in output when operated in current mode (upper graph) or voltage mode (lower graph). All other conditions (light intensity etc) were identical.

is 1 kΩ. The input to the lock-in is then passed through an internal current-to-voltage converter, with a transimpedance of 10^6 Ω. The 1 kΩ input impedance of the lock-in is large compared to the 50 Ω resistor it is connected across, so it did not significantly effect the termination of the low-pass filter. This resulted in much lower noise for the same light intensity, as the upper graph in figure 6.41 shows.

Note that the waveforms in figure 6.41 are shown inverted compared to most others. The original data files do show a decreased output like that shown in figure 6.41, and not as shown in all other diagrams. This is simply due to the lock-in amplifier subtracting the output current when the SRD is on from when it is off - the reverse of what is most logical. This could have been corrected with an TTL inverter between the reference output of the lock-in amplifier and the GaAs SPDT switch. However, it was easier to invert the sign of the collected data, which is what has been done with all other graphs. Although it is not shown, a measurement of noise across a simple 50 Ω resistor using the lock-in amplifier in voltage input mode gives much the same noise levels as seen in the lower graph of figure 6.41, so the noise is not a function of the cross correlator. There are several sources of noise than contribute to the total noise, which will now be discussed.

6.11.1 Johnson noise of the 50 Ω source resistance.

The 50 Ω source resistance will add a Johnson[103,104,131,132] noise voltage of:

$$V_{Johnson\ noise\ (RMS)} = \sqrt{4kTR\Delta f} \qquad \textbf{6.16}$$

where k=Boltzmann's constant (1.38×10^{-23} J/K), T is the temperature in Kelvin, R is the resistance in Ohms and Δf is the measurement bandwidth in Hz. The input amplifiers in the

lock-in have a bandwidth of approximately 300 kHz, so with a 50 Ω source resistance and 300 K temperature, there is an input voltage noise of 498 nV (RMS), or about 5 times this (2.49 µV) peak-to-peak. However, the amount of Johnson noise measured by the lock-in amplifier would be less than this, due to the filtering action of the filters. Referring to table 6.1 it can be seen that using the 6 dB/oct filters and a 100 ms time constant gives an equivalent noise bandwidth of 1/(4T)=2.5 Hz. Hence the typical Johnson noise measured due to the 50 Ω resistor at the input would be 1.45 nV RMS.

6.11.2   Shot noise.

The photocurrent which was 1.5 µA will create shot noise[131,132] given by

$$I_{Shot\ noise\ (RMS)} = \sqrt{2qI\Delta f} \qquad\qquad \textbf{6.17}$$

where q is the charge on an electron (1.6×10^{-19} C), Δf is the bandwidth and I is the RMS photocurrent. This gives a shot noise of 1.09 pA (RMS). Since this current flows through a 50 Ω resistor, it will develop a voltage of 54.5 pV RMS.

6.11.3   1/f noise.

Since the measurements were made at a frequency of 1000 Hz, 1/f noise is insignificant[131,132].

6.11.4   Input voltage noise of the lock-in amplifier.

The specification of the lock-in amplifier states the input voltage noise is a maximum of 8 nV/√Hz, although 5 nV/√Hz is said to be typical. The lock-in amplifier, when used at a sensitivity of 10 µV, has a gain of 1000, so the noise at the input of the low-pass filters following the phase sensitive detector will be 5 µV/√Hz. Referring to table 6.1 it can be seen that using the 6 dB/oct filters and a 100 ms time constant gives an equivalent noise bandwidth of 1/(4T)=2.5 Hz. Hence the typical noise at the output of the low-pass filter will be 5 µV/√Hz or 3.16 µV RMS.

6.11.5   Sum of the noise sources.

The four noise sources identified (shot noise, Johnson noise, 1/f noise and the input noise of the lock-in amplifier) are incoherent, so must be added in quadrature:

$$V_{Total\ noise\ (RMS)} = \sqrt{V^2_{Johnson\ noise} + V^2_{Shot\ noise} + V^2_{1/f\ noise} + V^2_{Lock-in\ noise}} \qquad \textbf{6.18}$$

which gives

$$V_{Total\ noise\ (RMS)} = \sqrt{(1.45 \times 10^{-9})^2 + (54.5 \times 10^{-12})^2 + 0^2 + (3.16 \times 10^{-6})^2} = 3.16\ \mu V\ (RMS) \qquad \textbf{6.19}$$

By far the most significant of these noise sources is the lock-in amplifiers internal noise, which explains why the lock-in did not work very well when operated in voltage input mode. When changed to current input mode, the input current to voltage converter increases the input voltage considerably, so the input voltage noise of the lock-in amplifier becomes less significant.

The noise shown in the lower graph of figure 6.41 is approximately 3 μV peak-to-peak, or about 3/5=0.6 μV RMS. However, the data acquisition software (see section 6.16) averages 10 measurements before outputting the mean of the 10 measurements to a file, so the actual noise measured was 0.6x√10=1.9 μV RMS. This is within 60% of the theoretically predicted value. The difference between theory and measurement is likely to be due the lock-in amplifier being slightly less noisy than assumed.

6.12   Time-base calibration.

The data acquisition software written to control the cross-correlator, could scan any section of the TPSF from t=0 to t=12200 ps in steps of 15.3 ps. The data acquisition programme initially computed the value of the phase ϕ needed on the phase shifter using:

$$\phi = t * 360 / 12200 \qquad \textbf{6.20}$$

where t is in ps. This arises since the time between laser pulses is 12200 ps, and for a complete span of 12200 ps, the phase shifter must be moved a full 360°. The data acquisition programme then computed the voltage necessary on the phase shifter to cause this phase shift, using equation 6.8. This voltage was then output by the DA card. To advance the section of the TPSF collected, the phase was increased, and the DA output voltage increased. It was soon realised that the time-base went backwards! This was because increasing the phase delay through the system causes a delay, not an advance, so a revised method of calculating the phase shift required was used, as given in equation 6.21

$$\phi = 360 - (t * 360 / 12200)$$ **6.21**

Equation 6.8 was still used to calculate the output voltage to send to the phase shifter from the DAC.

Data is usually required at known times. When the source fibre and detector fibre are placed in line, touching each other the light leaving the end of the source fibre arrives at the detector fibre with a time delay defined to be zero. This time is used as a reference in all the measurements. The system is calibrated for time by first performing an impulse response measurement with the source and detector fibres touching, then moving them apart in the medium of interest (tissue, or some phantom material), and subtracting from the latter measurements the peak time of the impulse response. This is shown in the following example. A typical impulse response measurement is shown in the centre graph of figure 6.42, where the source and detector fibres are touching. In this instance the impulse response peaks at 3626 ps, not 0 ps. Only a small part of the

Figure 6.42 Graphs showing how the time-base is calibrated, by measuring the temporal position of the impulse response peak and subtracting this from the measurements.

system impulse response was collected as only the peak position was needed. Hence all measurements made need to have 3636 ps subtracted in order to calibrate the system. The top graph in figure 6.42 shows some data collected experimentally in a scattering medium, with a μ_a of 0.0192 mm^{-1}, μ_s'=3 mm^{-1} at a source-detector spacing of 10 mm. The peak occurs at a time of 3749 ps, but after subtracting the 3636 ps, the corrected data is shown in the lower graph of figure 6.42, where the peak is then at 3749-3626=337 ps. If the system has been left for some time, or anything changed, a new calibration should be done. Since a lot of light is available for the calibration, it takes very little time. When the system was first assembled, we had no idea where the peak would be, so it was necessary to examine the whole 12200 ps. Subsequently,

when only small changes to the positions of the optical components or bias voltage were made, the zero point will be known fairly accurately (say within 100 ps), so it was only necessary to scan over a small range to find the zero point reference. This is why only a small section of the system impulses response is shown in the centre graph of figure 6.42. Sometimes it was inconvenient to do a reference measurement at a source-detector spacing of zero, due to the difficulty in aligning the fibres sufficiently accurately to couple light into the detector fibre. Often it was more convenient to set the source-detector spacing at 20 mm in air, and allow for this. Since light travels at 3.333 ps/mm in air, at a source-detector spacing of 20 mm, the reference file would need to have a correction of 20*3.333=66.7 ps.

It is desirable that the zero point is at no more than say 9000 ps, otherwise it will be impossible to measure TPSFs exceeding 12200-9000 = 3200 ps. If by chance this is the case, it can be shifted by either changing the length of one of the RF cables driving the SRD, or changing the optical paths externally - whichever is the more convenient.

6.13　　　Time-base linearity.

Section 6.6 discussed how the non-linearity of the R&K phase shifter caused non-linearity in the time base, and how this was corrected in software, by fitting a 4th order polynomial to the data in figure 6.14. Whilst a quick check was made at the time to check that this appeared to work, a thorough check was not made at the time. Hence before the system was used more fully, the time base linearity was checked carefully. The output of the laser (which emits a beam in air, and is not fibre coupled), was coupled to a fibre optic coupler fitted with a fibre terminated in an FC-PC optical connector. The fibre coupler was mounted on top of a 300 mm slide, Aerotech model PR-12/M 300 (Aerotech, 101 Zeta Drive, Pittsburg, PA 15238, USA). The impulse response of the system was then measured. The impulse peaks at some time, when the SRD pulses are gain modulating the APD at the same time as the laser pulses arrive. The fibre coupler was then moved 50, 100, 150, 200 and finally 250 mm further away from the laser on the slide. The impulse response was measured each time. Theoretically, one would expect the
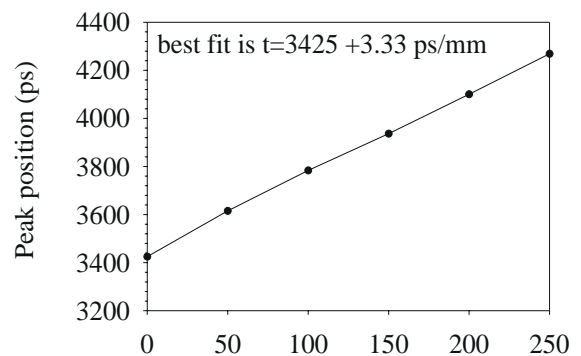
Figure 6.43 Diagram showing the shift in time of the peak of the impulse response as the detector is moved further from the laser.

peaks of the subsequent impulse responses to be shifted by 3.333 ps for every mm moved, or 166.66 ps for each 50 mm step. Figure 6.43 is a diagram of the shifts. One can see the impulse response advances in time in an almost linear manner as the arrival time of the laser pulses is delayed by moving the coupler further from the laser. The line of best fit has a slope of 3.33 ps/mm - virtually zero error.

6.14   **Amplitude Linearity.**
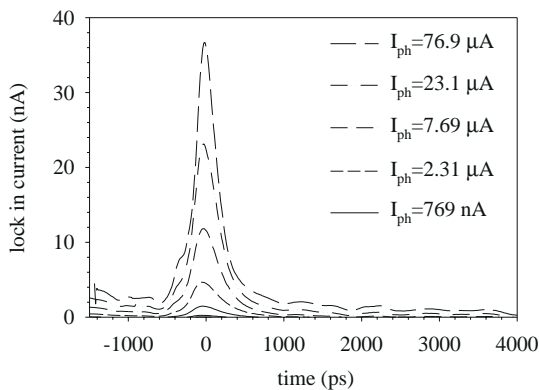
Figure 6.44 Graphs showing the impulse response for a range of different detected light levels.

Figure 6.45 Graphs showing the variation of lock-in amplifier current at the peak of the impulse response as a function of the photocurrent I_{ph}. Shown on both logarithmic and linear scales.

The amplitude linearity of the cross-correlator was checked, by measuring a system impulse response at various photocurrents, at a fixed DC voltage of 125.8 V from the APD bias supply. As the photocurrent was increased, so the voltage across the 13 kΩ series resistor increased ($\Delta V=IR$), and so the actual voltage at the APD fell. Since the gain is a strong function of voltage, the gain of the APD decreases at high light levels. Figure 6.44 shows the system impulse response collected at several different photocurrents (76.9 µA, 23.1 µA, 7.69 µA, 2.31 µA, 769 nA, 231 nA and 76.9 nA), although the lower photocurrents are not easily visible (these are shown more clearly in figure 6.46).

Figure 6.45 shows the peak lock-in current, at the peak of the impulse response, as a function of photocurrent on both linear and log scales. It can be seen that the photocurrent should be kept below a few µA, as otherwise gain compression was occurring, due to the voltage drop across the load resistor causing a reduction in the bias voltage. This was not a problem in

Figure 6.46 Impulse response measured at three small photocurrents (769 nA, 231 nA and 76.9 nA).

practice, as the signal to noise ratio of the system was acceptable at currents below this. Figure 6.46 shows the impulse response measured at three photocurrents of 76.9, 231 and 769 nA, where the S/N ratio is seen to be acceptable even at the lowest of these.

Figure 6.46 also shows that the baseline is not exactly at zero, but in the example given is at a current of approximately -0.15 nA. The base line is not completely flat either, but is somewhat more positive at early times than at later times. It is seen to be dependent on the average photocurrent, becoming more negative for small photocurrents. Several possible causes of this DC offset have been considered, such as the finite isolation of the GaAs RF switch, but no satisfactory explanation for this can be found.

6.15   Effect of APD bias voltage.

As discussed in section 3.2.2, the bias voltage of an APD sets the gain - the higher the voltage, the higher the gain. In this instrument, the bias voltage also sets the increase in gain achievable by the 200 mV of SRD pulse that is added to the bias voltage by the bias-tee. A pulse voltage of say 200 mV will produce a far greater increase in gain if the normal DC bias voltage is close to the breakdown voltage, than it will if the voltage is low. Graphs illustrating this are shown in section 3.2.8.

What was surprising was the way that the bias voltage effected the temporal offsets in the system. To test this, the laser was coupled via some neutral density filters to the APD, and the DC photocurrent adjusted to 1 μA by rotating the variable neutral density filter. The bias voltage was increased in stages, starting at 64.48 V, and the impulse response of the system was measured. This had a peak at an arbitrary time, which was noted. Each time the bias voltage (and hence gain) was increased, the light intensity was dropped to keep a constant 1 μA

photocurrent, and the impulse response again measured. As the voltage was increased, so the peak of the impulse response shifted positive in time - equivalent to a delay in the light reaching the APD. This data is shown in figure 6.47. Clearly as the bias voltage gets closer to avalanche, the apparent delay through the APD is increased. The reason for this is likely to be that the time for the avalanche to develop increases for increasing gain, as predicted by equation 3.49. However, since the physical construction of the APD is

Figure 6.47 Variation of the peak time of the system impulse response, as a function of APD bias voltage.

not accurately known, it is not possible to verify this for certain. Voltages below 64.48 were not used, as the system barely functions at these low voltages - there is not significant modulation pulse voltage to cause a measurable increase in APD gain.

If the APD bias voltage was too high, the impulse response of the system increased in width. At a voltage of 125.8 V, the system response was reasonably narrow at 380 ps. There was little change in the width of the impulse response if the voltage was lowered - only a loss of sensitivity. However, if the bias voltage was increased to 130.6 V, the impulse response widened to 800 ps FWHM. The most likely explanation for this is that the avalanche takes longer to develop, as equation 3.49 shows. Although another possible explanation for this is that at high bias voltages, a small change in voltage will cause a large change in the zero reference position, as was shown in figure 6.47. Hence even small amounts of noise on the bias voltage, or small changes in the APD characteristics with temperature will cause a large jitter on the reference position and so could cause a broadening in the impulse response.

6.16   Software.

The data acquisition software was written in Standard C, using the Microsoft Quick C compiler, version 2.0. The programme was called *tpsf*, and was written to run under DOS (Disk Operating System), version 5.0. The programme had approximately 3700 lines of C code and had four main functions:

1) Set the lock-in amplifier to the correct sensitivity, time constant, filter characteristics, input mode (current or voltage), etc. This was done via a General Purpose Interface Bus (GPIB) interface.

2) Controlled the phase delay caused by the phase shifter, by outputting the correct analogue control voltage from the DAC board fitted inside the computer.

3) Acquired the data from the lock-in amplifier, using the GPIB interface. 10 measurements are made at each point and the average and standard deviation calculated.

4) Saved the measurements of either voltage or current from the lock-in amplifier to the computers hard disk. The standard deviation was also saved.

5) Displayed the TPSF. Currently the software only does this once the full TPSF has been collected. It would have been preferable to display this point by point, since if there was a problem, it would have been easier to determine this without waiting for the measurement to be completed.

6.16.1   Software usage.

The programme *tpsf* was not interactive, but set the instrument from parameters passed as command line arguments to the programme. The programme took the following 10 command line arguments. The first 4 do not require any values, the last 6 do. The command line options were processed using a version of the *getops* command found on unix systems. The C source code for *getopts* was located, and ported to the PC. The use of *getopts* makes handling command line arguments quite easy.

6.16.1.1   Command-line arguments.

i - Sets the input mode to current (default was voltage).

X - Acquires the component of voltage or current in phase with the reference.

Y - Acquires the component of voltage or current at 90° to the reference.

R - Acquires the magnitude component $(X^2+Y^2)^{0.5}$. This was the default behaviour of the software, although it was not frequently used. Ideally the software needs changing to make option X the default.

The programme also takes several arguments that required values be used with them.

F - Sets the full scale deflection in either fA or nV, depending on whether the instrument was in voltage or current mode. The default was 1 nA (current mode) or 10 μV (voltage mode). The range was 2nV to 1 V (voltage mode) or 2 fA to 1 μA (current mode).

r - Sets the reference frequency (in Hz) for the lock-in amplifier. The default was 1 kHz, the range was 0.001 Hz to 102 kHz.

s - Sets the step size (in ps) for data to be collected. The default was to collect data every 15.3 ps. The range was 1 to 12200 ps.

T - Sets the time constant of the low-pass filter. The default was 100 ms, the range was 1 ms to 30,000 seconds. Note, the SR-830 lock-in amplifier was capable of having time constants from 30 μS to 30 ks, but the *tpsf* programme was not able to set the time constant below 1 ms. Time constants below 1 ms were never needed - 100 ms was the shortest time constant used.

S - Sets the data acquisition rate in mHz. The default was 8000 (8 Hz), with the range being 62.5 mHz to 512 Hz.

O - Sets the order of the low-pass filter, which affected the rate of fall off, as shown in figure 6.2. The default was 6 dB/octave, but it could be set at 6, 12, 18 or 24 dB/octave.

6.16.1.2   Examples of data acquisition.

A minimal command to the software would be:

tpsf 3s2_5m 2000 4000

which would collect data using the defaults (voltage mode, R signal, 100 ms filter time constant, 10 μV FSD, 8 Hz sampling rate, 6 dB/octave filter, 15.3 ps step size) into a file called 3s2_5m.dat, between 2000 and 4000 ps.

A more typical collection routine would be

tpsf -i -X -F5000000 3s2_10m 2500 6000

which would collect data in current mode (i), in phase (X) with the reference, at a full scale deflection of 5000000 fA or 5 nA (F=5000000), 100 ms filter time constant, 8 Hz sampling rate, 6 dB/octave filter, 15.3 ps step size, from 2500 ps to 6000 ps.

If the light intensity was weak, due to large source-detector spacings or high μ_a or μ_s', then the sensitivity of the lock-in would be increased, the time period of the filter would be increased from the default 100 ms, and the data collected more slowly to give the filters chance to settle. The following would achieve this.

tpsf -i -X -F500000 -S500 -T1000 3s2_20m 2500 6000

which would use current mode (i), in phase (X) with the reference, a FSD of 500 pA (F=500000), collect data at 500 mHz (S=500), using a time period of 1 second (T=1000).

6.16.1.3 An example data file.

The software produced a data file, with the extension .dat, that has the following format, if collected in current mode.

t(ps)   DAC_voltage(V) Current(nA) SD_of_current(nA) text_parameter value unit

The top and bottom of a typical file would look like:

```
2723.5 8.368090 0.061437 0.007367 SENS 5 nA
2738.8 8.357950 0.060527 0.010596 FREQ 1000.000000 Hz
2754.1 8.347810 0.063256 0.013392 RATE 8.000000 Hz
2769.4 8.337669 0.074033 0.002319 TC 0.100000 secs
2784.7 8.327528 0.057935 0.014870 ORDER 6 dB/oct
2800.0 8.317388 0.049204 0.013347 Version 1.3 software
2815.3 8.307246 0.049431 0.004957 RESERVE 1 level
...............
7635.0 4.870941 -0.128740 0.010436 RESERVE 1 level
7650.3 4.858833 -0.120645 0.026307 RESERVE 1 level
```

indicating data is collected at 5 nA FSD, a reference frequency of 1000 Hz, a sample rate of 8 Hz, a time constant of 100 ms and 6 dB/octave filtering and version 1.3 of the software. The data is collected in 15.3 ps steps. The voltage sent from the DAC to the analogue phase shifter was reduced from 8.368 V to 4.859 V during the data acquisition. (In practice, the 12 bit DAC, with a range 0 to 10.25 V, would have had a step size of 10.25/4095=2.5 mV, so the voltages in the file would not be the exact voltages output by the D/A board).

6.17 **Data collection duration.**

The time to collect a TPSF depended on several parameters.

1) The time steps used. The default was 15.3 ps, but it could be changed from 1 ps to 12200 ps. The default of 15.3 ps was generally suitable.

2) The duration over which the TPSF was collected. It was beneficial to collect data both before the TPSF starts to rise and after it had fallen to a level below the noise of the system. This is because the baseline was not flat and two flat regions were necessary in order to correct for this. The correction procedure is explained more fully in section 6.12.

3) The number of data points that were collected at each time step and subsequently averaged. Ten points were always averaged.

4) The sample rate of the lock-in amplifier, which needed to be set appropriately for the time-constant of the lock-in amplifier filters. The time constant of the filters was set depending on the light level reaching the APD.

At a source to detector spacing of 10 mm or less, using typical tissue optical properties, a 100 ms setting for the time constant of the filters was adequate. The sample rate was 8 Hz. While the settling time (0 to 99%) of the filters would be just under 500 ms (see figure 6.2), the 10 data points were collected within 1.25 seconds (the first after only 125 ms). This is justifiable, as the data doesn't change dramatically over a 15.3 ps time step, due to the system response (380 ps FWHM). Hence to scan a 15.3 ps section of TPSF would take 1.25 seconds. A typical scan of 2000 ps would take (2000*1.25)/15.3=163 s or 2 minutes and 43 seconds, which is 81.5 ms/ps. The FWHM of the system has been compared at shorter lock-in amplifier filter time constants, while still sampling at the default 8 Hz. No improvement was seen in the temporal resolution of the system, indicating the time allocated for the filters to settle was adequate.

If the light level droped too low, the 100 ms filter time constant would be too small. In this case the filter time-constant needed to be increased and the sample rate decreased. The time to acquire data then increased. Data collected at a μ_s' of 3.0 mm^{-1}, μ_a of 0.0192 and a source to detector spacing of 15 mm, using an Intralipid^{TM} and dye phantom (see chapter 7) used a lock-in amplifier time constant of 1 s and a sample rate of 500 mHz. Since the sample rate was 0.5 Hz, the 10 measurements at every time step took 20 seconds. The data was collected over a time period of 3840.4 ps, or (3840.4/15.3)+1=252 steps. Hence the total time for the data collection was 252*20=5040 s, or 1 hour 24 minutes. This was the most extreme example. A few minutes was more typical to measure a full TPSF.

Chapter 7.

Design and Construction of a Tissue-like Optical Phantom.

The process followed in designing and constructing the cross-correlator has been described in chapter 6. This chapter concentrates on the design of the optical phantom, which was necessary to test techniques of deriving μ_s' changes from TPSFs measured with the cross-correlator. Chapter 8 describes a computer model that can simulate light propagation in the phantom. The experimental results obtained with this phantom are described in chapter 9, along with the simulations from the computer model.

7.1   Optical phantom construction.

During early development of the instrument, a solid phantom with a low absorption coefficient and high scattering coefficient (μ_a=0.001 mm^{-1} and μ_s'=2 mm^{-1}) was used, fitted with two FC-PC optical connectors, as described on page 144. By using a low absorption coefficient long TPSFs could be generated, which were easier to measure when the instrumentation was under development and its temporal response still poor.

Although solid phantoms, made by mixing scattering objects within a stable epoxy material provide the most stable optical phantoms, they are not particularly suitable to act as a phantom where the source and detector are below the surface. Whilst it would have been easy to cast the optical fibres into the phantom, it would not have been possible to subsequently move them. Since the TPSF obtained will obviously be dependent on the source-detector spacing, it is necessary to make measurements with various source-detector spacings. This prevents the use of solid phantoms for the studying of the μ_s' monitoring technique.

A liquid phantom was therefore developed to allow the optical fibres to be moved readily. Two FC-PC to FC-PC fibre-optic patch cords, with 100 μm core multi-mode fibre were purchased. These consisted of the fibre, surrounded by Kevlar[*1] for strain relief, with a PVC outer jacket. They were cut in half, and the Kevlar™ and PVC stripped back for approximately

---

[*1] Kevlar is a trademark of Du Pont
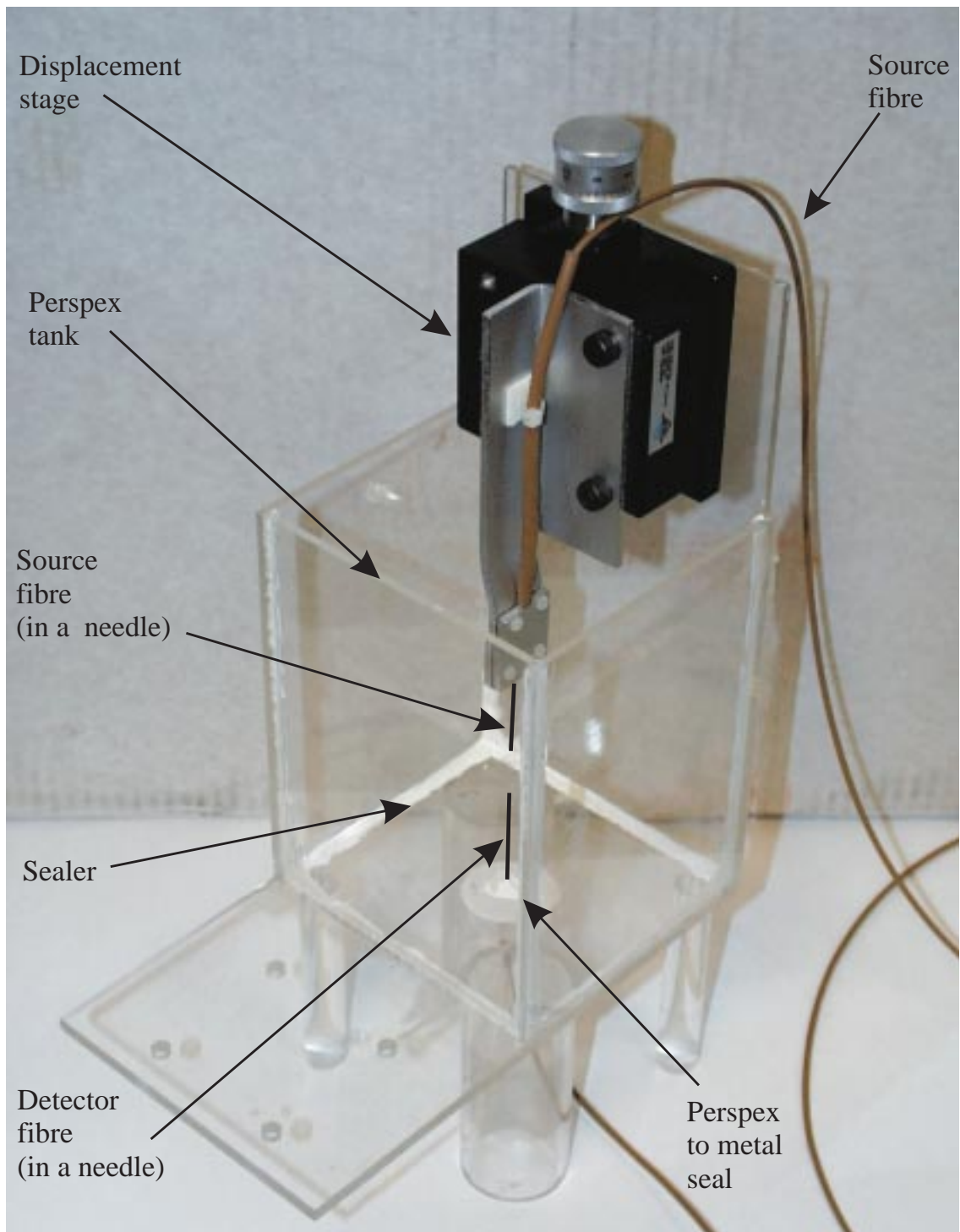
Figure 7.1 Diagram showing the tank built to contain a liquid phantom. The very thin hypodermic needles have been enhanced to aid visibility.

75 mm. The inner fibre was then threaded though a number 25 SWG hypodermic needle, 50 mm long. (The number 25 SWG needle, which is an imperial size, has an outside diameter of approximately 0.50 mm, and an inner diameter of 0.25 to 0.31 mm). Obviously, 25 mm of fibre

extended beyond the needle. Araldite (Bostik Ltd, Ulverscroft Road, Leicester, LE4 6BW) was then applied to the fibre and the fibre gently pulled back and forth, to get as much adhesive as possible into the needle. The Kevlar strain relief was glued with Araldite to the outside of the needle, and the whole arrangement left to set hard. Heatshrink tubing was shrunk over the transition between the fibre and the needle. Finally, the fibre ends were polished. The standard (rather than the quick-set) Araldite was used as it is more stable under long-term water emersion.

The fibres, inside the hypodermic needles, were then inserted into a tank, shown in figure 7.1. The tank consisted of a Perspex™ container 115 mm wide x 116 mm deep x 149 mm high, which was fabricated from 3 and 5 mm thick Perspex™ sheets (different thicknesses were used as they were offcuts). The back panel is somewhat taller than the other three sides, at 233 mm high. The tank rested on 4 legs, each 83 mm high (again different diameters were used, just to use offcuts of material). The Perspex™ sheet was cut as accurately as possible on a band saw, then glued together using Tensol 12 one-part cement (Evode Speciality Adhesives Ltd, Anglo House, Scudamore Road, Leicester, LE3 1UQ, England), as shown in figure 7.1. A couple of small leaks remained after the Perspex™ was glued together, which required sealing with a silicon based sealer. (This was a normal kitchen/bathroom sealer, made by EVO-STIK). Machining the edges of the Perspex™ accurately would have prevented this being necessary, but it would have been a time-consuming process. In contrast the small leaks were relatively fast to repair.

A vernier displacement stage (Areotec Model ATS300M) was mounted on the deeper back panel. The source fibre inside the hypodermic needle was mounted on the vernier displacement stage with a bracket. A small hole, 0.6 mm in diameter, was drilled in the bottom of the tank to allow the detector fibre (connected to the cross-correlator), to be inserted into the tank. Some strain relief was added under the tank and the detector fibre was sealed with the same sealer used in the tank.

An arrangement of four holes on a horizontal plate at the bottom of the tank can also been seen in figure 7.1. This was originally intended to allow the source fibre to be moved into a horizontal position and the detector fibre to be attached to the displacement stage, allowing investigation of the effects of having the source and detector fibres positioned at 90° to one another. However, this was not done to save time.

On the back of the phantom, where the vernier displacement stage is attached, a large T-section piece of aluminium was attached to the Perspex™ with silicon sealer to provide extra rigidity. Without this, the position of the upper fibre moved under the strain in the Perspex™

whenever the phantom was touched. The T-section is not shown in the photograph. The capacity of the tank was approximately 1700 ml, although it was originally decided to fill this to only 1300 ml. Subsequently, to allow the source and detector to be spaced further apart and still remain well below the surface, the tank was filled to 1500 ml.

7.1.1　　　　　　Scattering component of the optical phantom.

While several researchers[133] in the field of tissue optics have used milk as a scattering component in tissue phantom, this is not very suitable, as the optical properties of milk vary with the type of milk and processing method (pasteurised, skimmed, semi-skimmed, full fat, sterilised etc). However, nowadays the popular choice of scattering material in liquid phantoms is Intralipid^{TM} (Pharmicia Laboratories Ltd, Davy Avenue, Milton Keynes, MK5 8PH). Intralipid^{TM} is purified soybean oil, which is widely used as an intravenous nutrient. It is fairly inexpensive (approximately £7 for a 500 ml bottle in 1998) and non-toxic. It is usually obtained as a 10% solution (containing 50 g of purified soybean oil per 500 ml of Intralipid^{TM}), although a 20% solution is available.

Figure 7.2 Graphs of μ_s, g and μ_s' for 10% stock Intralipid diluted 1000 times, to 0.01%.

The optical properties of Intralipid^{TM} have been studied by van Starvansen et al[134] and by Flock et al[135]. The former paper is more useful for infra-red studies, as the wavelength range considered in the latter paper is more restrictive. Van Staveren et al[134, 136] used electron microscopy to study the size and distribution of the particles and used Mie theory to calculate the optical properties. These calculations were compared to the optical properties measured experimentally at 457.9 nm, 514.5 nm, 632.8 nm and 1064 nm. They found that the particles have a range of sizes up to 700 nm (mean size is 97 nm), and a density of 3.97×10^{19} m^{-3}. They found that the scattering coefficient (μ_s) dependence on wavelength (λ) was given by equation 7.1

$$\mu_s(\lambda) = 0.016\lambda^{-2.4} \qquad (\pm 6\%) \qquad\qquad \textbf{7.1}$$

and the anisotropy factor g was given by equation 7.2

$$g(\lambda) = 1.1 - 0.58\lambda \qquad (\pm 5\%) \qquad\qquad \textbf{7.2}$$

for $0.4 < \lambda < 1.1$, where λ is in μm, and μ_s is in mm^{-1}, when the 10% solution is diluted 1000 times with distilled water. Figure 7.2 shows graphs of μ_s, g and $\mu_s{}'$ plotted over the range of 400 nm to 1100 nm, for 1 ml of the 10% Intralipid^{TM} diluted in 999 ml of water, to make 1000 ml of solution. From the bottom graph in figure 7.2, it can be seen that at 800 nm the transport scattering coefficient of the solution is approximately 0.01 mm^{-1}. Tissue has a $\mu_s{}'$ of typically 1 mm^{-1}. To obtain a solution with $\mu_s{}' = 1$ mm^{-1} one needs to mix a solution comprising 1 part of the 10% Intralipid^{TM} to 9 parts distilled water. During these experiments, which had the aim to determine if $\mu_s{}'$ could be determined from the instrumentation, three different values of $\mu_s{}'$ were used - 0.3 mm^{-1}, 1 mm^{-1} and 3 mm^{-1}. These required respectively 45, 150 and 450 ml of Intralipid^{TM}, with the remainder of the 1500 ml being distilled water. The Intralipid^{TM} and water were both weighed to determine the amounts to use (relative density of Intralipid^{TM} is close to 1.0), rather than measured volumetrically, as weighing was more accurate.

Van Starvansen et al[136] state the equations 7.1 and 7.2 are valid only until the concentration of 10% Intralipid^{TM} is above 4% (at 1100 nm) to 17% (at 400 nm), as beyond this, the scattering events are not independent. The Intralipid^{TM} concentration used here is much more than this, so lies outside their stated range of validity. Work by Lovell et al[137] have shown that with red blood cells the concentration can be much higher than this while scattering is still proportional to concentration of red blood cells. For this reason, it was felt acceptable to use concentrations of 30% of 10% Intralipid^{TM}.

The background absorption of diluted Intralipid^{TM} is close to that of the water it is diluted in. Figure 7.3 shows the absorption coefficient of distilled water over the range 650-1050 nm. This data was measured in our own laboratory, and has been published by Matcher et al[138]. It can be seen that at 800 nm, the absorption coefficient of water is approximately 0.0018 mm^{-1} - well below that of any tissue[34]. Extra absorber must therefore be added to the Intralipid^{TM}, to make its optical properties closer to that of tissue.

7.1.2 Absorbing component of the optical phantom.

To add extra absorption to the tissue phantom, an infra-red absorbing dye, S109564,
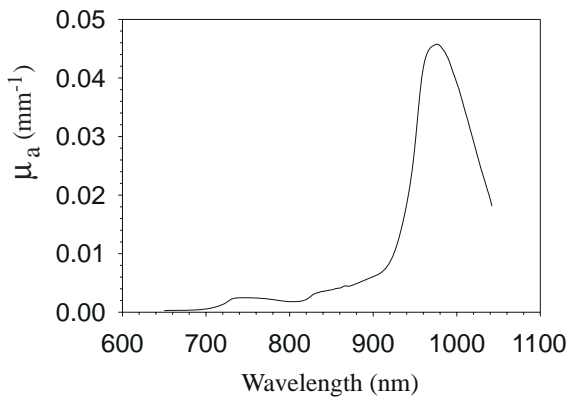
Figure 7.3 Absorption coefficient of water.

made by ICI (Imperial Chemical Industries PLC, Organics Division, Blackley, Manchester, England) was added. This is a pure absorber and so would not add any scattering, unlike India Ink, which has been used by several groups[139] as an absorber (but does introduce some scatter[140]). It was assumed that the scattering coefficient of the phantom mixture was independent of the amount of dye added. This would only be true if the total volume of diluted dye added to the tank was small compared to the total volume of Intralipid™ solution in the tank, which was originally intended to be 1300 ml. A set of pipettes were available that could measure from 10 μl to 2500 μl. A step size of 200 μl of dye was decided upon, as it was neither too small to make measurements inaccurate, nor so large that even with say 20 additions of dye (20 x 200μl = 4 ml), the scattering coefficient changes would only be about 100*4/1500=0.27%. A change in μ_a of 0.002 mm^{-1} for each 0.2 ml (200 μl) was chosen, so this required a μ_a for the dye of

$$\mu_a = \frac{\Delta\mu_a \; V_{tank}}{\Delta V} = \frac{0.002 \times 1300}{0.2} = 13 \; mm^{-1} \qquad \textbf{7.3}$$

where $\Delta\mu_a$ is the step change in absorption coefficient required (0.002 mm^{-1}), V_{tank} is the volume of the tank (1300 ml) and ΔV is the amount of dye added to cause the step change (200 μl = 0.2 ml).

Water had previously been added to the S109564 ICI dye, as the lid had been left off
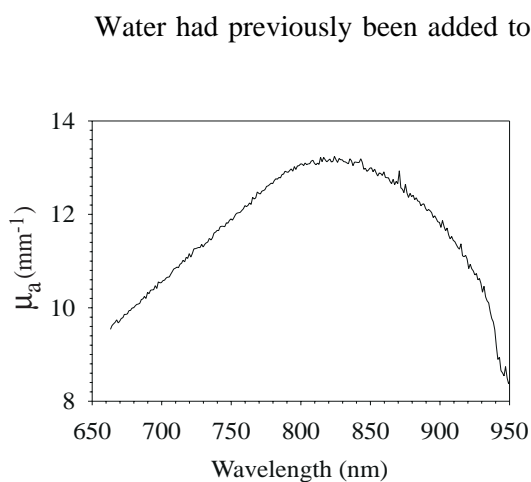
Figure 7.4 Absorption coefficient of the diluted dye, as measured on an optical spectrometer using a CCD camera detector.

and all the water had evaporated! Hence the optical properties of the dye were unknown, and could not be expected to be similar to any specifications published by ICI. The dye was far too optically dense for the intended use, so had to be diluted. The dilution required was determined retrospectively, after taking measurements of the dye at a known dilution using an infra-red spectrometer attached to a CCD camera. First the absorption spectrum of a 1 cm cuvette, filled with distilled water was

obtained. Then 1% ICI dye was mixed with 99% water (2.4 ml of dye + 237.6 ml of water = 240 ml total), as it was suggested to me by a colleague, Dr. Mark Cope, that 1% might be a sensible starting point. Next the 1% dye solution was further diluted by 50:1, before being put into the cuvette and measured in the CCD, as a 1% dilution would have been too optically dense to measure in a reasonable time. The μ_a of the 1 in 5000 mixture was measured at 0.036 mm^{-1} at 800 nm, indicating that the 250 ml of dye that had been mixed had a μ_a of 50*0.036=1.8 mm^{-1}. From this it was deduced that the amount of dye needed to be increased by a factor 13/1.8= 7.22 to increase the μ_a to 13 mm^{-1}. Hence extra dye was added. The absorption spectrum of the final dye mixture used is shown in figure 7.4, which is seen to be 13.08 mm^{-1} at 800 nm. Although originally it was assumed the tank would be filled with 1300 ml, in fact 1500 ml of phantom solution was mixed up. This was to allow the upper fibre to be moved upwards more than the 20 mm available on the slide mechanism, while still being well below the surface. The tank was filled with 1500 ml, so the step change in μ_a that was caused by the addition of 200 μl of dye with a μ_a of 13.08 mm^{-1}, was not 0.002 mm^{-1}, but instead 0.2*13.08/1500=0.00174 mm^{-1}.

Although it had originally been intended to use 200 μl additions of dye, it soon became apparent that making measurements with such small steps of μ_a would have been too time consuming. For this reason, it was decided to use additions of 500 μl, which caused step changes in μ_a of 0.5/1500*13.08=0.00436 mm^{-1}. As the diluted Intralipid^{TM} had a baseline μ_a of 0.0018 mm^{-1}, 500 μl of dye increased this to 0.0018 + 0.00436 = 0.006164 mm^{-1}. Table 7.1, on the next page, shows the absorption coefficient of the Intralipid^{TM}, for volumes of dye from 0 to 3000 μl, in steps of 500 μl.

Table 7.1 Absorption coefficient of the Intralipid^{TM}/dye mixture

Amount of dye added (μl)	μ_a (mm^{-1})
0	0.00180
500	0.00616
1000	0.0105
1500	0.0149
2000	0.0192
2500	0.0236
3000	0.0280
3000	0.0324

7.1.3   Refractive index of the optical phantom.

The refractive index of an Intralipid^{TM} phantom will be approximately that of water, which is 1.33[141] at 800 nm - slightly less than that of tissue (1.4)[30]. However, the differences in the refractive index can be taken into account in the any numerical or analytical fitting procedure.

Chapter 8.
Monte Carlo Model.

As was discussed in chapter 1, a Monte Carlo simulation is necessary to predict the TPSFs in the tissue. A Monte Carlo model is also required to deduce the TPSFs expected from the tissue-like phantom described in chapter 7, at close source-detector spacings. This chapter describes the development of a Monte Carlo computer model, measurements performed to gather parameters used by the model and procedures carried out to verify the accuracy of the model. The main set of results from the model are in the next chapter.

8.1   Fibre characteristics.
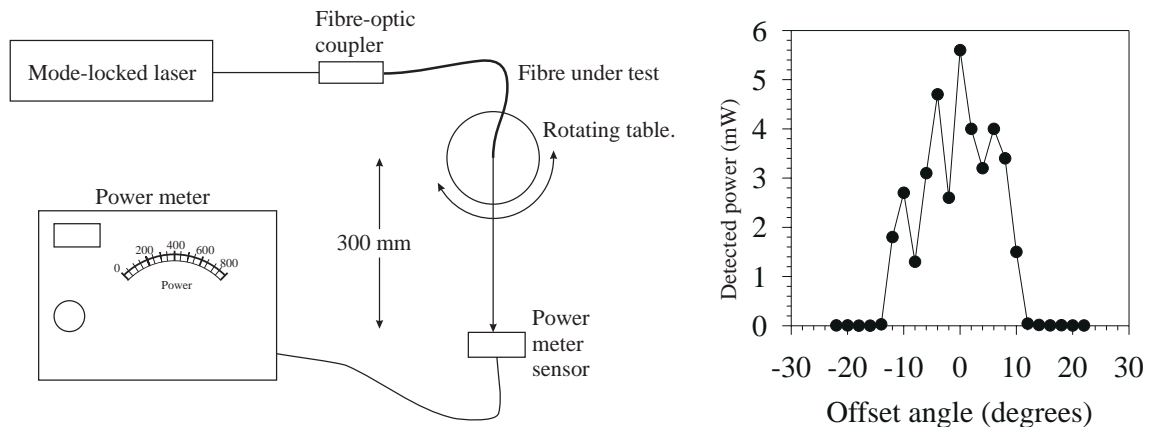
Figure 8.1 Experimental setup to measure the fibre characteristics, and the results obtained.

As the source and detector fibres would be fairly close in some of the actual measurements performed with the cross-correlator, the radiation characteristics of the optical fibres, which were 100μm graded index multi-mode fibres, would have a significant effect on the detected light. No information was available from the fibre manufacturer, so firstly the angular distribution of intensity was measured in the laboratory. The setup used is shown in figure 8.1. One end of the fibre under test was fitted with an FC-PC connector and mounted onto a fibre optic coupler which focused the laser light into the fibre. The far end of the fibre, which was 2 m long, was polished and mounted on a rotating table. The table was rotated in 2° increments as the intensity was measured on a Coherent model 212 optical power meter (Coherent, 2301 Lindberg St, Auburn, CA, 95603, USA). The Tsunami pulsed mode-locked laser

used in all the experiments was used, set to 800 nm, since the exact conditions of the input to the fibre - wavelength, degree of collimation, mode contents etc, may be expected to have some effect on the radiation pattern of the fibres.

The results are shown to the right hand side of figure 8.1. Although the radiation pattern from a single mode fibre can be described analytically, this is not so when several modes interact in a fibre. The individual modes superimpose in a random manner, which explains the apparent noisiness of the data in figure 8.1. Moving the fibre slightly, would generate a different pattern. Due to the speckle which is evident if the laser was directed onto a white card, the intensity showed huge changes for very small angular displacements.

The measurement of the fibre characteristics was then repeated, but this time using an electromagnetic vibrator attached to the fibre, to scramble the modes. The vibrator was essentially a speaker voice coil, with a DC resistance of 3 Ω. It was first driven from a 1 V, 100 A transformer, to avoid damaging the vibrator, which had no maximum voltage marked on the case, and no data sheet that could be found. Later this was replaced by a 6.3 V, 3 A transformer, as the 1 V generated insufficient movement. Power was only applied during the measurement, to reduce the possibility of thermal damage to the vibrator. The experimental setup is shown in
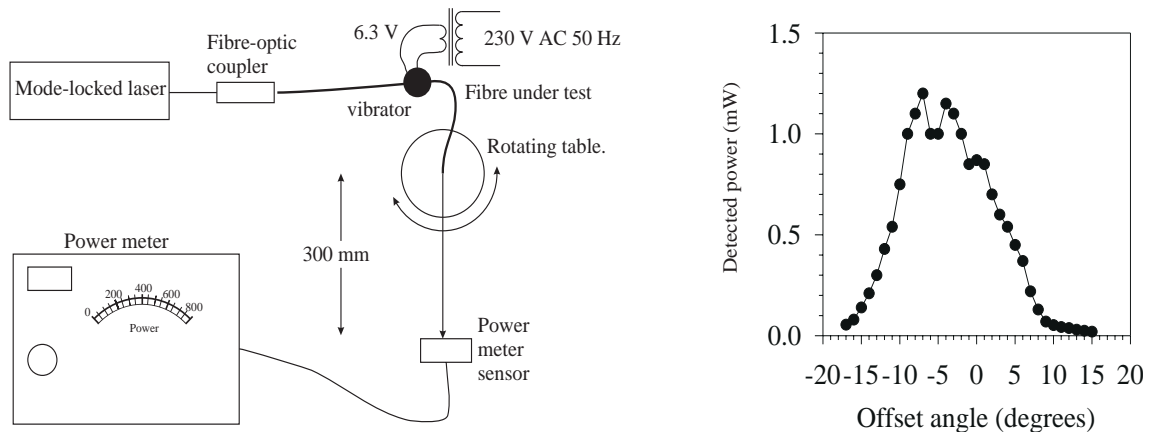
Figure 8.2 Experimental setup, and the results obtained, after attempts to scramble the fibre modes.

figure 8.2, along with the results. The fibre was assumed to be a Gaussian radiator, with a FWHM of 13°. The dip in the centre was ignored, as it was not reproducible between different samples of fibre.

It should be noted that since these results were measured in air, not tissue, a correction was necessary in accordance with Snell's Law before they could be applied in tissue. The

following relationships can be derived from Snell's Law:

$$n_{fibre}\sin\theta_{fibre} = n_{air}\sin\theta_{air}$$

8.4

$$n_{fibre}\sin\theta_{fibre} = n_{tissue}\sin\theta_{tissue}$$

8.5

where n_{fibre} is the refractive index of the fibre core, θ_{fibre} is the internal angle at which the light makes with the end of the fibre, n_{air} is the refractive index of air, n_{tissue} is the refractive index of tissue, θ_{tissue} is the angle at which light leaves the fibre in the tissue. Hence we can show that the following equation is satisfied:

$$\theta_{tissue} = \sin^{-1}\left[\frac{n_{air}}{n_{tissue}}\sin\theta_{air}\right]$$

8.6

Assuming the light exits the fibre with a FWHM of 13° in air, then substituting for n_{fibre}=1.5, n_{air}=1.0, n_{tissue}=1.4 and θ_{air}=13°, we can calculate that the fibre will radiate with a FWHM of 8.6° in tissue.

8.2   Experimental model simulated.

The experimental system simulated is shown in figure 8.3. A tissue volume of x by y by z mm, was assumed, with all sides orthogonal. The volume is broken into a number of voxels, X along the x axis, Y along the y axis and Z along the z axis, each with dimensions x/X by y/Y by z/Z (mm). The source and detector fibres usually had the same y and z coordinates (ie they were in-line, which was the case in the experimental phantom), although the detector fibre could be moved off axis from the source fibre. The Monte Carlo model was designed such that different regions could have different optical properties. Individual voxels could have the background optical properties (shown white) or any alternate value (shown grey to black). This was desirable, since it would allow testing of the effects of tissue heterogeneities. However all the results presented here were obtained assuming homogeneous optical properties. Additionally, the reflection of light at the tissue boundaries has not currently been taken into account, so the model currently assumes an infinite volume, although it is designed to allow easy extension to finite volumes. Currently, the model accurately mimics what happens in the liquid phantom described in the last chapter. That phantom is effectively an infinite volume, with two fibres in line.
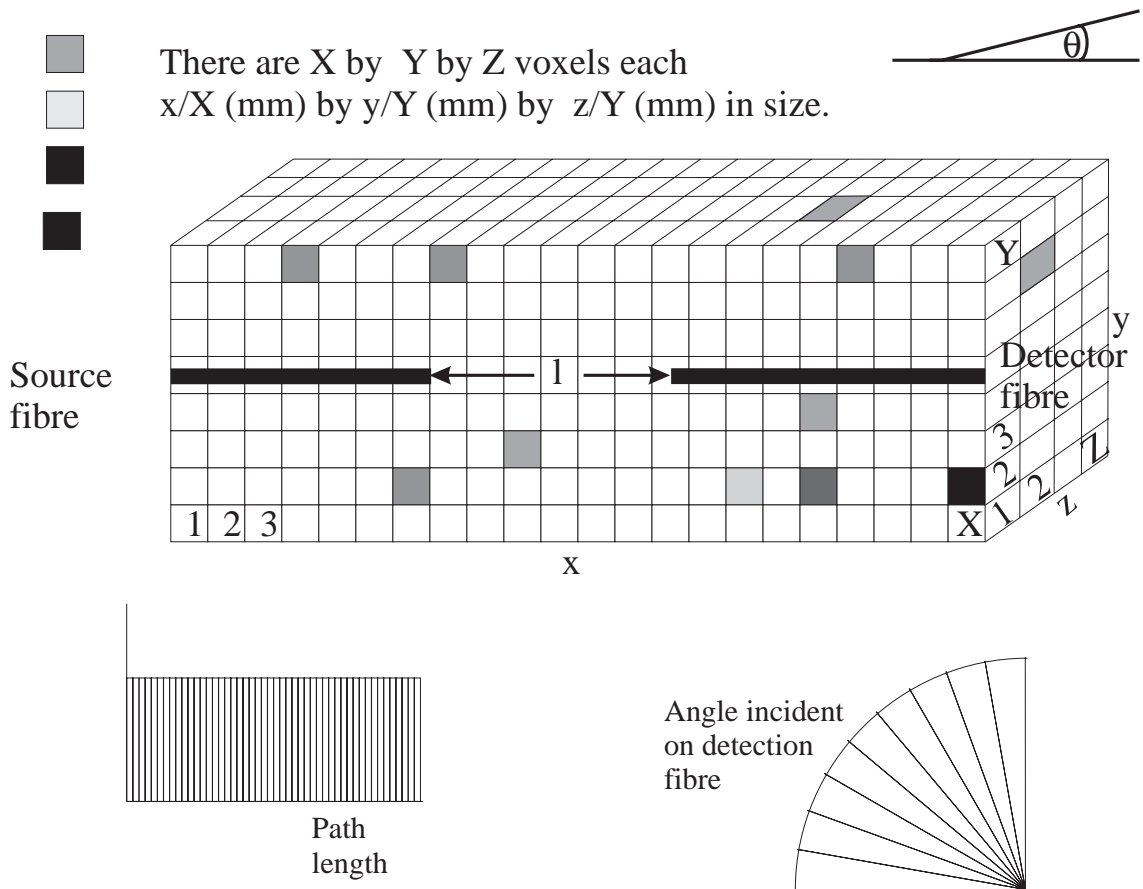
Figure 8.3 Diagram showing the experimental arrangement modelled with the Monte Carlo programme. A block x by y by z mm is represented by an arrangement of X by Y by Y voxels.

8.3 Mathematical details of the Monte Carlo model.
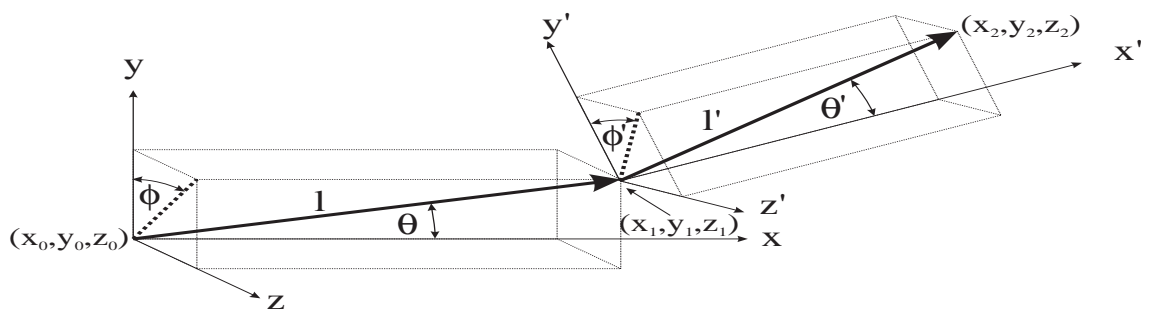
Figure 8.4 Diagram showing the definition of the scattering angles θ,ϕ and θ',ϕ'

Photons leave the source fibre at a position (x_0,y_0,z_0), inside the tissue. The photon's initial path was determined by the direction at which it left the fibre. This was described by two angles θ and ϕ (see figure 8.4). Photons were assumed to be launched from the source fibre with a mean

angle θ of $0°$, and a FWHM of $8.6°$ in accordance with the measured fibre characteristics shown in figure 8.2. The function *gasdev* (Gaussian Deviate) from the book Numerical Recipes in C[142] was used to generate such a random angle. The angle ϕ was set randomly between 0 and 2π by use of a random number N_{rand} uniformly distributed between 0.0 and 1.0.

$$\phi = 2\pi N_{rand}$$ **8.7**

The x, y, and z direction cosines of the photon relative to the fixed coordinate system x,y,z are then:

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta)\sin(\phi) \\ \sin(\theta)\cos(\phi) \end{pmatrix}$$ **8.8**

If the original position of the photon is (x_0, y_0, z_0), and the photon travels a distance l, its new position (x_1, y_1, z_1) is then given by equation 8.9.

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + l. \begin{pmatrix} \cos(\theta) \\ \sin(\theta)\sin(\phi) \\ \sin(\theta)\cos(\phi) \end{pmatrix}$$ **8.9**

After leaving the fibre at the angle (θ, ϕ), and travelling a distance l, the photon will scatter for the first time in the tissue, at an angle (θ', ϕ') relative to its previous direction as shown in figure 8.4. The scattering angle (θ', ϕ'), did not depend on the fibre characteristics (unlike θ), but instead on the scattering anisotropy factor g of the tissue. To find the scattering angle θ', the Henyey Greenstein formula[53] was used (equation 8.10), along with a random number N_{rand} uniformly distributed between 0.0 and 1.0.

$$\cos(\theta') = \frac{1}{2g} \left[1 + g^2 - \left(\frac{1-g^2}{1-g+2 g N_{rand}} \right)^2 \right]$$ **8.10**

To calculate the angle ϕ', which was assumed to be randomly distributed between 0 and 2π, another random number, again uniformly distributed between 0.0 and 1.0 was generated.

$$\phi' = 2\pi N_{rand} \qquad \textbf{8.11}$$

The direction cosines of the photon relative to its previous direction, but not the base coordinate system were then given by:

$$\begin{pmatrix} \cos(\theta') \\ \sin(\theta')\sin(\phi') \\ \sin(\theta')\cos(\phi') \end{pmatrix} \qquad \textbf{8.12}$$

The absolute direction cosines of the photon (ie relative to the base coordinate system) was then given by the matrix multiplication:

$$\begin{pmatrix} -\sin(\theta) & 0 & \cos(\theta) \\ \sin(\theta)\cos(\theta) & \cos(\theta) & \sin(\theta)\sin(\phi) \\ \cos(\theta)\cos(\phi) & -\sin(\theta) & \sin(\theta)\cos(\phi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta') \\ \sin(\theta')\sin(\phi') \\ \sin(\theta')\cos(\phi') \end{pmatrix} \qquad \textbf{8.13}$$

Hence the final position of the photon after assuming it then travels a further distance l' was:

$$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + l' \cdot \begin{pmatrix} -\sin(\theta) & 0 & \cos(\theta) \\ \sin(\theta)\cos(\theta) & \cos(\theta) & \sin(\theta)\sin(\phi) \\ \cos(\theta)\cos(\phi) & -\sin(\theta) & \sin(\theta)\cos(\phi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta') \\ \sin(\theta')\sin(\phi') \\ \sin(\theta')\cos(\phi') \end{pmatrix} \qquad \textbf{8.14}$$

where (x_1, y_1, z_1) was given in equation 8.9. This process of generating scattering angles and scattering lengths can be repeated as necessary. At each time the new position is found using the old position, and the scattering angles.

The length that a photon travelled after leaving the fibre before undergoing scattering was determined. This was the same as the distance from any one scattering site to the next. If the scattering coefficient is μ_s, the probability that no interaction occurs for a photon travelling a distance l is $\exp(-\mu_s l)$. The probability of a photon being scattered in a length less than l, is then given by $p(l) = 1.0 - \exp(-\mu_s l)$, where $p(l)$ must obviously be between 0.0 and 1.0. In order to find a number of random scattering lengths with the correct distribution, $p(l)$ can be replaced by a random number N_{rand}, such that $0.0 \leq N_{rand} \leq 1.0$. Therefore

$$N_{rand} = 1.0 - \exp(-\mu_s l) \qquad \textbf{8.15}$$

Inverting this gives

$$l = \frac{-\ln(1.0 - N_{rand})}{\mu_s} \qquad \textbf{8.16}$$

Given N_{rand} is uniformly distributed between 0.0 and 1.0, we can replace $1.0 - N_{rand}$ by N_{rand}. Hence the formula used to calculate the distance between scattering sites is:

$$l = \frac{-\ln(N_{rand})}{\mu_s} \qquad \textbf{8.17}$$

Photons in this model were assumed to have an initial *weight* of 1.0, but reduced with distance travelled in the tissue as $\exp(-\mu_a \Sigma l)$, where Σl is the sum of the distances travelled. A photon therefore became progressively weaker and was followed until it became very weak. However, it was never disregarded, as this would have resulted in lost energy. Instead, if the *weight* of a photon fell below an internal parameter called INTENSITY_CUTOFF (which is 10^{-12} in the programme), the photon was then given a 10% chance of surviving, and increasing its *weight* by a factor of 10. Ninety percent of the time the photon did not survive, and so was terminated, and a new photon path followed. This method of photon termination is called *Russian Roulette*, and is described by Lux et al in greater detail[51].

Using weights for photons is much more computationally efficient than giving each photon a probability of surviving after each scattering event. If this were done, a lot of computational effort would have been expended in following a photon, that eventually became absorbed and therefore contributed no data.

The programme calculated the total path length Σl a photon travelled, in 1 mm increments from 0 to 499 mm, to enable TPSFs to be displayed. The time taken for a photon to reach the detector is:

$$t = \frac{n \; \Sigma \; l}{c} \qquad \textbf{8.18}$$

If, for example, a photon travelled a total distance of 51.1 mm, to cross the fibre 10 mm away from the source, this would have been put into a bin with pathlengths from 51 to 52 mm. The mean pathlength of this bin, 51.5 mm, was used to calculate the arrival time of the photon. This would be 240.33 ps assuming a refractive index of 1.4. The position on the detector fibre was recorded in 1 mm increments from 0 to 79 mm, so for this example would have been recorded as 10 mm. If the absorption coefficient of the tissue was 0.03 mm^{-1}, the *weight* recorded for this

photon would be exp(-μ_a Σl)=exp(-0.03*50.1)=0.222. The angle the photon made to the detector fibre was also recorded. The photons are binned into bins 1° wide up to 12°. Any photons arriving outside 12° are put into a bin for photons from 12 to 90°. The detection fibre would not be expected to detect such photons, as they would fall outside the acceptance angle.

Once a photon was detected crossing the fibre and this recorded in the programme, this was not the end of the photon. Instead the photon was continued until it was terminated by the Russian Roulette technique described on page 191. It was assumed that to continue following the photon would introduce negligible error with very small detection fibres, yet lead to a large increase in computational efficiency. This enabled the Monte Carlo model to calculate the intensities at all points along the x axis (in 1 mm increments), without having to know exactly where the detection fibre was to be placed. However, if a photon was detected twice (or more times) at the same position on the fibre, subsequent detections were ignored.

The programme calculated the TPSFs for a fibre in line, as shown in figure 8.3. It also calculated the TPSFs for a detector fibre 5 mm in front of the source fibre, but at 90°, as shown in figure 8.5. Data was collected in 1 mm steps from 0 mm to 79 mm, although only the first 9 mm are shown in figure 8.5. It was intended to use this data, as one might expect the TPSFs to be longer, for the same source-detector spacing, if the detector is at 90° to the source fibre. For photons to be detected on this orthogonal fibre, they would have needed to be scattered through approximately 90°, which one might expect to require more scattering events, and so take longer. This might be an advantage when using instrumentation with a wide impulse response, as the TPSFs would be subject to

Figure 8.5 Diagram showing how a second detection fibre, at 90° to the first one, is used to measure light that is likely to be more highly scattered.

less blurring by the instrument. Unfortunately, time did not permit analysis of this data, although it has been collected.

8.4   **Parallel processing of the Monte Carlo computer programme.**

Given the small area of the detector (200 μm diameter), most light emitted from the source fibre would not find its way to the detector. Hence any simulation of the problem would be very slow. For this reason, the programme was written so that it could be run simultaneously on a number of computers, of various types (PC's, workstations etc) on one computer network that was accessible to all the computers. This reduced considerably the time needed to perform the simulation, as it was performed on a number of computers in parallel.

This method was developed as a result of problems encountered with an earlier attempt to use Monte Carlo technique to simulate Laser Doppler blood flow measurements, where thirteen different Sun workstations were used[143]. Unfortunately, with many computers writing to many different data files, it became difficult to monitor the progress of the calculation. The random nature of the Monte Carlo method meant that each computer could be writing data to its files at anytime, making it impossible to safely add the data files from many computers for analysis, without first stopping the processes, summing the data, then restarting the computation, which was very tedious and error prone. A brief description of the parallel implementation of the programme developed to overcome this problem follows.

The description of the problem to be solved was held in one small ASCII file. This contained the measurement geometry, tissue optical properties (μ_a, μ_s and g), position of source and detector, fibre characteristics etc. The Monte Carlo simulation software read this from disk, performed the simulation, then wrote the photon history results, consisting of intensity, position, pathlength and angle with respect to the detection fibre into a 4 dimensional array of memory. After a minimum of 1000 photons were launched, the results were written to a file on a disk accessible to all computers on the network. One copy of the programme ran on each computer available, usually at reduced priority when being executed on Unix based machines that implement scheduling priorities.

Only one results file was stored on the computer network. Each computer performed the Monte Carlo calculations and wrote the photon history results into this one results file. Since two computers editing the one file could easily result in file corruption, the time of the day at which any computer could update the data file was fixed. The first computer to execute the programme wrote a file named *0* into the directory where the data was stored. The file did not need to contain any data at all, although when the computer was running under the Unix operating system, the machine's name and the process identification number (PID) were written into the file in ASCII to enable the user to determine on which machines the programme had been

started. The second execution of the programme, on finding the file named *0*, wrote a file named *1*. A third execution of the programme wrote a file named *2* and so on, up to the 24th execution, which wrote a file named *23*. The source code was written such that the programme that wrote file *0*, could only store data to disk between the times of midnight and 00:55 am (until this time the programme continued to run and accumulate data into the local machines memory). The second programme to run (which wrote the file *1*) could only store data between the times of 01:00 and 01:55, the third stored data between the times of 02:00 and 02:55 etc. This ensured that each execution of the programme could only edit the data file at a time when no other programme was doing this. Obviously, this limited the number of separate executions of the programme to 24, but even this could be extended if necessary, by using smaller time steps than one hour. Conversely, if the number of computers available was smaller (say 6) it would probably be preferable to write the data once every six hours, rather than once per day, as this would reduce the time that could elapse without the data file being updated. For safety, no programmes could edit the data file in the last 5 minutes of any hour, since time must be allowed for writing the file. If for any reason (such as hardware failure) a process was unable to write the data to disk in the 55 minutes allocated for the task, the data was retained (and continued to further accumulate) in the computer's random access memory (RAM) until the next day, when another attempt would be made to write the data to disk.

An important feature of the programme is the way that one data file could be updated by any computer on the network, which consisted of both Sun workstations and PC's. Initially, the photon data was stored in the random access memory (RAM) of each computer in whatever format the computer normally used. This would be *big-endian* on a Sun computer, but use *little-endian* on a PC[144], as shown in figure 8.6.

0	7	8	15	16	23	24	31
12_{16}		34_{16}		56_{16}		78_{16}	

4-byte integer 12345678_{16} in 'big-endian' memory format

0	7	8	15	16	23	24	31
78_{16}		56_{16}		34_{16}		12_{16}	

4-byte integer 12345678_{16} in 'little-endian' memory format

Figure 8.6 Diagram showing how a 4-byte integer (standard integer on 32 bit computers) would be stored in memory in both *big-endian* and *little-endian* formats.

Before writing data to disk, there was a choice of using either binary or text. Binary is much more efficient in its use of disk space, using only 8 bytes to write a double precision floating point number (the data type used for most of the data), compared to 16 bytes if text were chosen. It is also simpler to read and write binary data, since one can also know exactly where the data will be in a disk file - something that is not so easy to achieve when data is written as

text. Unfortunately, since PCs and Suns store binary data in a different format, storing binary data in a way that both PCs and Suns can access it, is not without its problems. To circumvent this, data written to disk was always stored in the Sun (*big-endian*) format. Any PC version of the programme stored data internally in RAM in the *little-endian* format, but reversed the order of the bytes to convert to *big-endian* format before writing to disk.

Before a programme wrote data to a disk, it first ensured that the data in the data file was for the same conditions - ie geometry, tissue optical properties etc. Both the number of launched and detected photons were incremented along with all intensity and pathlength data from the programme.

A paper describing this parallel Monte Carlo programme was published in 1997[145].

8.5   Random number generation.

Ideally, a true random number generator was required for the Monte Carlo model. Such a set of random numbers is obtainable by for example counting the decay interval from a radioactive source, the fluctuation in current from a diode (shot noise) or other physical process. A set of numbers generated by using the radioactive decay method is available on CD-ROM. However, there is only a finite number of these, and it was decided to use a pseudo random number generator instead. These are much more convenient to use, requiring no CD ROM to be mounted and also make the code much more portable.

With the obvious difficulty of producing random numbers from a computer that performs only deterministic operations, a different approach was needed. The formula in equation 8.19 is known as a linear congruential generator[142,146].

$$X_{n+1} = (A\,X_n + C)\bmod M \qquad\qquad \textbf{8.19}$$

Starting with some seed (starting value) X_0, this can with appropriate choice of A, C and M, generate a set of pseudo random numbers with a period of up to M-1, although if A, C and M are poorly chosen, the period would be much less. Typically the initial seed X_0 would be set to the number of seconds that have elapsed since 1st January 1970, as most computers have a function that will give this information. It is important to note that the correct values to use for A, C and M depended on the word length of the computer. For example, the following is used in the Turbo Pascal Compiler,

$$X_{n+1} = (8088405X_n + 1) \bmod 16304 \qquad \textbf{8.20}$$

and while this is acceptable for the 16 bit integers used in its computation, its performance is poor if implemented on a 32 bit word machine, with a 32 bit compiler.

No pseudo random number generator of the form in equation 8.19 is good enough for Monte Carlo simulations, although they are fine for games and other non-critical applications. The usual method used to rectify this problem is to use a linear congruential pseudo random number generator and subsequently shuffle the numbers obtained[142]. There are other methods of generating good sequences of pseudo random numbers, that do not use a linear congruential generator at all, such as the function ran3() in the second edition of the book Numerical Recipes in C[142].

The pseudo random number generator used was a modification of the C function ran1()[146]. The published design returned one of 2^{32} different numbers, with a repeat period of 2^{32}-1. This obtained its random numbers using the system supplied random number generator (RNG) of the form shown in equation 8.19. Using integers on a 32 bit computer, the system RNG can only return one of 2^{32}-1 different numbers. No amount of shuffling can improve upon this. The approach taken here was to use instead the function *drand48* available on unix systems, that can return one of 2^{48}-1 numbers, then shuffle them. Thorough analysis of this method was performed in an M.Sc. thesis in 1993[143]. The method was concluded to be excellent for Monte Carlo simulations.

There is one area of concern with the programme, which has not been completely addressed. If several copies of the programme are running on a computer network, all started at different times, there is no guarantee that the sequence of random numbers generated by the random number generators are independent. It could happen that two copies of the programme are using the same sequence of numbers. This is most unlikely, as the period of the random number generator (2^{48}-1=2.8x10^{14}) is very large compared to the number of random numbers used. Each copy of the programme simulated at most 25x10^6 photons (see table 8.1) which scattered at most 300 times, requiring at most 900 random numbers per photon launched (one for θ', one for ϕ' and one for the scattering length l). Hence the largest simulation required less than 2.25x10^{10} random numbers. It is most unlikely that two simulations that required 2.25x10^{10} random numbers, from a generator with a period of 2.8x10^{14}, would use any numbers in common.

One possibility considered was to generate all the random numbers first and put these

into one file. Then each programme could use part of this file. However, the largest simulation, that required 2.25×10^{10} random floating point numbers, each 8-bytes long, would have required a disk file of 180 Gb - not very practical with current computers.

8.6 Data compression.

The Monte Carlo programme was called *fibremc* and saved data in one binary file of length 8,323,320 bytes. This consisted of a header (3,320 bytes) and the data itself (8,320,000 bytes).

The header consisted of 70 4-byte integers (280 bytes) and 380 8-byte floating point numbers (3040 bytes), making a total size for the header of 3320 bytes. The integer numbers contained such things as the number of photons launched, number of photons detected, number of objects with different optical properties to the background, number of position slots (usually 80), number of time slots (usually 500). The floating point header contained such things as the refractive index of the tissue (usually 1.4), the version of the programme, the dimensions x, y, and z, μ_a, μ_s, g, μ_s', μ_{eff}, μ_t, source position, detector position, detector diameter, the source fibre radiation pattern etc.

The 8,320,000 bytes of data following the header consisted of a 4D array of 80 position slots (0 to 79 mm), 500 time slots (0 to 499 mm), 13 angular slots and 2 detection fibre geometries (one along the x-axis as in figure 8.3, and the other at 90° to this, as in figure 8.5). Each of the data elements was an 8-byte floating point number, making a total storage requirement for the data (excluding the header) of 80 x 500 x 13 x 2 x 8 = 8,320,000 bytes, or 8,323,320 bytes including the header. The data was always stored with the same byte ordering (*big-endian*). While storing one such 8 Mb file was not a problem with current computers, it did become a problem if there are lots of different data files, all around 8 Mb in length, as would happen when running many simulations with different properties on the same computer network. The Monte Carlo programme was written such that the 8 Mb of data was always stored in RAM. When it was necessary to write it to disk (as it is advisable periodically), the 8 Mb data file was written, then a Public Domain compression utility (zip or pkzip) was used to compress the data by removing redundant entries. Then the original uncompressed data file was removed. Typically the data could be compressed to 20 % of its original size. If the programme needed to read the data from disk, it first decompressed it, then used the decompressed data, then re-compressed it before again removing the large uncompressed file. Using this method, it was possible to have 10 or more copies of the main Monte Carlo programme running, while using a maximum disk

storage space of around 25 Mb rather than 80 Mb for data files. Obviously, if all the programmes happened to require the uncompressed file at the same time, it would need the full storage requirement, but in practice using compression, the total disk storage requirement could be kept within manageable limits. Since the time required to compress or decompress the data was a very small fraction of the time between which the full uncompressed data file was needed, the system saved storage space without a significant performance penalty.

Given that the source code for compression programmes is readily available, it would have been more computationally efficient if the data was compressed by the Monte Carlo programme directly, rather than using an external programme. Since the time spent compressing/decompressing data was quite small, any gains from more efficient processing would be modest.

8.7   Testing of the programme.

With any numerical programme such as this, it is clear that programme bugs can easily occur and further that the effect of these will often go undetected. This can result in the data being inaccurate. A number of tests were performed to increase confidence in the model, although it is impossible to say with absolute certainty that there are no errors remaining.

8.7.1   Testing of the Pseudo-Random number generator.

If all scientific papers whose results are in doubt because of bad rand()s were to disappear from library shelves, there would be a gap on each shelf about as big as your fist.

W.H. Press et al, Numerical Recipes in C, Cambridge.

The random number generator has been tested previously[143], although the results of some of the tests are shown here.

8.7.1.1   Serial Correlation between the digits of the RNG.

The basic concept of this test is to call the random number generator k times and use the k different values to define a point in k dimensional space. One then plots the point on a k dimensional graph. This is repeated n times, making a total of k n calls to the random number generator. Obviously, using k > 2 gives problems in plotting the data and k > 3 gives serious problems! However, according to Wood[147] plotting k > 2 gives little more information than the k = 1 or k = 2 case. Figure 8.7 is a 2D plot for 100,000 calls from the generator dran0(). The random number generator was called twice per point - once for the x coordinate and once

for the y coordinate, as shown by the following pseudo-code.

```
for(i=1 to 100000)              /* Loop 100,000 times */
{
        x=random();             /* Use dran0() to generate a random number */
        y=random();             /* Use dran0() to generate a random number */
        plotpoint(x,y);         /* Plot the point (x,y) in white on a black background */
}
```

Note that there was no obvious pattern to the image in figure 8.7.
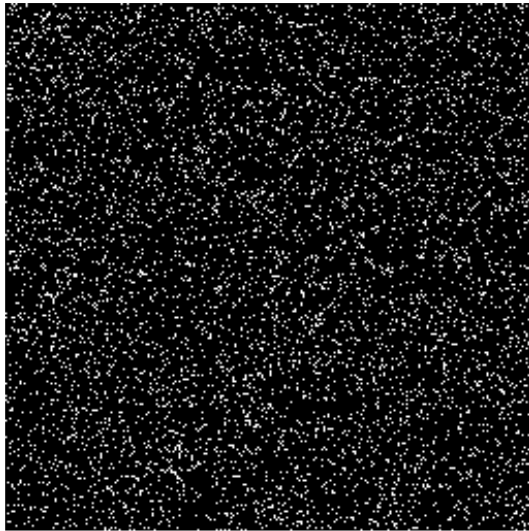
Figure 8.7 Two dimensional plot showing the absence of any serial correlation on the random number generator used in the Monte Carlo simulation software. Compiled with a 32-bit compiler and executed on a 32-bit computer.
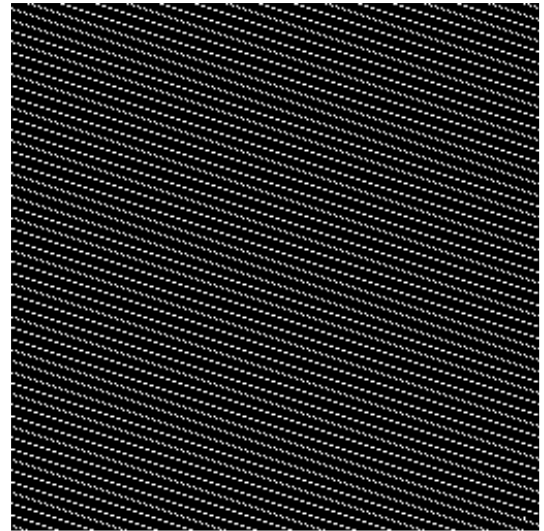
Figure 8.8 A scatter diagram for the RNG given in equation 8.21, compiled with a 32–bit compiler and executed on a 32-bit computer. If compiled with a 16-bit compiler, the results are not as poor.

An example of serious serial correlations in the following simple linear congruential pseudo random number generator:

$$X_{n+1} = (8088405X_n + 1) \bmod 16304 \qquad \textbf{8.21}$$

is shown in figure 8.8. There are clearly nowhere near 100,000 different points on the graph! Each point was plotted over and over again. The random number generator in equation 8.21 is used in the 16 bit Turbo Pascal 4.0 compiler[148]. However, the results are much better if implemented on a 16 bit compiler, indicating that a random number generator that works well on one machine cannot be guaranteed to continue to work well if the word length of the computer and/or compiler is changed.

8.7.1.2 Frequency test.

The random numbers produced by the generator should be uniformly distributed. To test this, the random numbers range (in this case 0.0 to 1.0) was split into a number (n=1000) of equally sized intervals (0.000 to < 0.001, 0.001 to <0.002, 0.002 to <0.003 ... 0.999 - 1.000). The random number generator was then called a number (N=2,000,000) times. The expected population in each interval is roughly equal to N/n=2000. The χ^2 test[149] was used to compare the observed frequency distribution with the expected distribution. The actual distribution is shown in figure 8.9. The statistic χ^2 was computed for the data, and found to be 897.0. The

values of χ^2 was computed for 1000-1=999 degrees of freedom, with the aid of Mathematica[150]. $\chi^2_{0.99}$ =1105.5 and $\chi^2_{0.01}$=898.0. Since the measured value of χ^2 < 1105.5, it can be concluded that the observed distribution of numbers does not differ significantly from the expected distribution. Since χ^2 > 898.0, it follows that agreement is not so exceptionally good that it should be looked upon with suspicion. Hence there is no reason to reject the random number generator used here on the grounds of its frequency distribution.

Figure 8.9 Frequency distribution of the random number generator used for the simulations.

8.7.1.3 Gap test.

There are two forms of the gap test[147]. For the test used here, a large number N of random numbers was produced and grouped into b bins, depending on the range of the random number. For the tests, 50,000 random numbers were tested, each grouped into 1 of 10 bins (N=50000, b=10). Numbers less than 0.1 were grouped into the first bin, those at least equal to 0.1, but less than 0.2 were grouped into the second bin, and so on. The gaps between occurrences of the same digit were measured. A gap of 0 indicated adjacent digits, a gap of 1 indicated one different number between two similar digits and so on. For N digits, the minimum gap is 0 and the largest is N-2. Assuming a random sequence of numbers, the probability of any gap can be calculated. By multiplying the probability by the number N, the expected number for each gap could be calculated. By statistical comparison of the measured gaps, with those expected from theoretical calculations, using the χ^2 test, it can be determined if there is reason

to believe the pseudo-random number generator gives different results from a true random sequence. If we denote the probability of a gap of g as p(g), then:

$$p(0)=(\frac{1}{b})^2 \qquad p(1)=(\frac{1}{b})^2(1-\frac{1}{b}) \qquad p(2)=(\frac{1}{b})^2(1-\frac{1}{b})^2 \qquad \textbf{8.22}$$

and in general:

$$p(g)=(\frac{1}{b})^2(1-\frac{1}{b})^g \qquad \textbf{8.23}$$

so the expected number with a gap g, E(g) is:

$$E(g)=N(\frac{1}{b})^2(1-\frac{1}{b})^g \qquad \textbf{8.24}$$

In the following example, which is included only to illustrate the method, 20 numbers were converted into 3 groups (N=20, b=3). The gap test allows any number of groups to be used. As was stated, 10 were used to test the random number generator, but only 3 are used in the following example. Assuming the pseudo random number stream was:

```
2 1 2 0 2 2 0 0 2 0 1 0 1 0 2 0 0 2 2 1 0
then the gaps are:
```

| | | | | | | | Gap length | | | | | | |
Digit	0	1	2	3	4	5	6	7	8	9	10	11	>11
0	2	4	1	1	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	1	0	1	0	0	0	0
2	2	1	2	0	0	1	0	0	0	0	0	0	0
Expected	2.2	1.5	1.0	0.6	0.4	0.3	0.2	0.1	0.1	0.1	0.0	0.0	0.1

χ^2 = 28.4

This trivial example, has (3x13)-1=38 degrees of freedom,

$\chi^2_{0.99}$ for 38 degrees of freedom = 61.1

$\chi^2_{0.01}$ for 38 degrees of freedom = 20.7

Since the calculated value of χ^2 is < $\chi^2_{0.99}$, there is no reason to suspect that the gaps do not come from a random source of numbers. As χ^2 > $\chi^2_{0.01}$, the measured gaps are not so perfect to suspect that this example data is dubious.

For the tests of the random number generator, the 50,000 numbers were put into 10 bins and gap lengths measured from 0 to 17, and for >17. Hence there were (19x10)-1=189 degrees of freedom. Calculations of χ^2 used a floating point format, accurate to 16 significant figures.

χ^2= 164.9

$\chi^2_{0.99}$ for 189 degrees of freedom = 237.1

$\chi^2_{0.01}$ for 189 degrees of freedom = 146.7

The value of χ^2 is between that of $\chi^2_{0.99}$ and $\chi^2_{0.01}$ and therefore shows no cause for concern.

Work then commenced on testing the data produced by the Monte Carlo programme *fibremc*. Whenever problems were noted, the cause was ascertained and a quick re-check of previous results performed, since fixing one problem can create another - and occasionally did!

8.7.2 Comparison of the Monte Carlo model with the Beer-Lambert Law.

For the case of zero scatter, the intensity distribution predicted by the Monte Carlo model can be tested against an analytical model - the Beer-Lambert Law given on page 39, which predicts a fall in intensity as $\exp(-\mu_a \Sigma l)$, where Σl is the total distance travelled from the source. The same intensity distribution should result if the model includes scatter, but the scatter is always in the forward direction (ie g=1).

The model was tested with three different scattering coefficients of 0.0, 0.2 mm^{-1} and 1000 mm^{-1} (with g=1), and three different absorption coefficients of 0.1 mm^{-1},

Figure 8.10 Graphs showing the intensity theoretically predicated by the Beer-Lambert Law (solid lines) and results from the Monte Carlo simulations (filled circles). g=1.

1 mm^{-1} and 10 mm^{-1}. In each of the nine tests, for which 10,000 photons were launched, the intensity was calculated at up to 49 mm from the source. The resulting intensity distributions were independent of μ_s but showed a variation with μ_a as shown in figure 8.10. Also plotted on figure 8.10 are solid lines showing the theoretical distribution. The agreement between theory and the computer simulation is excellent.

8.7.3 Testing the accuracy of the fibre detection geometry.

As a second test, the detector fibre diameter was set to 1.99 mm and the detector moved off axis from the source by 1.0 mm. The anisotropy factor *g* was again set to 1.0, so all photons should be forward scattered and hence miss the detector by 0.01 mm. No photons should then have been detected, and this result was confirmed. The detector diameter was then increased to 2.01 mm, so all photons should have been detected, which they were. This test, like all others,

does not prove the model's computer implementation to be free from errors, but the results do provide a significant degree of confidence in the computer programme.

8.7.4 TPSF of a forward scattering (g=1) medium.
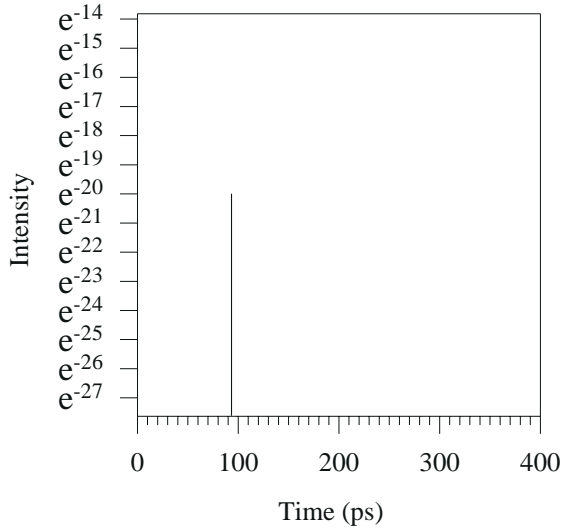
Figure 8.11 Temporal profile of light predicted by the Monte Carlo programme for a detector placed 20 mm from a source in a medium of μ_a=1 mm^{-1}, and a refractive index of 1.4.

Having proved that the intensity distribution for the case when g=1, was producing the correct results, the temporal distribution then needed to be tested. To do this, the intensity distribution as a function of time was plotted for the photons detected at 20 mm in front of the source, using g=1, μ_a=1 and μ_s=1. Assuming unit input intensity, the detected intensity should be exp(-$\mu_a \Sigma l$) = e^{-20}. Since the source and detector are separated by 20 mm, the light photons should arrive at a time given by equation 8.25. The refractive index n was assumed to be 1.4 for this simulation.

$$time = \frac{distance}{velocity} = \frac{l}{c/n} = \frac{nl}{c} = \frac{1.4 \times 0.02}{3 \times 10^8} = 93.33 \; ps \qquad \textbf{8.25}$$

Figure 8.11 shows the TPSF obtained from the Monte Carlo programme. The programme predicted photons to arrive at the correct time (93 ps after leaving the source) and with the correct intensity (e^{-20}). By using μ_s=1 mm^{-1}, the photons will have scattered a number of times before reaching the detector, so any problem with photons being scattered with the wrong intensity or direction would have shown up in this test. However, since g=1 for this test, θ was always zero, so it is possible that problems may exist when *g* is not 1.0. Hence further tests were performed.

8.7.5 Comparison with the time-independent diffusion equation.

The output intensity predicted by the Monte Carlo programme was then compared to that

derived from the time-independent diffusion equation for an infinite medium, given in equation 1.37. Figure 8.12 shows the diffusion equation prediction and the Monte Carlo data for μ_a=0.025 mm^{-1}, μ_s=20 mm^{-1}, g=0.9, μ_s'=2 mm^{-1}, when the intensities were matched at two mean free path lengths in front of the source. Since the mean free path, is 0.5 mm (given by $1/\mu_s'$ =1/2 mm^{-1} = 0.5 mm), the data is matched at 1 mm from the source. The trend of the Monte Carlo data follows the diffusion model well.
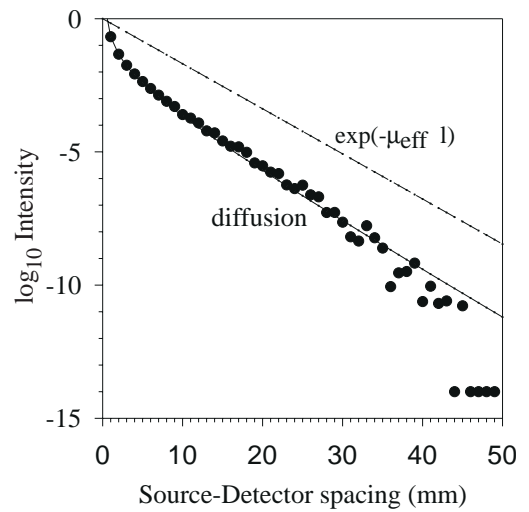
Figure 8.12 Graphs showing a comparison between the Monte Carlo data (filled circles) and that from the time-independent diffusion equation (solid line) when matched at 2 mean free path lengths in front of the source.

8.7.6 Comparison with the time-dependent diffusion equation.
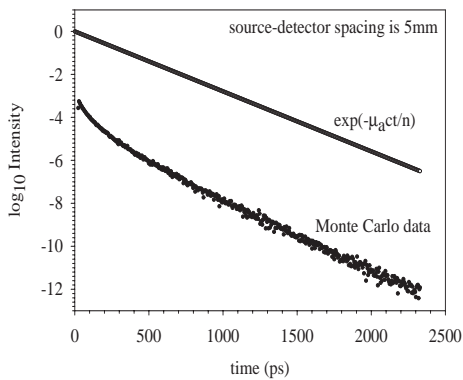
Figure 8.13 Comparison of the time-resolved Monte Carlo data with the slope of exp(-$\mu_a$ct/n). μ_a=0.03 mm^{-1}, μ_s'=1 mm^{-1}. Source-detector spacing is 5 mm.
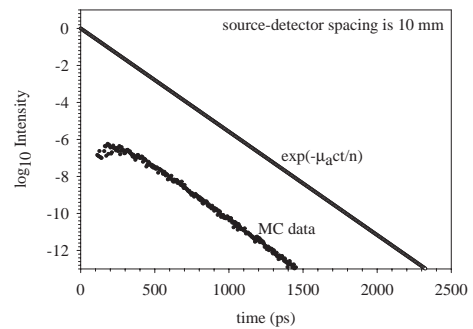
Figure 8.14 Comparison of the time-resolved Monte Carlo data with the slope of exp(-$\mu_a$ct/n). μ_a=0.06 mm^{-1}, μ_s'=3 mm^{-1}. Source-detector spacing is 10 mm.

Although the time-independent diffusion equation (equation 1.37) agreed well with the Monte Carlo predictions, further tests were performed to compare the results against the time-dependent diffusion equation for an infinite medium given in equation 1.34. Two graphs (figures 8.13 and 8.14) are shown, at different values of μ_a, μ_s' and source-detector spacing. Figure 8.13 shows the comparison with optical properties of μ_a=0.03 mm^{-1}, μ_s=10 mm^{-1}, g=0.9, μ_s'=1 mm^{-1},

at a source-detector spacing of 5 mm. Figure 8.14 shows a different simulation, using $\mu_a=0.06$ mm^{-1}, $\mu_s'=3$ mm^{-1} and at a source-detector spacing of 10 mm. Also plotted in figures 8.13 and 8.14 is the line exp(-μ_a c t/n). The limiting gradients at large times, of the Monte Carlo data can be seen to be equal to the gradients of exp(-μ_a c t/n), as predicated by the time-dependent diffusion equation for an infinite medium (equation 1.34).

8.8   Choice of Computers for the simulations.

As previously stated, there was a network of Sun workstations and PCs on which the simulations could be performed. Since photons scatter far less at low μ_s than at high μ_s, simulations with a low μ_s require less computational effort for the same statistical significance of results. Up to 21 computers were used on occasions for the simulations. The fastest machines (Sun Ultras running at 300 MHz) were used for simulations with the highest μ_s' and the slowest computers at the lowest μ_s'. Table 8.1 lists the number of photons launched, detected and the speed of the computers used. In all cases g was 0.9, so $\mu_s = 10 \mu_s'$.

Table 8.1 Numbers of photons simulated for the Monte Carlo data files

μ_s' (mm^{-1})	CPU speeds (MHz)	Launched	Detected
0.2	25-40	15,219,000	31,505,268
0.3	40-50	25,111,000	41,737,296
0.5	60-70	15,260,000	22,099,537
1.0	70-150	10,603,000	16,637,567
1.5	233-300	10,055,000	19,436,065
2.0	270-300	9,677,000	19,930,914
3.0	300	5,260,000	12,608,161

Even though all the Monte Carlo programs ran for approximately the same time (about 6 weeks) and the fastest computers were used for the highest values of μ_s', the number of photons simulated for the highest value of μ_s' is significantly less than the number at the lowest μ_s'. Since the exact run time of the different programs was not identical, and the load the machines were under from other users would not be identical, the numbers should only be used

as a guidance. In particular, it should be realised that others in the department were aware of the relative performance of the various Sun workstations and so the fastest machines are more heavily used.

8.9   Post-processing of Monte Carlo data.

A number of other programmes (some in C, some unix shell scripts) were written to post-process the 8,323,320 byte binary output file generated by the Monte Carlo program *fibremc*, to get the data in convenient forms. Several of the programs are described here. The programme *musfit* is described in the next chapter.

8.9.1   Post-processing programme *anal6*.

The 8,323,320 byte binary file created by *fibremc* could not easily be analysed with any common graph plotting tool. In order to get the data into a text file, a programme *anal6* was written, that read the binary file and wrote the data as text in a convenient form. Generally the data was written into a file with 17 columns, the first of which was position (mm), the second time (ps). The next 13 columns were data either plotted against position (column 1) or time (column 2). The final 2 columns gave some extra information, such as filenames, optical properties, number of photons simulated etc.

8.9.2   Post-Processing programme *muascale*.

When calculations were being performed, a set of optical parameters (μ_a, μ_s and g) were used. If data for a different μ_s was required, then it was necessary to re-run the simulation - a potentially very time consuming process. If however, a new simulation is required with the same μ_s, *g* and hence μ_s', but different μ_a, then re-running of the simulation can be avoided, if the medium is homogeneous. For the first simulation, using $\mu_a = \mu_{a1}$, light is attenuated due to absorption, so at any one position, incident angle, time etc, the intensity is given by

$$I = I_0 \exp(-\mu_{a1} \Sigma l_i)$$

where Σl_i is the sum of the individual pathlengths. A second simulation with a new absorption coefficient μ_{a2} would give an intensity at the same position, detector angle etc of

$$I = I_0 \exp(-\mu_{a2} \Sigma l_i) \hspace{3cm} \textbf{8.27}$$

Hence the two simulations should differ in intensity by a factor

$$\frac{I_2}{I_1} = \frac{I_0 \exp(-\mu_{a2} \Sigma l_i)}{I_0 \exp(-\mu_{a1} \Sigma l_i)} = \exp((\mu_{a1} - \mu_{a2}) \Sigma l_i) \qquad \textbf{8.28}$$

To avoid having to re-run simulations every time only μ_a is altered, a C programme called *muascale* was written to change the apparent μ_a of the simulation. This read in the 8323320 byte binary file, and wrote it out to a new file, scaling the data to a new μ_a and changing the header to show the new μ_a.

   The idea of re-scaling the data for a change of μ_a can be extended to the inhomogeneous case if the pathlengths are recorded in each of the regions of differing μ_a's. However, time did not permit simulations with inhomogeneous media, so the programme *muascale* has currently only been tested for homogeneous media.

Figure 8.15 shows some data collected for μ_a=0.03 mm^{-1}, μ_s=10, μ_s'=1.0 mm^{-1} after post-processing with *muascale* to increase the μ_a by a factor of 10 to μ_a=0.3 mm^{-1}. Also plotted is some data collected at a μ_a of 0.3 mm^{-1}, but at the same scattering coefficient and *g*. The data are as expected very similar. Slight differences between 17 and 23 mm source-detector spacing, can be attributed to statistical variations. The data collected at a μ_a of 0.3 mm^{-1} shows nothing beyond 23 mm, whereas that rescaled from 0.03 to 0.3 mm^{-1} shows some data, although it is noisy. The data collected at a μ_a of 0.3 mm^{-1} is believed to have fallen below a cutoff used internally.

Figure 8.15 Data collected at μ_a=0.03 mm^{-1} but re-scaled to μ_a=0.3 mm^{-1} (solid line) and data collected at μ_a=0.3 mm^{-1} (dashed line).

8.9.3   Post-processing programme *addmc*.

   The main Monte Carlo programme *fibremc* could be run in parallel on many machines, producing only one 8 Mb data file containing the results from all the parallel simulations. Sometimes data was collected on several computers in different locations, including several PC's

and Sun workstations at the university, and one PC and one Sun workstation at home. A programme *addmc* was written to add the two sets of data from the two simulations into one file. The programme read in the two 8 Mb binary files, then wrote the data out as a third 8 Mb binary file. Checks were made to ensure the two input data files were of the same optical properties and geometry, as otherwise adding two such data sets would be a worthless exercise. The programme needs approximately 24 Mb of memory, so would only run on a reasonably well specified PC, when first written, but now most PCs have 64 Mb of memory, so this is not a problem. The programme could be re-written to use memory more efficiently, but there is little point due to the improved specification of modern computers.

The graphs at the top and centre of figure 8.16 show the output from two runs of the computer programme. These were independent runs, performed at separate locations. The lower graph in figure 8.16 shows the results after post-processing with

Figure 8.16 Diagram showing how *addmc* could be used to reduce the noise by adding two data files. μ_a=0.03 mm^{-1}, μ_s'=1.0 mm^{-1}.

addmc, where the statistical noise on the data is reduced as expected.

8.9.4 Post-processing programme *avgmc*.

A typical graph of intensity versus time is shown in the upper graph of figure 8.17. This data was collected by simulating a 1.1 mm diameter detector fibre, 10 mm from the source fibre, with tissue properties of μ_a=0.03 mm^{-1}, μ_s=10 mm^{-1}, μ_s'=1 mm^{-1}. To avoid this discontinuous result, it was usually necessary to avoid plotting data points for which no data is available. To aid subsequent plotting, a programme *avgmc* was used to average the data from one of the files

produced by *anal6*. This worked by averaging the data from several time steps, then plotting this. The 1 mm pathlength steps were increased to 3 mm, corresponding to 14 ps in tissue, or 13.33 ps in the phantom. Eventually a different approach, using a program *rmzero* was found to be more favourable than *avgmc* and is described in the next section.

8.9.5   Post-processing programme *rmzero*.

The program *rmzero* post-processed the data file, removing all entries for which no data was collected - no averaging was performed. An example of this is shown in the lower half of figure 8.17.

8.9.6   Post-processing   programme *plotmany*

The text data files produced by programme *anal6* were for a specific source-detector spacing. Sometimes it was desirable to plot a graph with several different source-detector spacings. While it was possible to create files using *anal6* for all the required source-detector spacing, then read each into a graph plotting package, this was a very tedious process. Hence a unix shell script *plotmany* was used to automate the process. This used *anal6* to put the simulated TPSFs at 1 mm, 2 mm ... 20 mm, 25 mm, 30 mm ... 50 mm into 26 different files, named 1.txt, 2.txt, 3.txt ... 50.txt, then used the standard unix utilities of *awk* and *paste* to put the required data from the 26 files into one file, which was easier to read into the graph plotting programme.

Figure 8.17 Graphs of the data produced by *anal6* (upper graph), and latter post-processed with programme *rmzero*. μ_a=0.03 μ_s'=1.0 mm^{-1}.

The Sun Microsystems version of *paste* (which copies columns from multiple files into one larger file), as supplied with Solaris 2.6, had a limit of 12 input files. Hence it would not cope with the 26 different input files

needed. No such problem occurs with the linux operating system on a PC, as the GNU version of *paste* used by linux, has no such limit on file numbers. The GNU version of *paste* was then compiled on the Sun, and used in preference to the Sun version.

While the main Monte Carlo programme *fibremc* was tested under DOS 5.0 on a PC, linux 2.6.2 on a PC, and Solaris 2.6 on a Sun workstation, the programme *plotmany* has not been implemented on a PC under DOS. *Plotmany* needs several unix utilities and shell features, that make porting it to DOS a long process, so for now this part of the analysis routine has to be performed on a Sun workstation or a PC running linux.

Chapter 9.

Results - Experimental, Theoretical and their Comparison.

This chapter contains the experimental results measured with the cross-correlator on both the solid phantom described on page 144 and the liquid phantom described in chapter 7. The expected TPSFs were calculated by diffusion theory for both phantoms and by the use of the Monte Carlo model described in chapter 8 for only the liquid phantom. Five different fitting routines are described that were used to attempt to estimate the μ_s' of the liquid phantom by comparing experimental data with results from both the Monte Carlo and diffusion theory models. The relative merits of these different fitting algorithms are discussed.

9.1   Monte Carlo results.

Monte Carlo data was collected for 80 source-detector spacings of 0 to 79 mm, for 7 different values of μ_s' (0.2, 0.3, 0.5, 1.0, 1.5, 2.0 and 3.0 mm^{-1}). Table 8.1 on page 205 shows the number of photons launched and detected for each value of μ_s' simulated. The simulations can be re-scaled for an infinite range of different μ_a's. Although it is possible to generate huge amounts of data quite easily with Monte Carlo programmes, there is little point in including such a huge amount in this thesis. Hence only a few representative examples are plotted in the following pages. Data is plotted for the following range of optical properties:

1) Low (0.0018 mm^{-1}), medium (0.0192 mm^{-1}) and high (0.0324 mm^{-1}) values of μ_a.

2) Low (0.3 mm^{-1}), medium (1.0 mm^{-1}) and high (3.0 mm^{-1}) values of μ_s'. (In each case g was 0.9, so the $\mu_s'=\mu_s/10$).

Figures 9.1 to 9.9 show TPSFs for these 9 combinations of μ_a and μ_s', at source-detector spacings of 5, 10, 15 and 20 mm. (The source and detecting fibre are co-linear with each other and have been normalised to the peak value for ease of visualisation, since the absolute magnitudes of the signals vary by many orders of magnitude, depending on the tissue properties).
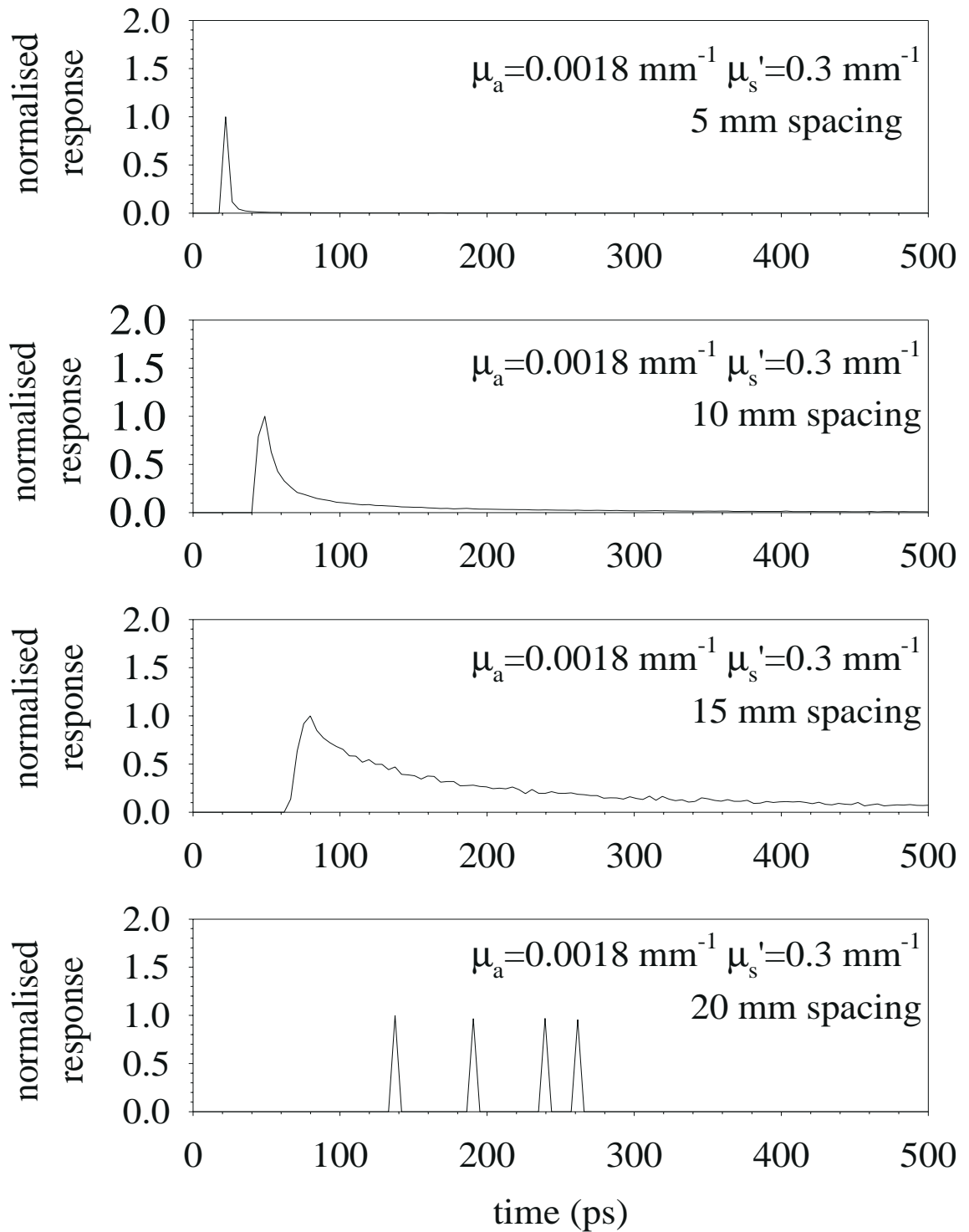
Figure 9.1 TPSF's calculated by the Monte Carlo programme for low μ_a (0.0018 mm^{-1}) and low μ_s' (0.3 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
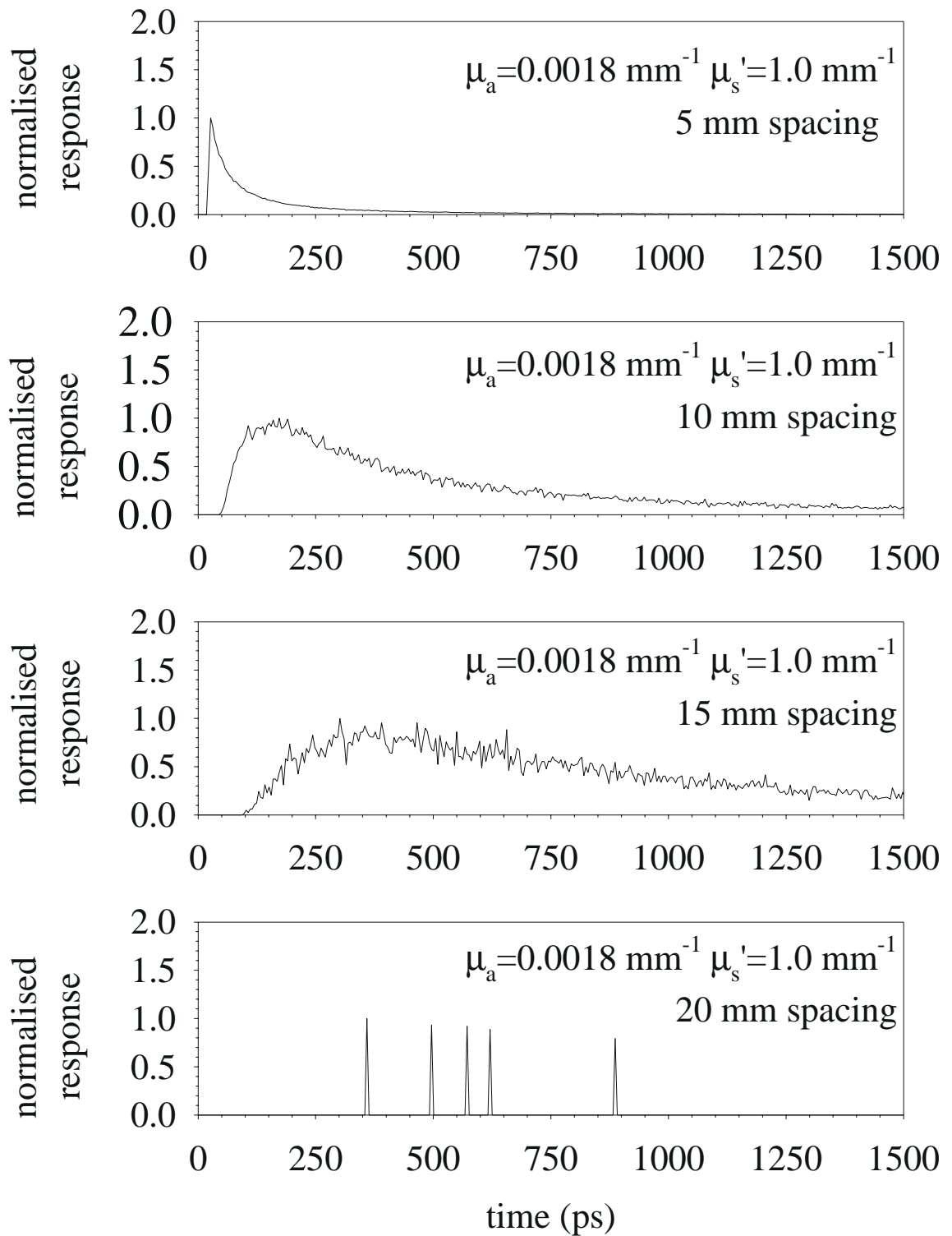
Figure 9.2 TPSF's calculated by the Monte Carlo programme for low μ_a (0.0018 mm^{-1}) and medium μ_s' (1.0 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
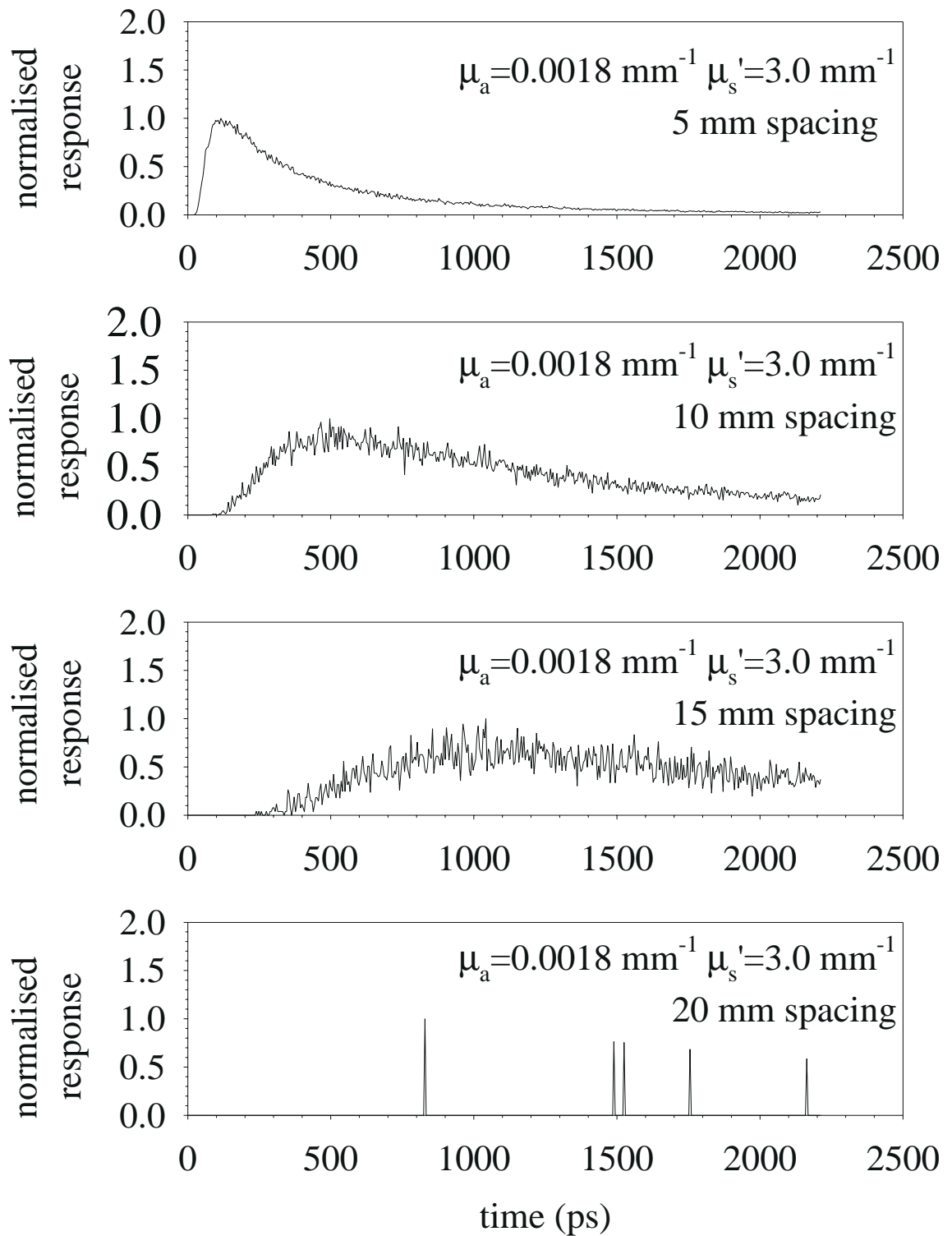
Figure 9.3 TPSF's calculated by the Monte Carlo programme for low μ_a (0.0018 mm $^{-1}$) and high μ_s' (3.0 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
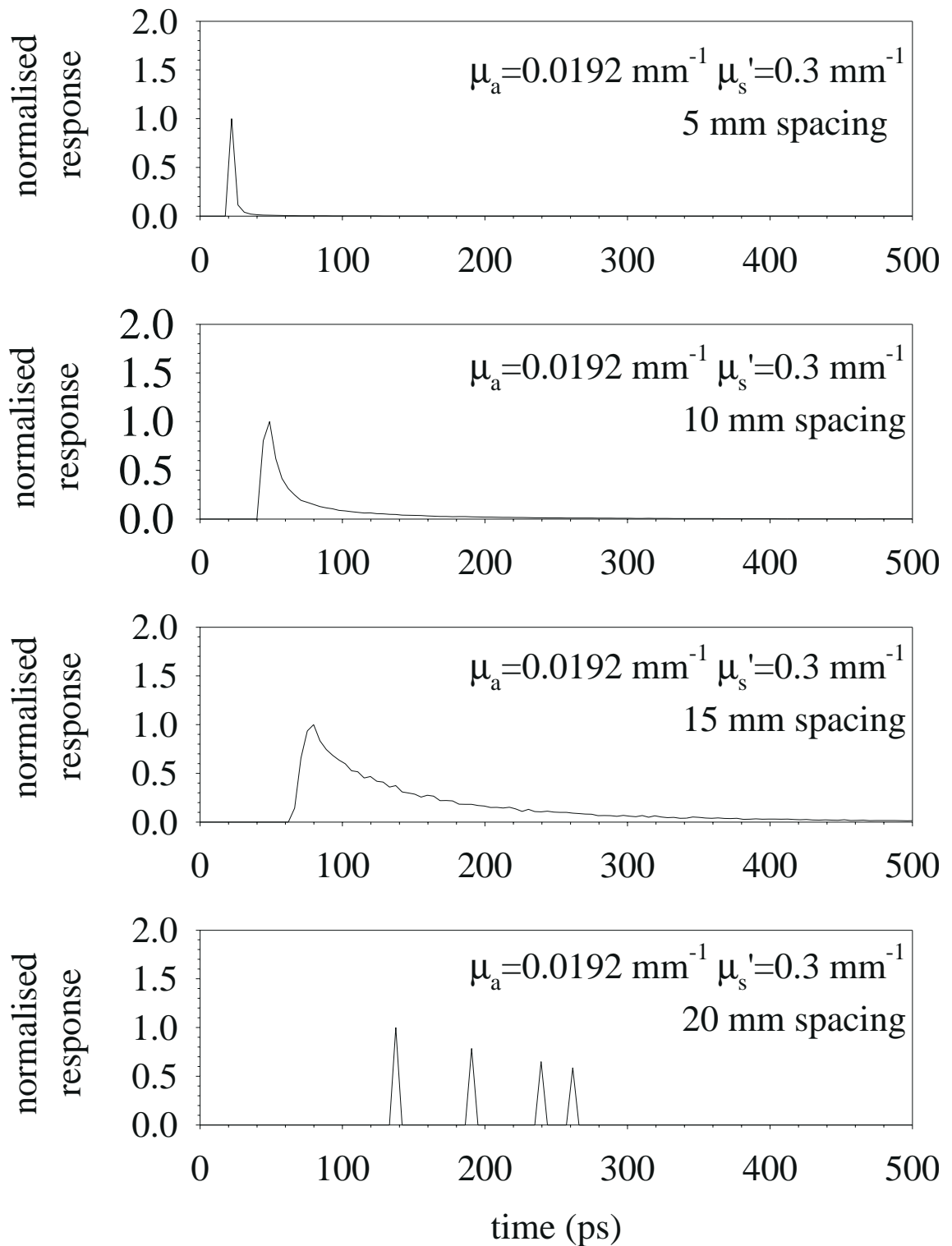
Figure 9.4 TPSF's calculated by the Monte Carlo programme for medium μ_a (0.0192 mm^{-1}) and low μ_s' (0.3 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
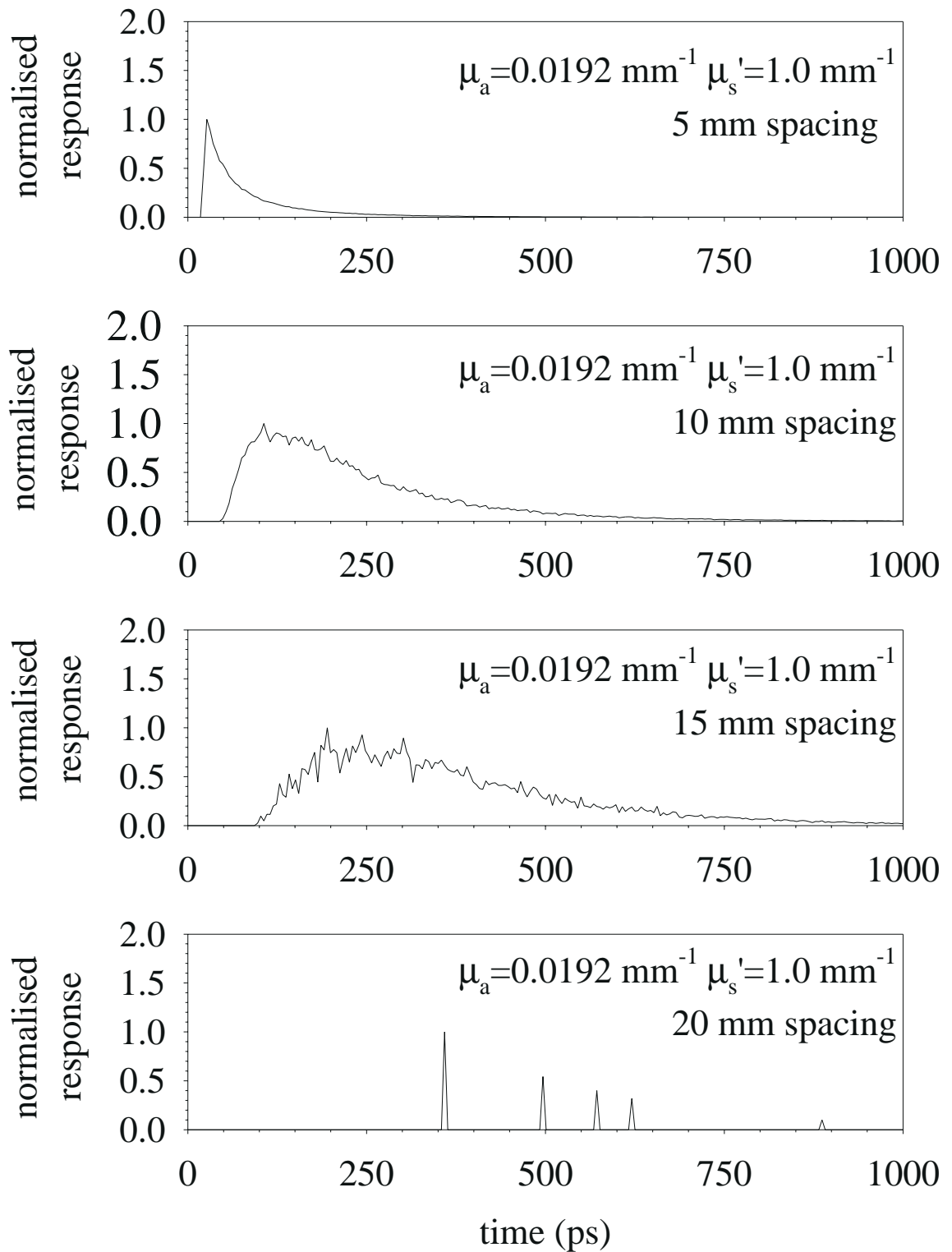
Figure 9.5 TPSF calculated by the Monte Carlo programme for medium μ_a (0.0192 mm^{-1}) and medium μ_s' (1.0 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
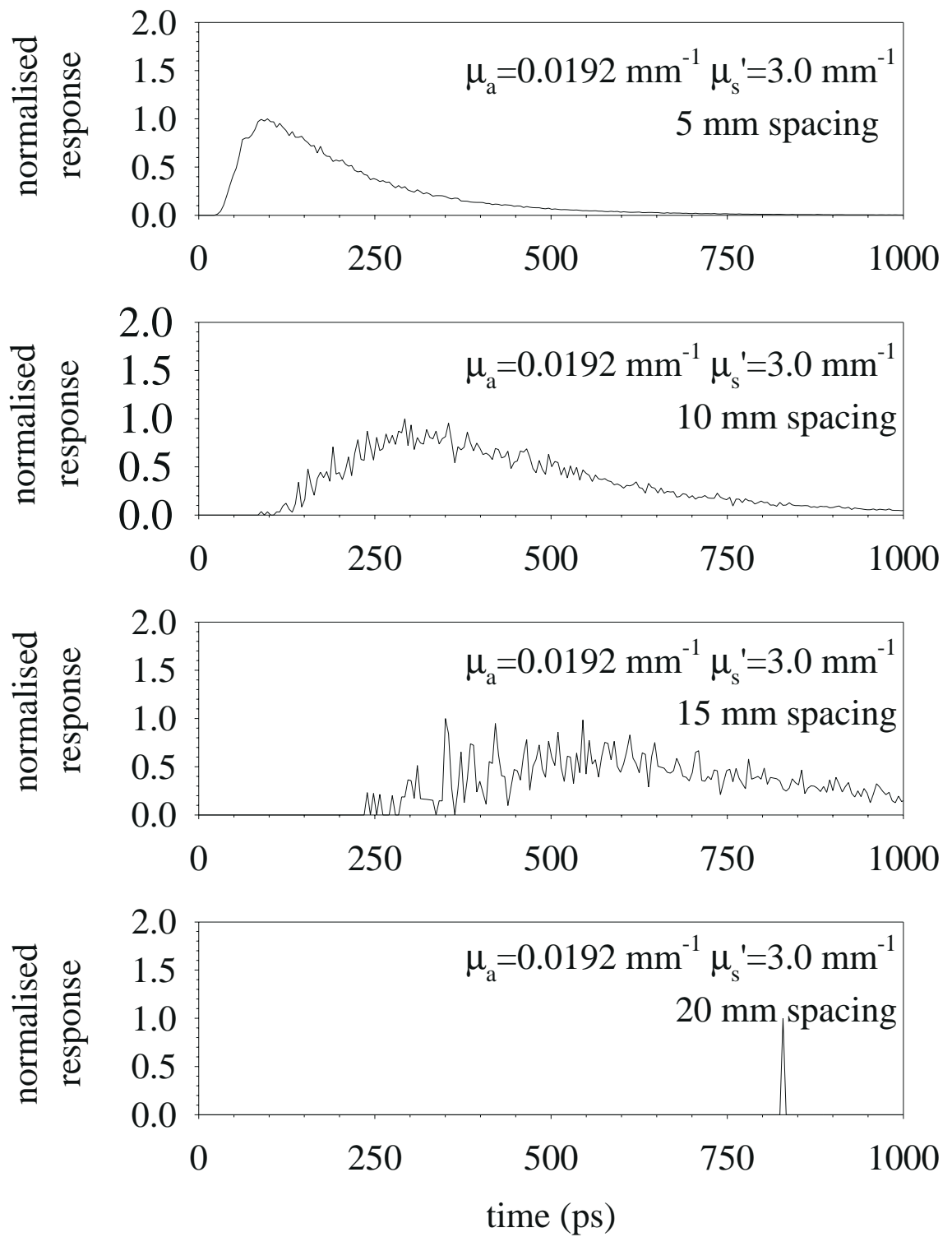
Figure 9.6 TPSF's calculated by the Monte Carlo programme for medium μ_a (0.0192 mm^{-1}) and high μ_s' (3.0 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
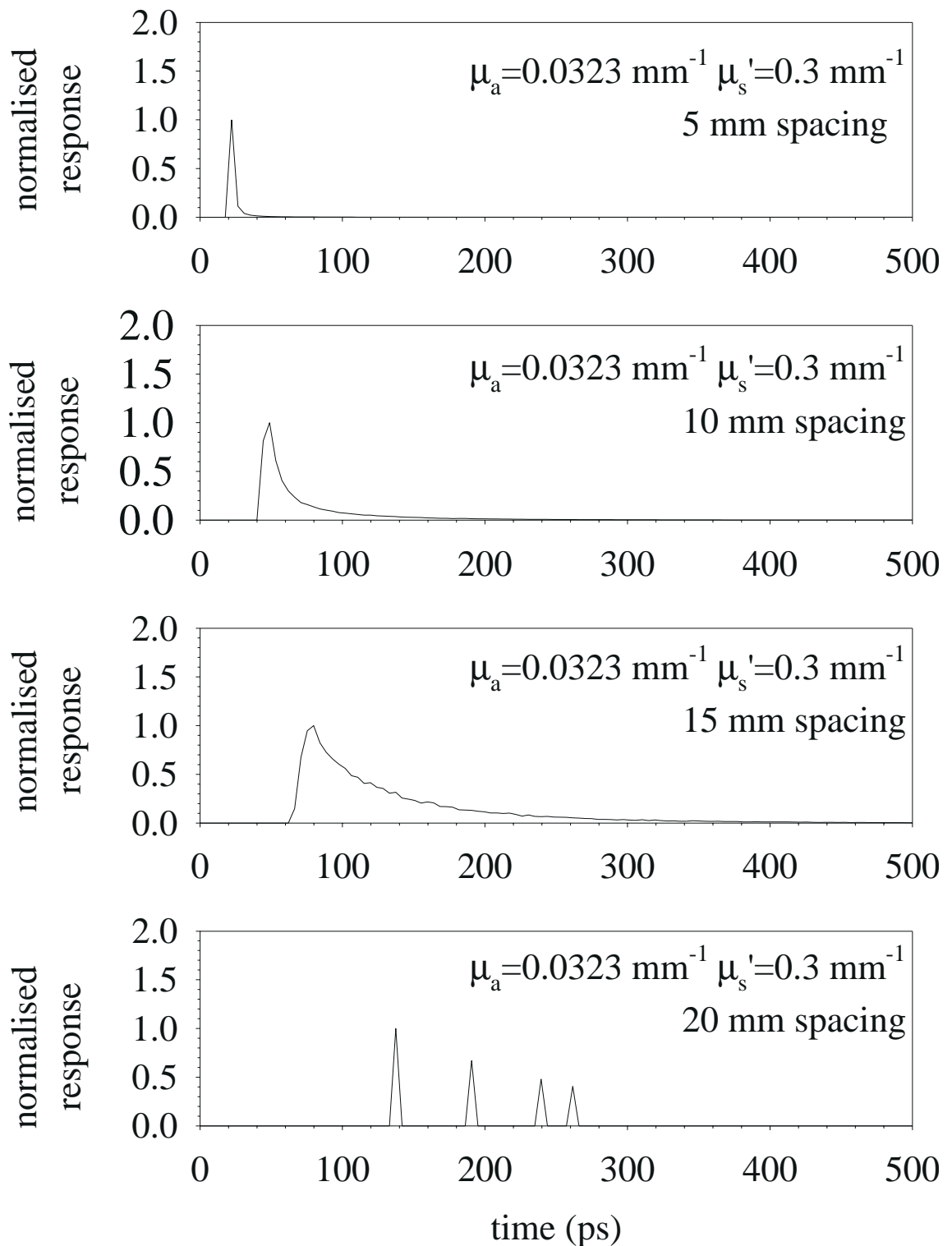
Figure 9.7 TPSF's calculated by the Monte Carlo programme for high μ_a (0.0323 mm^{-1}) and low μ_s' (0.3 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
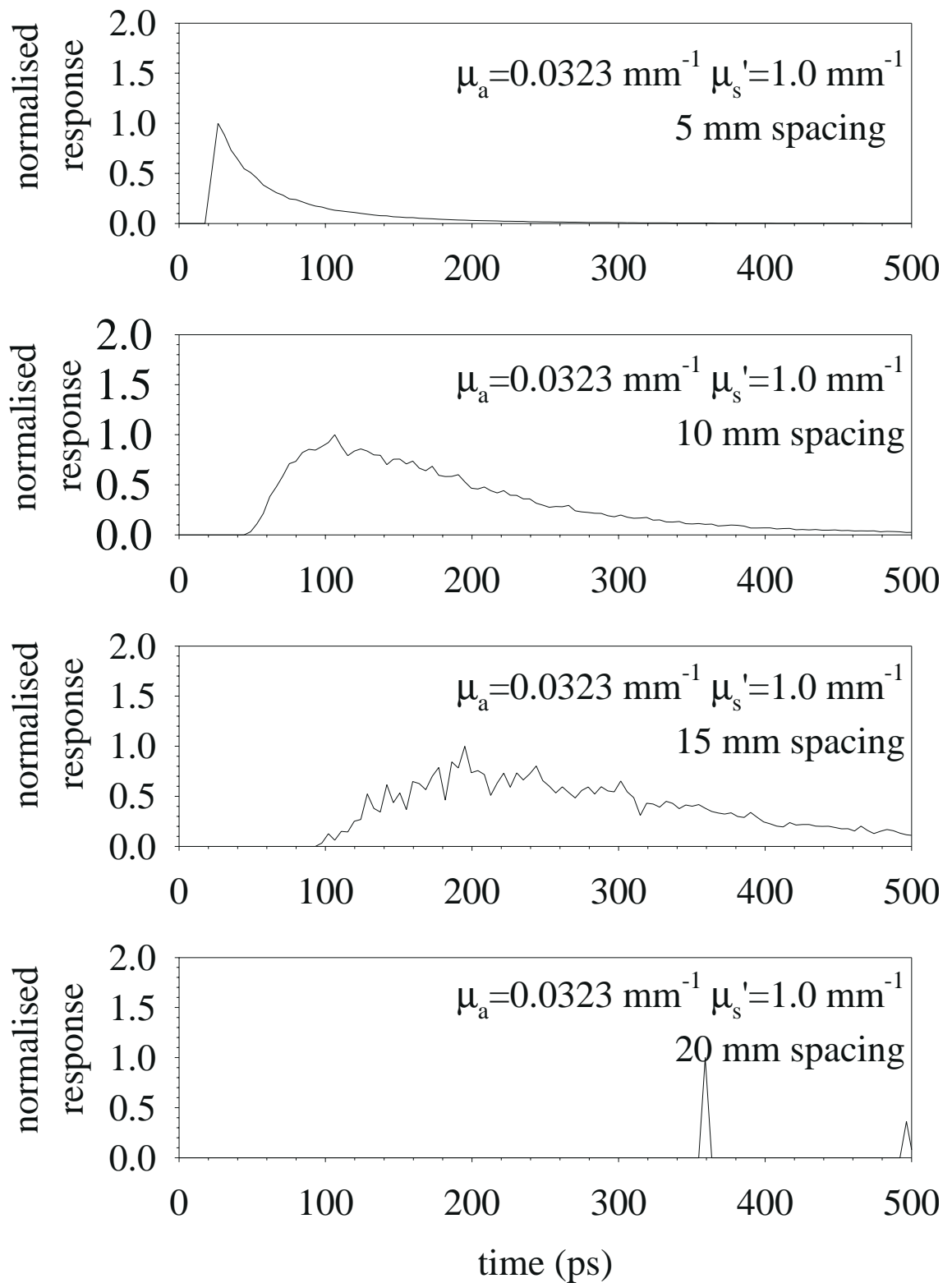
Figure 9.8 TPSF's calculated by the Monte Carlo programme for high μ_a (0.0323 mm^{-1}) and medium μ_s' (1.0 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.
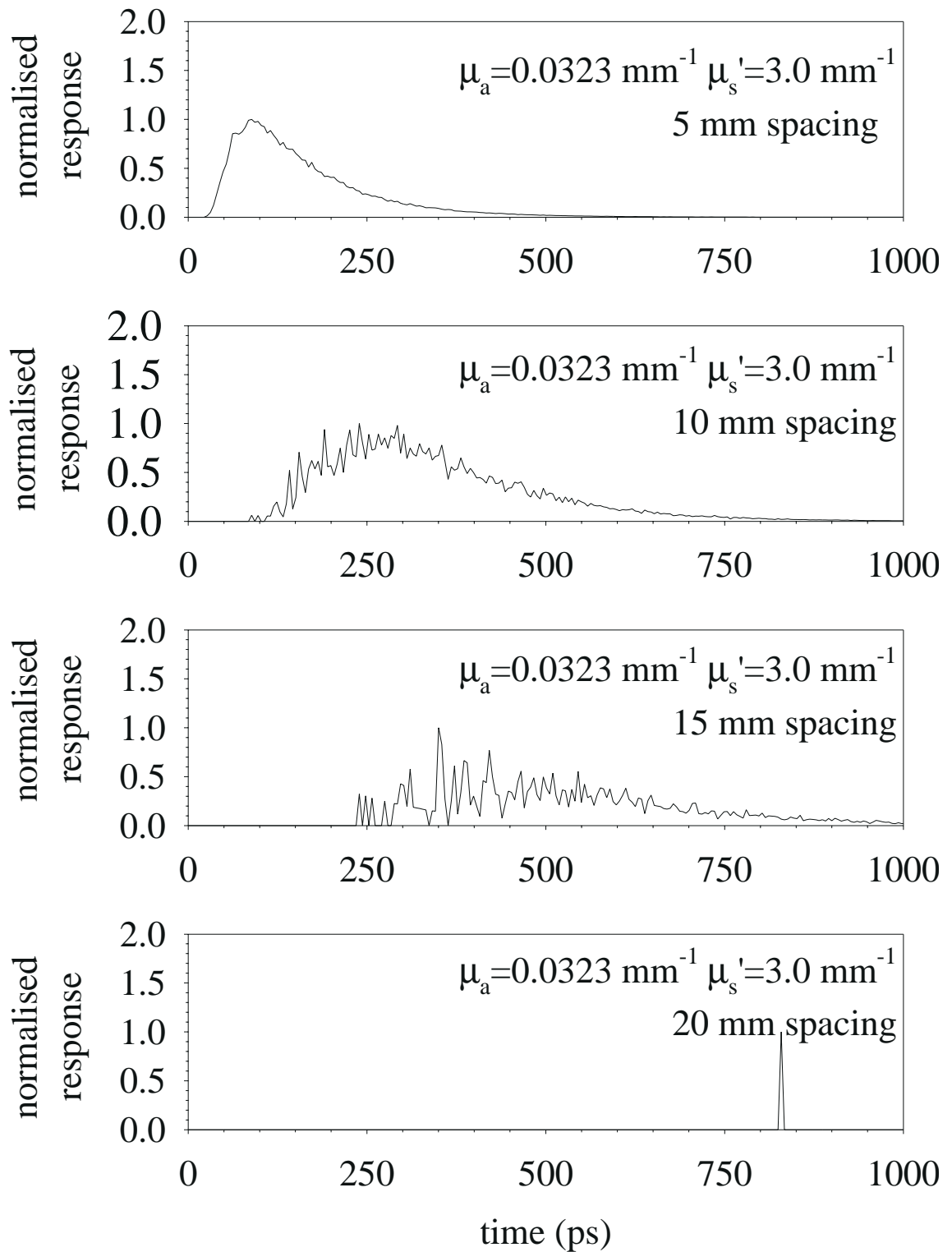
Figure 9.9 TPSF's calculated by Monte Carlo programme for high μ_a (0.0323 mm^{-1}) and high μ' (3.0 mm^{-1}) at source-detector spacings of 5, 10, 15 and 20 mm.

These Monte Carlo simulations show several properties of the TSPF.

1) The TPSF is always positively skewed - ie the rise time is faster than the fall time.

2) The duration of the TSPF increases for increasing μ_s'.

3) The duration of the TSPF increases for increasing distance between the source and the detector.

4) The duration of the TPSF decreases for increasing μ_a.

5) The time of the peak of the TPSF decreases as μ_a is increased.

6) The time of the peak of the TPSF increases as μ_s' is increased.

7) There is a minimum time before which no output is generated. This time can be calculated analytically as

$$t_{min} = \frac{n \times d}{c} \qquad\qquad \textbf{9.1}$$

where n is the refractive index, c the velocity of light in vacuum, and d the source to detector spacing. For large source to detector spacings and μ_s', there is no significant output until times much greater than this, due to photon diffusion.

8) The noise on the data is increased for increasing source to detector spacing as there is less detected light. Since this fact was known from the beginning, the fastest computers were used for the simulations at high μ_s' and the slower ones used for low μ_s'. The number of photons simulated is shown in table 8.1 on page 205.

9) The data at source to detector spacings of 20 mm is unusable. The noise naturally rises with increasing source to detector spacing, although at 20 mm the noise seems to have been larger than expected. This is probably due to an inappropriate (too high) choice of the intensity cutoff parameter, which is described on page 191. This is not however a major problem, since as explained later, diffusion theory can be used to accurately calculate the TPSF at this spacing.

9.2   Experimental results with the cross-correlator on phantoms.

The cross-correlator described in chapter 6 was used to make measurements on two types of optical phantom - a solid phantom and a liquid phantom. The solid phantom was used to carry out tests of the instrument performance, since it is highly stable and reproducible. The liquid phantom was used to test methods of estimating μ_s' from measurements performed with the instrument since it is relatively easy to change the optical properties of this phantom.

9.2.1   TPSF obtained on Solid Phantom.

Figure 9.10 shows one of the results on a solid phantom produced during development

of the instrument. For comparison, data was also measured on a Hamamatsu streak camera. The absolute time calibration of the cross-correlator data was unknown at the time this measurement was taken, so the time of the peak outputs of the cross-correlator and streak camera were set equal. The two sets of data are normalised to have the same peak amplitude. This data was presented at the Optical Society of America Conference at Orlando, Florida in 1996[151] and was later published[152]. Since then, the time axis has been calibrated properly, so subsequent data is shown at absolute times. Note that this early data was collected with the lock-in amplifier operating in voltage mode. Later measurements were made in current mode, as it was realised this was significantly better in terms of S/N ratio (see section 6.11).
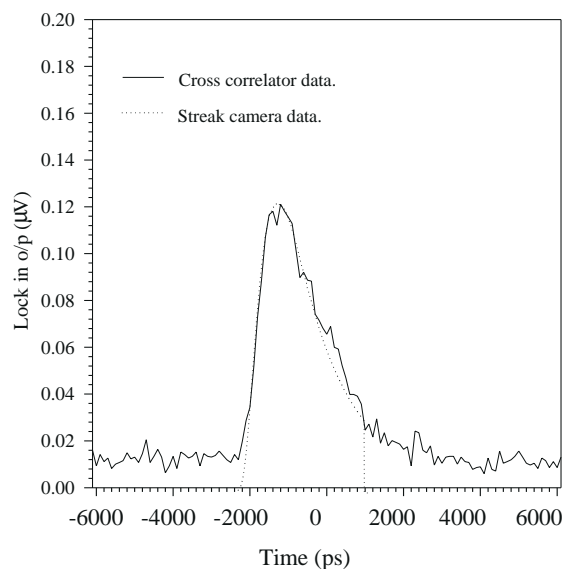
Figure 9.10 A comparison of the TPSF measured on both the cross-correlator described here and a Hamamatsu Streak camera.

9.2.2 TPSFs obtained on Liquid Phantoms.

After construction of the phantom described in chapter 7, and shown in figure 7.1, the cross-correlator was used to make measurements of TPSFs with various source-detector spacings on phantoms with a range of optical properties. The laser was tuned to 800 nm. The 10% Intralipid^{TM} was diluted to either 3%, 10% or 30%, to obtain reduced scattering coefficients of 0.3 mm^{-1}, 1 mm^{-1} or 3 mm^{-1} respectively at 800 nm, as discussed in chapter 7. Dye was added to these 3 scattering solutions enabling measurements to be made at μ_a's of 0.0018 (Intralipid^{TM} diluted with pure water), 0.00616, 0.0105, 0.0149, 0.0192, 0.0236, 0.0279 and 0.0323 mm^{-1}. In all cases except those where μ_a=0.0018 mm^{-1}, the dye with the absorption spectrum shown in figure 7.4 was added to increase absorption. The dye, with a μ_a of 13.08 mm^{-1} at 800 nm, was added to the Intralipid^{TM} in multiples of 500 μl, so the absorption coefficients were as stated in table 7.1. Measurements were usually made at source-detector spacing of 5, 10, 15 and 20 mm, although on one occasion measurements were taken at 30 mm. Sometimes when both μ_a and μ_s' were towards the higher ends of their values it was not possible to obtain data at a spacing of

20 mm in a reasonable time.

9.3 Estimating μ_s' from the experimental data.

The data from the cross-correlator is saved in an ASCII text file, an example of which was shown in section 6.16.1.3. In order that μ_s' could be estimated from the experimental measurements, several different methods were tried. Sections 9.3.1 to 9.3.5 show the results of estimating μ_s' using these five different methods.

9.3.1 Method 1 - fitting to the temporal position of the TPSF peak (t_{peak}).

Patterson et al[39] have shown that for a semi-infinite half space, the absorption coefficient μ_a can be found from the log slope of the TPSF and the transport scattering coefficient μ_s' can be found from the time of the peak of the TPSF, t_{peak} (see figure 1.8). The theory of this is discussed, starting on page 50.

If instead of starting from the solution of the diffusion equation in a semi-infinite slab, the solution of the diffusion equation for the infinite homogeneous medium (equation 1.34) is used very similar mathematics will give μ_a and μ_s' for the infinite homogeneous medium. The natural logarithm of equation 1.34 is

$$\log_e[\psi(r,t)] = \log_e[\frac{d}{2}(4\pi Dc/n)^{-3/2}] + \log_e[t^{-5/2}] - \frac{nr^2}{4Dct} - \mu_a tc/n \qquad \textbf{9.2}$$

If equation 9.2 is differentiated with respect to time, the following is obtained

$$\frac{d[\log_e[\psi(r,t)]]}{dt} = -\frac{5}{2t} + \frac{nr^2}{4Dct^2} - \mu_a c/n \qquad \textbf{9.3}$$

At large t, this allows μ_a to be found, although the result is no different from that of the semi-infinite slab (equation 1.42).

$$\mu_a = \frac{-n\left[\lim_{t\to\infty} \frac{d[\log_e[\psi(r,t)]]}{dt}\right]}{c} \qquad \textbf{9.4}$$

At the time of the peak of the TPSF t_{peak}, the rate of change of $\psi(r,t)$ is obviously zero, which leads to equation 9.5

$$0 = \frac{-5}{2t_{peak}} + \frac{nr^2}{4Dct_{peak}^2} - \frac{\mu_a c}{n}$$ **9.5**

Substituting for D from equation 1.32 and rearranging gives equation 9.6

$$\mu_s' = \frac{4ct_{peak}^2}{3nr^2}\left(\frac{5}{2t_{peak}} + \frac{\mu_a c}{n}\right) - \mu_a$$ **9.6**

Figure 9.11 shows the experimentally measured peak position times of the TPSFs measured on the cross-correlator at a μ_s' of 0.3 mm^{-1}, 1.0 mm^{-1} and 3.0 mm^{-1}, for various source-detector spacings from 5 to 30 mm. To gather this data, the time of the peak of the TPSF was noted and this subtracted from the peak time of the instrument impulse response, as described in section 6.12. One of the conclusions from the Monte Carlo results (see page 221) was that the peak time of the TPSF should decrease with increasing μ_a. This is generally the case (apart from the odd example), although there is a definite increase in this time for the highest 3 values of μ_a at a μ_s' of 0.3 mm^{-1}. This is believed to be due to the Intralipid^{TM} phantom suffering bacterial contamination. The mixture which was mixed on 24/12/1998 and kept in a refrigerator at 4°C, but it was noted 6 days later (30/12/1998) that it had developed a visibly different appearance and was disposed of. A second phantom mixture was then mixed up at μ_a=0.02796 mm^{-1} and μ_s'=0.3 mm^{-1}. It is probable that at some time during the 6 days the mixture had deteriorated due to bacterial growth. The last measurements made with the original mixture were on 29/12/1998 when it was 5 days old.

Figure 9.12 shows the μ_s' estimated from the data in figure 9.11 using equation 9.6 at source-detector spacings of 5, 10, 15, 20 and 30 mm. For the upper graph the true μ_s' is 0.3 mm^{-1}, for the centre graph it is 1.0 mm^{-1} and for the lower graph 3.0 mm^{-1}. The mean estimate of μ_s' when the true value was 0.3 mm^{-1} is 0.58 mm^{-1}, indicating a 93% over-estimation. At a true μ_s' of 1.0 mm^{-1} the mean estimate is 2.01 mm^{-1} indicating a 101% over-estimation. At a true μ_s' of 3.0 mm^{-1} the mean estimate is 4.25 mm^{-1}, indicating an over-estimation of 41.6% on average. For all three graphs, the data points most in error are not surprisingly those measured at a source-detector spacing of 5 mm. The data at a μ_s' of 3.0 mm^{-1} is less in error than the other two sets of data, since the system response is less significant at the longer TPSFs generated when μ_s' was 3.0 mm^{-1}.

Virtually all the predicted values of μ_s' were larger than the true transport scattering coefficients. Analysis of the 3 main possible causes of error show 2 out of 3 causes would result
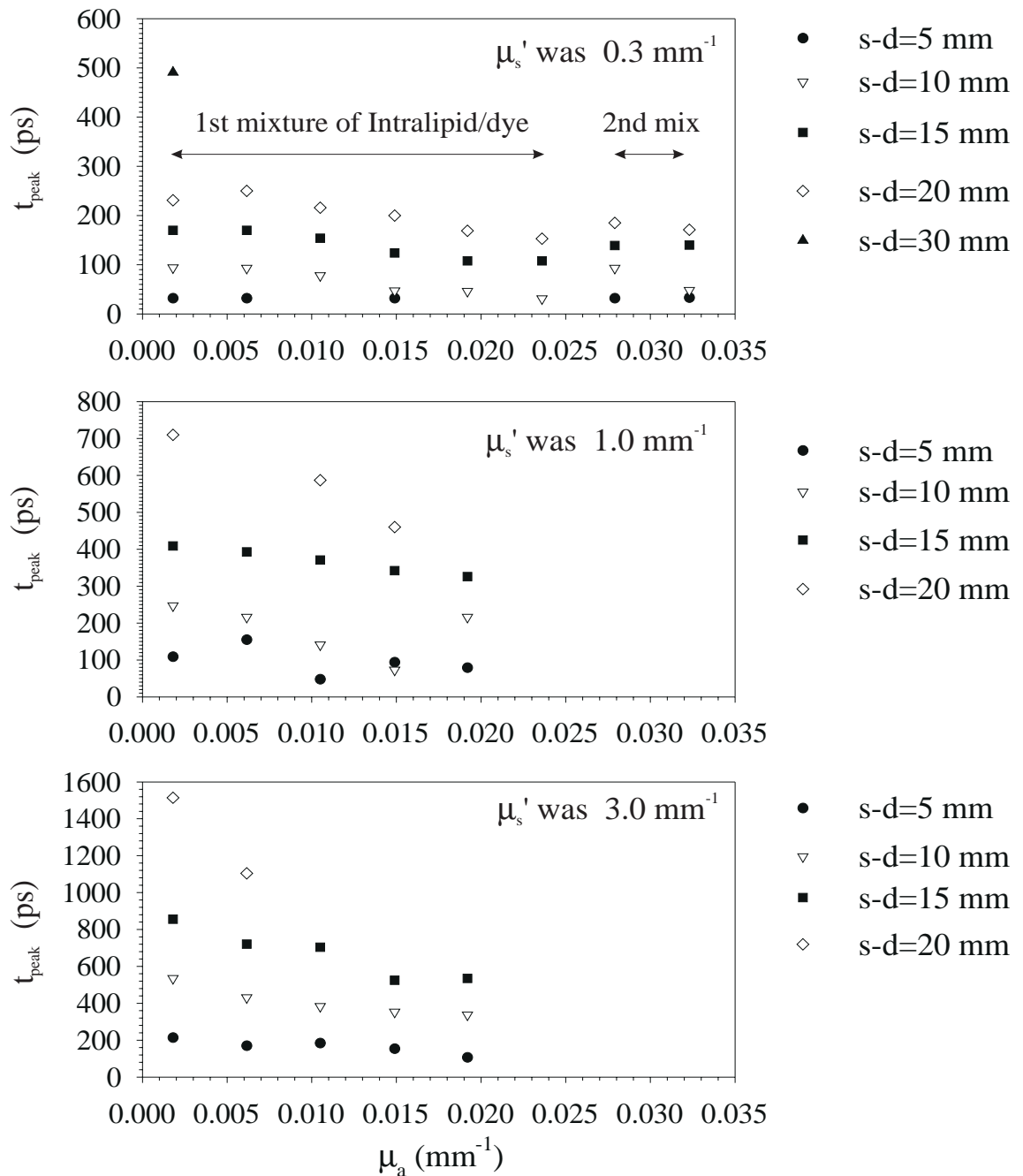
Figure 9.11 Graphs showing the time of the peak of the TPSF's at $\mu_s' = 0.3$, 1.0 and 3.0 mm^{-1} for a range of absorption coefficients and source-detector spacings.

in an over-estimation of μ_s', while the third source of error is random, and so would expect to lead to an equal chance of underestimation or overestimation.

1) The source-detector distance recorded in the experiments was the distance the vernier scale was moved from when the source and detector were in the same plane. Sometimes the two fibres were not exactly in line at a spacing of zero. This was almost impossible to achieve, as no
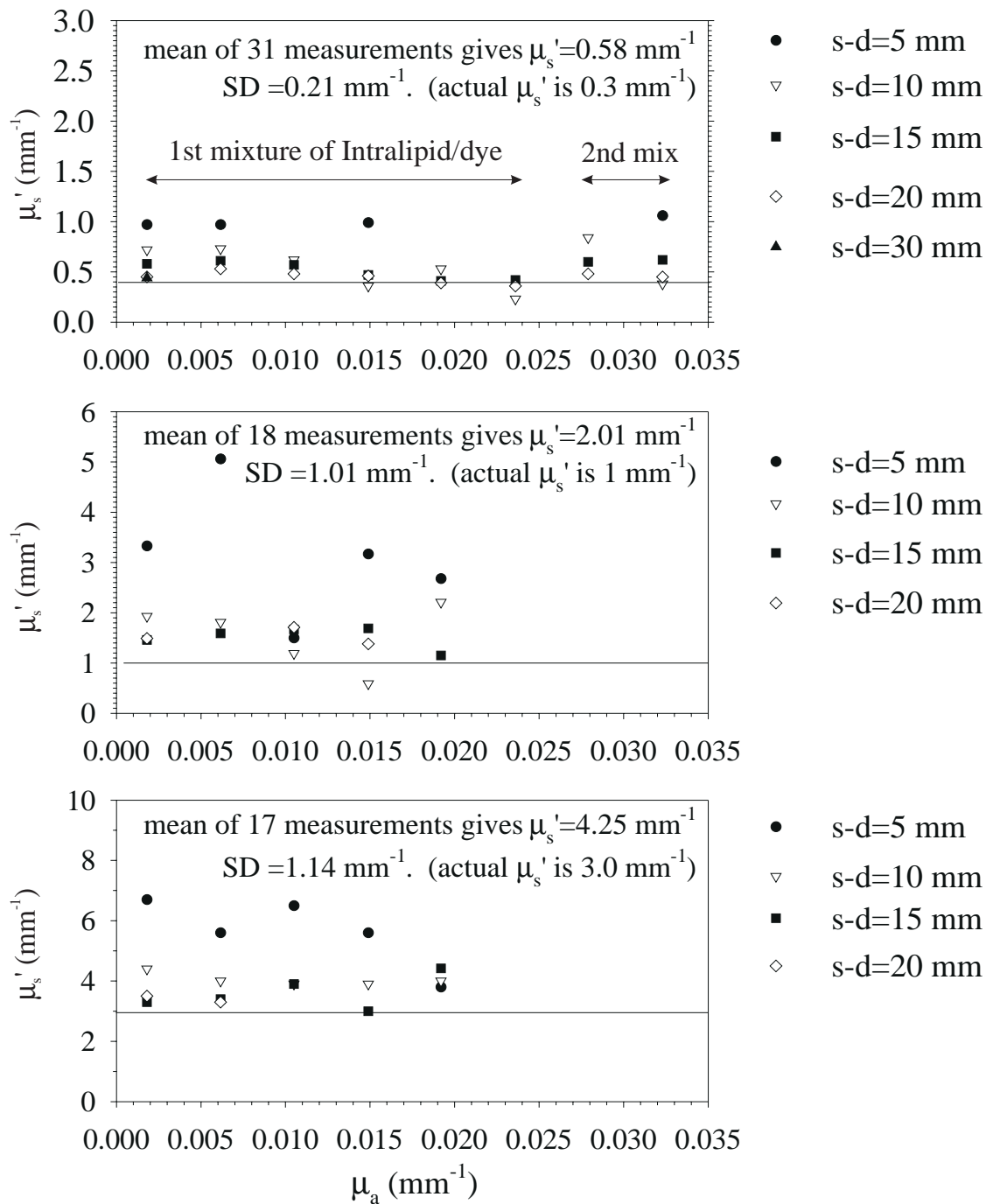
Figure 9.12 Graphs showing the predicated μ_s' using method 1 for known μ_s' of 0.3 mm^{-1} (upper graph), 1.0 mm^{-1} (centre graph) and 3.0 mm^{-1} (lower graph).

amount of bending of the fibres could guarantee they would be exactly in line. Therefore, if the vernier scale was moved 5, 10, 15 or 20 mm, the true source-detector spacing would always be greater than this. It is estimated that the worst-case error would be a lateral displacement of the fibres of 2.5 mm in x and 2.5 mm in the y planes. At the shortest z distance used (5 mm), the overall source-detector spacing could be as much as:

$$r = \sqrt{2.5^2 + 2.5^2 + 5^2} = 6.1 \ mm \qquad \textbf{9.7}$$

This could at worst-case mean the source-detector spacing having an error of 22% at 5 mm, although this would be a worst-case and the error would be insignificant with the vernier set at more than 5 mm. This will result in the TPSF being shifted towards somewhat later times than calculated and result in over-estimating μ_s'.

2) The TPSFs predicted by both the Monte Carlo and diffusion models are all positively skewed - their rise time is shorter than the fall time. The measured data is the convolution of the true TPSF and the system response, which has zero skew (the rise time is approximately the same as the fall time). This will result in the measured peak of the TSPF being shifted to the right (later time) of the true peak, which will result in over-estimating μ_s'.

3) Random errors of timing will be expected to cause random fluctuations in the estimates of μ_s'.

A revised method was therefore tested to estimate μ_s', taking into account the finite temporal resolution of the system, which is expected to be the most significant cause of error.

9.3.2   Method 2 - fitting to the position of t_{peak} after convolution.

The system has an approximately Gaussian shaped temporal response of 380 ps FWHM, which is not insignificant, except at large source-detector spacings and high scattering coefficients. The true TPSFs (TPSF(t)) will be smeared by the system response R(t) to give an experimentally measured set of data $TPSF_{exp}$(t). A fitting algorithm was needed that took this into account. Perhaps the most obvious way would be to deconvolve the experimental data $TPSF_{exp}$(t) with the system response R(t), to get the true TSPF(t). However, deconvolution only really works well if the system response is much faster than the measured data and the measured data is of low noise. Neither criteria is satisfied here, so it was not surprising that attempts at deconvolution were unsuccessful.

Instead, TPSFs were generated by the diffusion equation in an infinite volume (equation 1.34) for the known μ_a, n and source-detector spacing, but at various transport scattering coefficients from 0.025 mm^{-1} to 10 mm^{-1} in steps of 0.025 mm^{-1}. These TPSFs (called $TPSF_{diff}$(t)) were then convolved with the system response R(t) (assumed to be Gaussian of 380 ps FWHM) to give a signal $TPSF_{diff}$(t)⊗R(t). The time of the peak of $TPSF_{diff}$(t)⊗R(t) was then compared to that of the experimentally measured TPSFs. This effect is illustrated in figure 9.13,
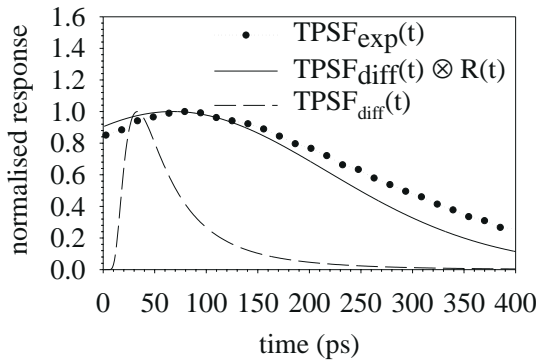
Figure 9.13 Comparison of simulated data generated with the diffusion equation both before and after convolution with the system response. Also shown is experimental data for the same conditions.

which shows $TPSF_{diff}(t)$ for μ_a=0.0192 mm^{-1}, μ_s' =1 mm^{-1}, n=1.33 and a source-detector spacing of 5 mm both before (dashed line) and after convolution (solid line) with the system response. Also shown is the experimentally measured data at the same conditions. The experimentally measured data had peak at a time of 79 ps - very slightly after $TPSF_{diff}(t) \otimes R(t)$, but well after $TPSF_{diff}(t)$.

Figure 9.14 shows the time of the peak of $TPSF_{diff}(t) \otimes R(t)$ (dashed line) as a function of μ_s' over the range 0.025 mm^{-1} to 3 mm^{-1}. Also shown (solid line) is the difference between this time and the peak time measured experimentally. At a μ_s' of 1.25 mm^{-1} the signal $TPSF_{diff}(t) \otimes R(t)$ had a peak at 79 ps - the same as the experimental data, so 1.25 mm^{-1} was the estimated μ_s' for this particular piece of experimental data.
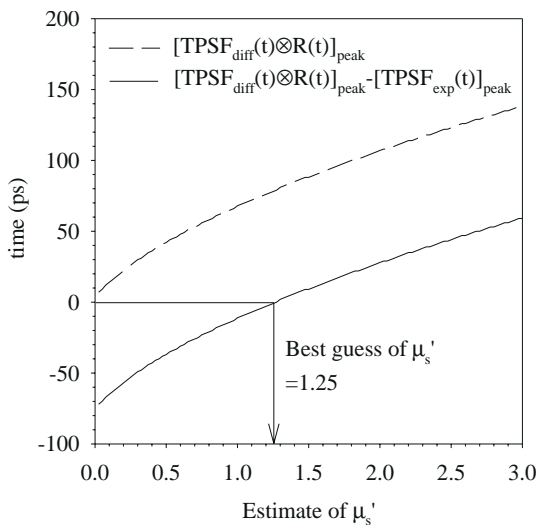
Figure 9.14 Graphs showing the temporal position of the TPSF peak as a function of μ_s' (diffusion theory), and the difference between this and the experimentally measured peak time (79 ps).

Figure 9.15 shows the results of the revised analysis, taking into account the system response and using the data from figure 9.11, which is the same data used to construct figure 9.12. The mean estimate of μ_s' when the true value was 0.3 mm^{-1} is 0.28 mm^{-1}, indicating a 6.7% under-estimation. At a true μ_s' of 1.0 mm^{-1} the mean estimate is 1.3 mm^{-1} indicating a 30% over-estimation. At a true μ_s' of 3.0 mm^{-1} the mean estimate is 3.02 mm^{-1}, indicating an over-estimation of 0.7% on average. Note that there is a very significant improvement in the accuracy to which μ_s' is estimated over that in figure 9.12. Hence it is very beneficial to take into account the system response.

The data in the upper graph of figure 9.15, which shows the estimates of μ_s' when the actual was 0.3 mm^{-1}, shows a decrease of apparent μ_s' as μ_a increases to 0.023 mm^{-1}, then makes a sudden jump upwards. This is almost certainly a result of bacterial contamination causing a
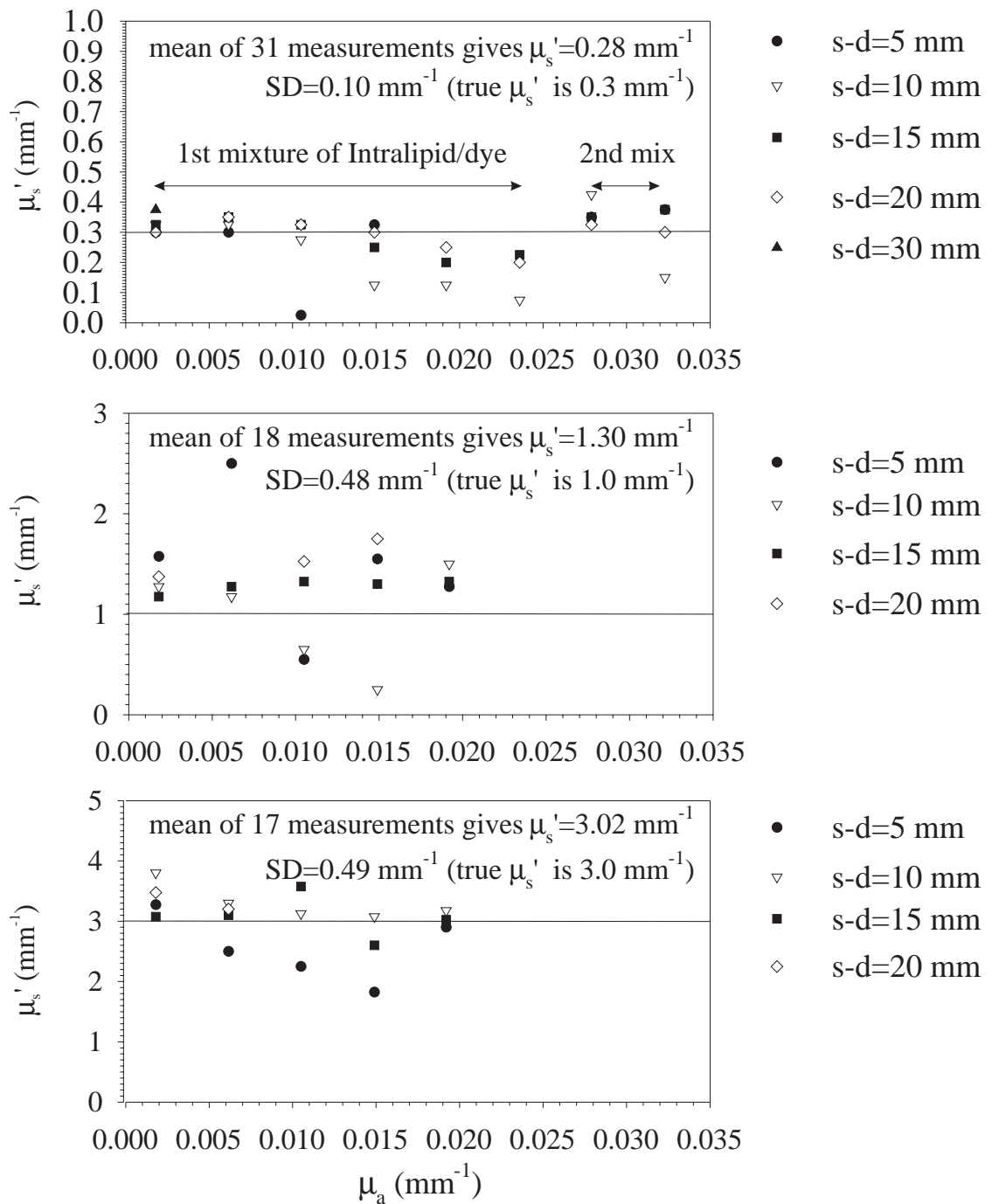
Figure 9.15 Graphs showing the estimated μ_s' using method 2 for phantoms with 3 known values of μ_s'. R(t) has been taken into account which improves accuracy over method 1.

change in the optical properties. It is difficult to estimate exactly what happens in this case. However, it does show that the Intralipid dye mixture can not be kept for 5 days, even if refrigerated. The mixture was mixed with distilled water, although the containers were not sterilised. The addition of 1 g/l of sodium azide would have reduced bacterial growth[153].

The problems with the Intralipid phantom show one of the main disadvantages of this

type of phantom compared to the solid phantom based on stable epoxy resin. Unfortunately, it is not possible to change the optical properties of the solid phantoms once they have been made.

9.3.3   Method 3 - using the Monte Carlo data.

A third method was used to estimate $\mu_s{}'$ from the experimental data measured with the cross-correlator. For this method the experimental data was compared to the results from the Monte Carlo simulations performed at seven different transport scattering coefficients (0.2, 0.3, 0.5, 1, 1.5, 2 and 3 mm^{-1}) by fitting the whole shape of the TPSF and not just the temporal position of the peak.

Since the impulse response of the system is approximately 380 ps FWHM, the Monte Carlo data needed to be convolved with a Gaussian of 380 ps FWHM. Without this convolution, the Monte Carlo simulations almost always generated a TPSF much faster than would ever be measured with the cross-correlator.

When the refractive index was 1.33, the Monte Carlo programme gave data at steps of every 4.43 ps, although the cross-correlator data was every 15.3 ps. These two time intervals were intended to be a convenient integer multiple of one another, although for various reasons they were not. A simple linear interpolation routine was therefore used to allow an estimate of the cross-correlator data at every 4.43 ps, even though it was measured every 15.3 ps.

Three corrections were applied to the experimental TPSFs before attempting a fit to the Monte Carlo data. Firstly, from the time axis, the time of the peak of the system impulse response at a source-detector spacing of zero was subtracted, as described in section 6.12. This is the same correction as used for all measurements. A second correction was necessary for this fitting method, which was not needed for the methods described in the last two sections (sections 9.3.1 and 9.3.2). As discussed in section 6.14 the cross-correlator data had a small non-zero baseline, that was more positive for earlier time than for later time, as shown in figure 9.16. This was even more apparent at very low light intensities. The average lock-in current for 4 samples at the start of the data acquisition (I_{start}) and 4 at the end (I_{end}) were therefore computed. For the data in figure 9.16, I_{start} would be +0.239 nA and I_{end} -0.078 nA. A linear baseline correction was then applied in software, so the start and end currents were zero (apart from noise). Since the start current was generally more positive than the end current, this effectively adds a tilt to the baseline. This tilting of the baseline of the collected TPSF would also move the peak position of the data slightly to the left, since the reference measurements were not subjected to baseline tilting. Any shift in the temporal position of the peak caused by the baseline compensation was

therefore allowed for, by moving all the data along the time axis so the temporal position of the peak was at the same time before and after the baseline correction.
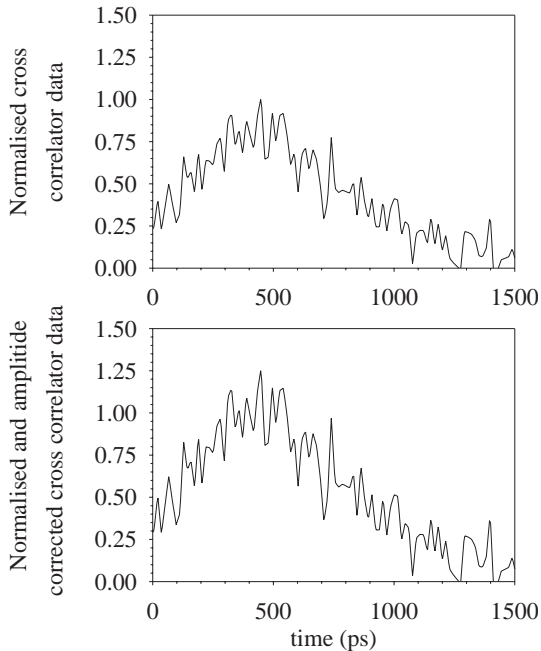
Figure 9.17 Experimental data collected at μ_s'=1.0 mm^{-1} μ=0.0105 mm^{-1} and a source-detector spacing of 20 mm, before (upper graph) and after rescaling (lower graph).
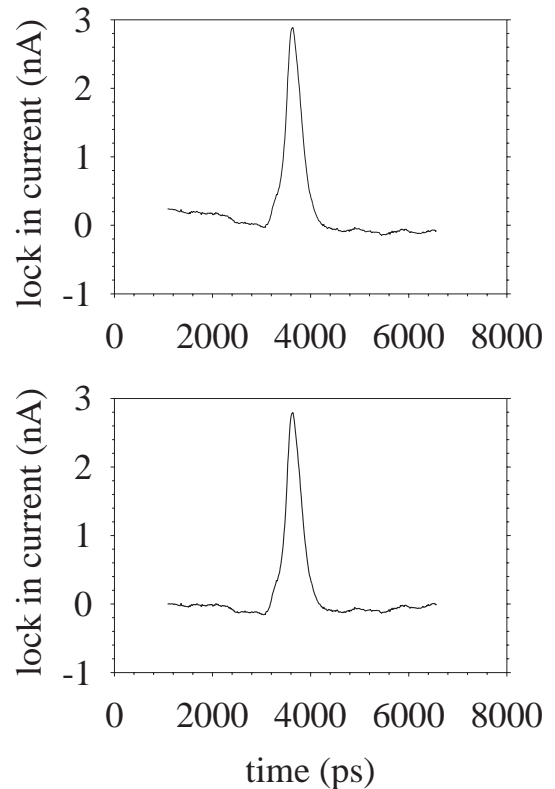
Figure 9.16 The upper diagram shows how the baseline is non-zero and that the offset is more positive for earlier times than for later times. The lower graph shows the same data after baseline correction.

Once the corrections had been applied to the cross-correlator data to allow for the baseline problems and to the Monte Carlo data to simulate a 380 ps FWHM system response, the experimental and convolved Monte Carlo data were both normalised to a peak height of 1.0. The normalisation was performed automatically by computer. The computer programme looked only for the peak amplitude of the data, and ensured all data was re-scaled so the peak was 1. Since the cross-correlator data was noisy at large source-detector spacing, it was sometimes necessary to make a manual correction of the amplitude of the cross-correlator data, so the peak was no longer 1.0. An example of where this was necessary is shown in figure 9.17. The normalised data has a peak height of 1.0, but this is not representative of the data. Hence all the data in the upper half of figure 9.17 was increased by 25% to that shown in the lower half of figure 9.17. The noise is now distributed equally both above and below 1 at the peak. It would be possible to automate this process, by fitting a curve through several points either side of the peak and scaling all data so the fitted curve had a peak amplitude of 1.0. Because of time limitations this was not done and therefore manual correction

of the data was necessary at 15 and 20 mm source-detector spacing, where the data was noisy. The noise on the data at 5 and 10 mm is insignificant, so this is not necessary.

Figure 9.18 shows (solid line with round symbols) experimental data collected at a source-detector spacing of 15 mm with a μ_a of 0.0018 mm^{-1} and a μ_s' of 1.0 mm^{-1}. Also shown (dashed line) is the Monte Carlo data at the same μ_a and source-detector spacing, but at either μ_s'=0.3, 1.0 or 3.0 mm^{-1}. Also shown (solid line, no symbols) is the convolution of the Monte Carlo data and the system response TPSF$_{MC}$(t)⊗R(t). Note that in the top graph, where the Monte Carlo data is for 0.3 mm^{-1}, the Monte Carlo data is much shorter in duration than the experimental data and even after being convolved with the system response, it is a poor match for the experimental data. In the lower graph, where the Monte Carlo data was simulated for μ_s'=3.0 mm^{-1}, TPSF$_{MC}$(t)⊗R(t) is much longer in duration than the experimental data. In the centre graph, where the Monte Carlo data was simulated for the same values of μ_s' as the experimental data was measured on, the fit between TPSF$_{MC}$(t)⊗R(t) and the experimental data is much closer. Since the Monte Carlo data is much longer in duration than the system response when $\mu_s' = 1.0$ or 3.0 mm^{-1}, convolving with the system response does not make such a significant difference at a μ_s' of 1.0 or 3.0 mm^{-1}, but it does in the upper graph, where μ_s' was only 0.3 mm^{-1}.

The differences between the two sets of data (TPSF$_{exp}$(t) and TPSF$_{MC}$(t)⊗R(t)) was then compared starting at t=0, to t_{max} in steps of 4.43 ps, where t_{max} was the time at which the experimental data had fallen to 5% of its peak amplitude. This cut-off avoided erroneous baseline data being included in the fitting procedure. A mean-square error ξ was calculated as:

$$\xi = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} [TPSF_{exp}(t) - TPSF_{MC}(t) \otimes R(t)]^2 \qquad \textbf{9.8}$$

Since there were 7 different Monte Carlo data files, the mean-square error ξ was computed at 7 different values of μ_s' using the known values of μ_a and source to detector spacing. Typical graphs of ξ against μ_s' are shown in figure 9.19, where the true μ_s' was 0.3 mm^{-1} for the top graph, 1.0 mm^{-1} for the centre graph and 3.0 mm^{-1} for the lower graph. A 4th order polynomial was fitted through the seven data points so that the mean-square error ξ could be estimated at any μ_s'. The algorithm used for fitting the polynomial was *mrqmin* from the book Numerical Recipes in C[142]. The minimum of the 4th order polynomial fit of ξ was then found using the routine *golden* from Numerical Recipes in C[142], which does a golden search for a minimum to a tolerance *tol* between two endpoints *ax* and *cx* that must be supplied. The parameter *tol* was
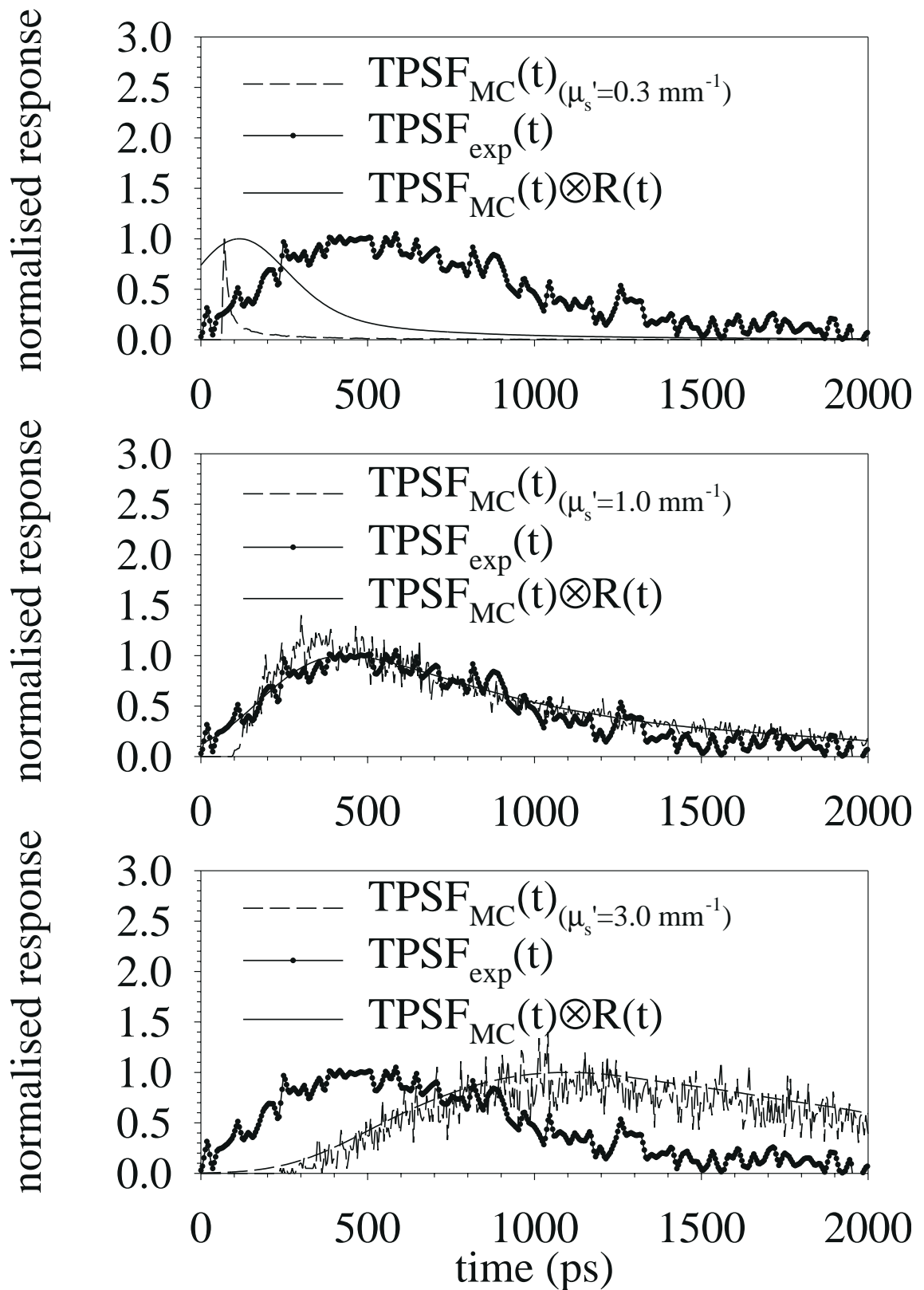
Figure 9.18 See text for explanation of graphs.

set to the square root of the computers floating point precision (3×10^{-8} since double precision
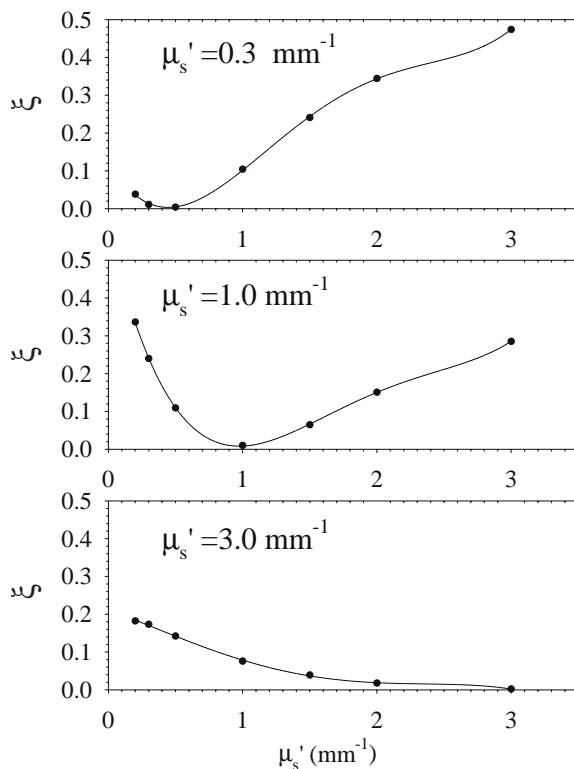
Figure 9.19 Graphs showing the mean-square error ξ between the convolved Monte Carlo data and the experimental data, for 3 different values of μ_s'.

was used), *ax* was set to 0.2 mm^{-1} and *cx* was set to 3.0 mm^{-1}.

The experimental data had been collected at transport scattering coefficients of 0.3, 1.0 and 3.0 mm^{-1}. The Monte Carlo data had been collected at seven transport scattering coefficients between 0.2 and 3.0 mm^{-1}. With the benefit of hindsight it is apparent that it would have been preferable to have the Monte Carlo data covering a significantly wider range of μ_s' than the experimental data. It is impossible to use the polynomial fits to evaluate μ_s' at 3.0 mm^{-1}, when 3.0 mm^{-1} is on the upper limit for which the polynomial is fitted assuming that we do not wish to risk the errors involved in extrapolation. The experimental data at a μ_s' of 0.3 mm^{-1} was not fitted to the polynomial either, since again the Monte Carlo simulations were not performed at a sufficiently low μ_s'.

An inspection of the experimental data for $\mu_s'=1$ mm^{-1} at 20 mm source-detector spacing is shown in figure 9.20, as a solid line, with round symbols. The Monte Carlo data is shown as a dashed line (also shown more clearly in figure 9.2). Clearly the Monte Carlo data at a source-detector spacing of 20 mm is virtually non-existent and therefore can not be used for the reasons discussed on page 221.

The fitting of the experimental data from the cross-correlator to the Monte Carlo data was performed by a C programme *fitmus*. The programme used standard C, so should compile without difficulty under linux, DOS, unix or any reasonable operating system, although this particular programme has only been checked on Sun workstations using gcc version 2.8.1 as the compiler. The programme *fitmus* required 4 command line parameters (the name of the data file generated by the cross-correlator, μ_a, the source to detector spacing and the temporal offset of all the data in the cross-correlator file. Several command line options could be used also. The only one regularly used was a scaling factor (-fscale) to multiply the experimental data
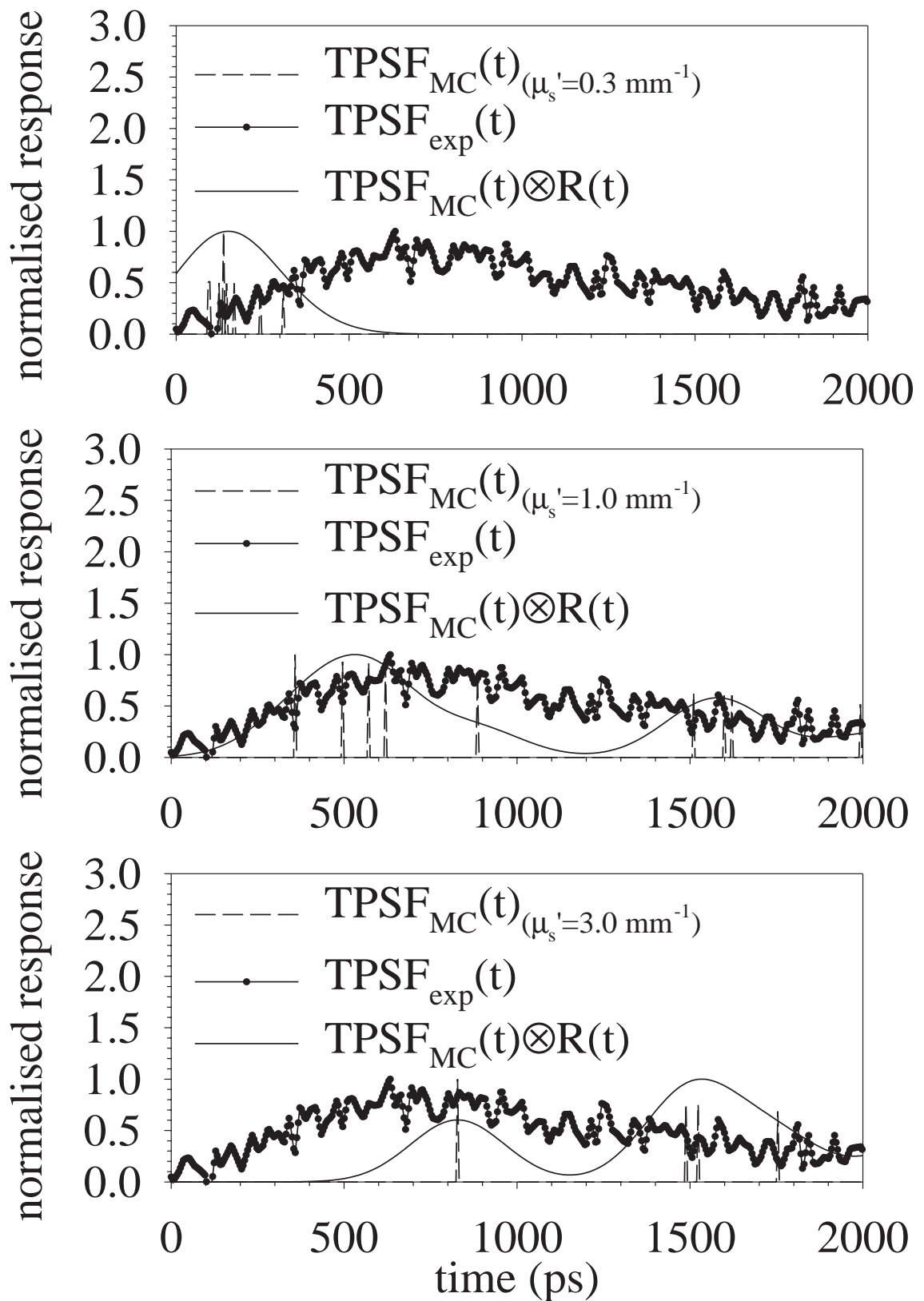
Figure 9.20 Graphs of the Monte Carlo data, experimental data and convolved MC data at 20 mm source-detector spacing. The noise on the MC data prevents it being used at 20 mm.

$TPSF_{exp}(t)$ before comparison to $TPSF_{MC}(t) \otimes G(t)$. For the data shown in figure 9.17, the option -f1.25 was used to scale all the cross-correlator data by 1.25. Other options allowed the order of the polynomial fit to be changed, although a 4th order polynomial was fine for all cases and so the option was not used. The following is a typical example of the output of the program. In this case it opens a data file produced by the cross-correlator called *s1_2_15.dat* which was collected with a μ_a of 0.0105 mm^{-1}, at a source-detector spacing of 15 mm, with an offset of 3622 ps on the data file. The estimated μ_s' was 0.982 mm^{-1} in this case. The polynomial fitting procedure used 431 iterations with a resulting χ^2 of 0.000017.

```
spice-girls /export/home/davek % fitmus -f1.15 s1_2_15.dat 0.0105 15 3622
t_peak=371.400000
Using musprime = 0.200000, error_sum=0.245094
Using musprime = 0.300000, error_sum=0.174914
Using musprime = 0.500000, error_sum=0.079574
Using musprime = 1.000000, error_sum=0.007060
Using musprime = 1.500000, error_sum=0.047533
Using musprime = 2.000000, error_sum=0.127013
Using musprime = 3.000000, error_sum=0.303172
chi^2=0.000017 itterations=431
mus'=0.982 spacing = 15 mua=0.010500 t_offset=3622.000000 (err=0.004479)
spice-girls /export/home/davek %
```

The computer programme took about 1 minute and 50 seconds to fit the experimental data against all 7 Monte Carlo data files and produce its estimate of μ_s'. This was using a rather old Sun IPC workstation (called *spice-girls*), running at approximately 25 MHz, with the Monte Carlo data files (8 Mb each) residing on the Sun's internal disk. Using a Sun Ultra 10 workstation with a much faster (300 MHz) processor did not significantly reduce the fitting time as the execution time was limited by the time to transfer the data files from disk to memory. Since the data files contain data at every source-detector spacing from 0 to 79 mm, but the data was fitted at only one spacing, it would be possible to reduce considerably the time by only reading in a small part of the data file. However optimising the fitting time was not an important component of this current work.

Figure 9.21 shows the estimated values of μ_s' at source to detector spacings of 5, 10 and 15 mm using the Monte Carlo data (method 3). This gives a mean estimate of 1.12 mm^{-1} which is an overestimation of 12 %. The standard deviation of the results are only 0.27 mm^{-1}. Again the errors are largest at 5 mm source to detector spacing, which is probably attributable to the accuracy of the system and not to the Monte Carlo data, which unlike the diffusion data, should be accurate close to the source. Unfortunately, since only one third of the experimental data has been analysed, it is not possible to say how this method performs other than at a true μ_s' of 1.0 mm^{-1}.
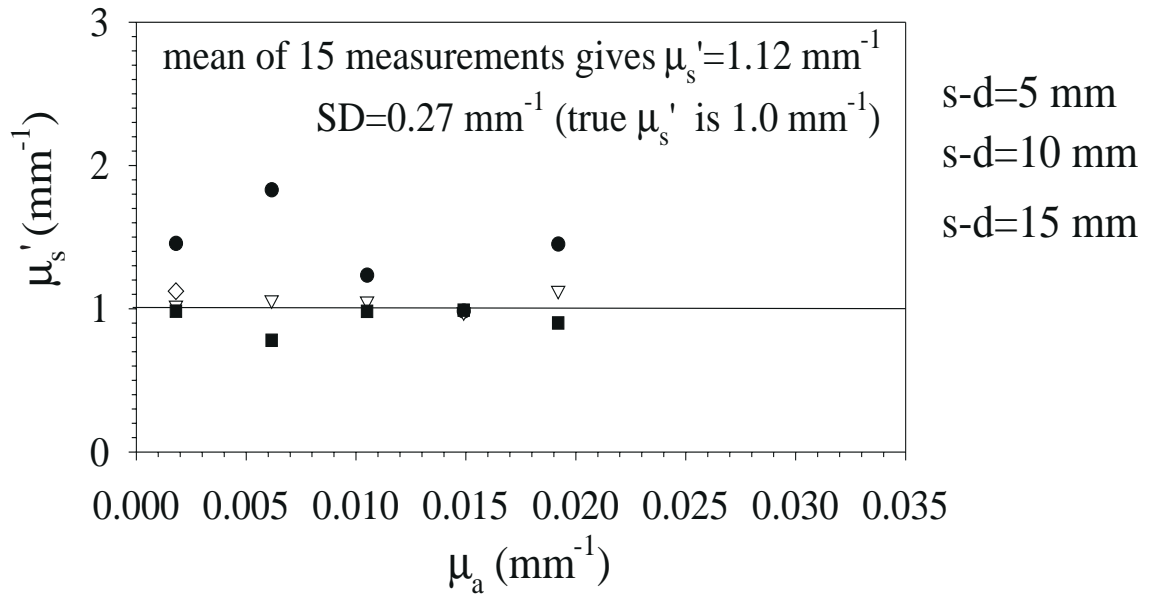
Figure 9.21 Graph showing the estimate of μ_s' from fitting to the Monte Carlo TPSF data (method 3). The line across the graph at $\mu_s'=1$ shows the actual μ_s'.

9.3.4 Method 4 - using diffusion theory to fit the full TPSF (not just t_{peak}).

Figure 9.22 shows the results of estimating μ_s' by fitting the experimental data to TPSFs calculated using the diffusion equation (equation 1.34). A programme *diffit* was written that generated TPSFs (called $TPSF_{diff}$) using the diffusion equation for an infinite volume (equation 1.34) for one fixed value of μ_a and one fixed source-detector spacing but at various values of μ_s' starting from 0.1 mm^{-1} and increasing up to a maximum of 5.0 mm^{-1} in steps of 0.5 mm^{-1}. These TPSFs were convolved with the system response R(t), to give a signal $TPSF_{diff}(t) \otimes R(t)$. A mean-square error ξ was computed between the experimental data $TPSF_{exp}$ and $TPSF_{diff}(t) \otimes R(t)$ in a similar way to used when fitting to the Monte Carlo data.

$$\xi = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} [TPSF_{exp}(t) - TPSF_{diff}(t) \otimes R(t)]^2 \qquad \textbf{9.9}$$

Since all experimental data was above the minimum μ_s' (0.1 mm^{-1}) used in evaluating the diffusion equation, the mean-square error ξ started off large, and gradually decreased as larger values of μ_s' were put in to the diffusion equation (equation 1.34). When the programme detected the mean-square error ξ starting to increase with increased μ_s', the programme backtracked two steps of μ_s' and changed the step size of μ_s' from steps of 0.5 mm^{-1} to steps of 0.05 mm^{-1}. Again when the closest fit using steps of 0.05 mm^{-1} was found, the step size was changed to 0.005 mm^{-1}. The following gives an example of the input and output of the
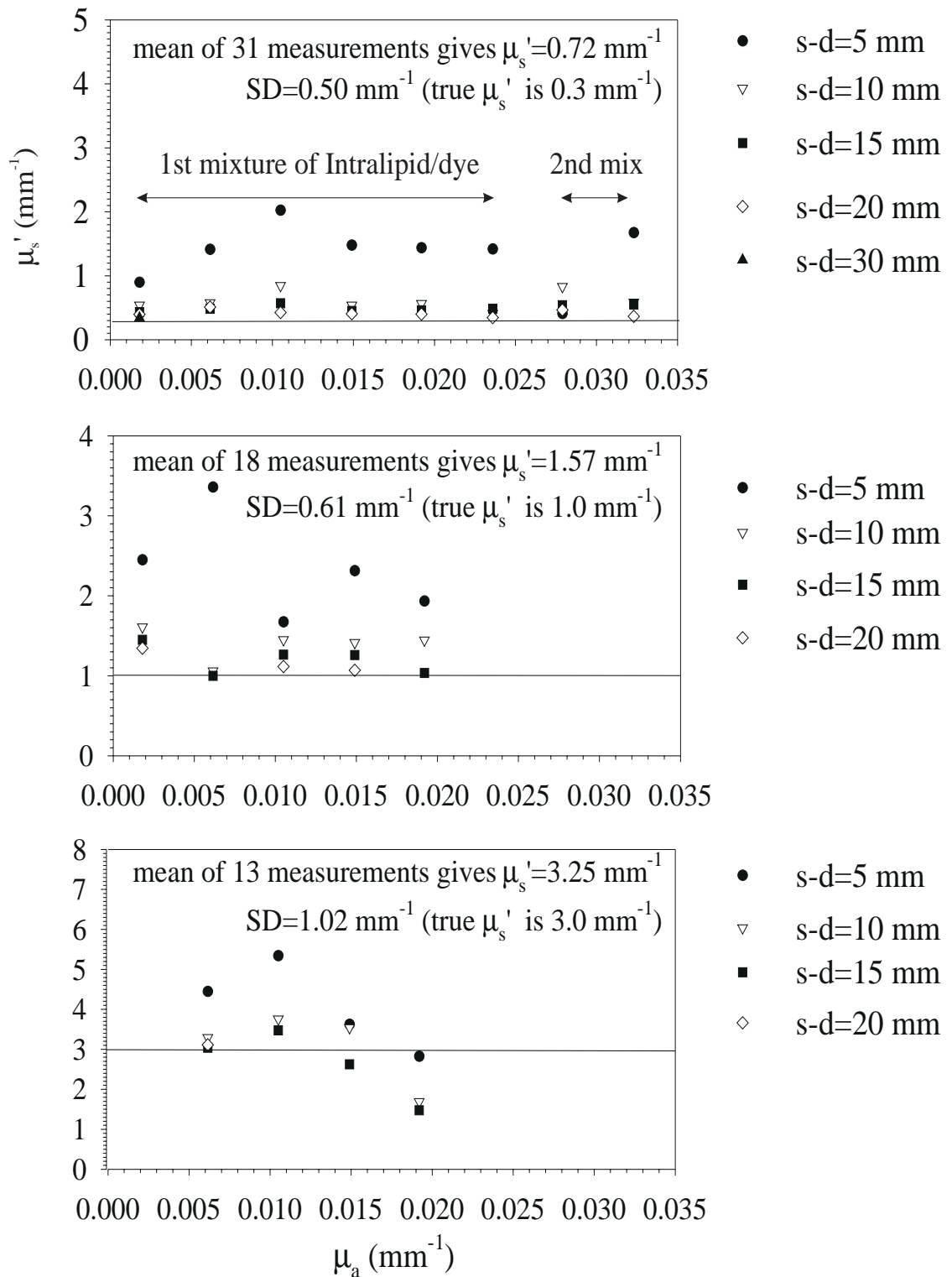
Figure 9.22 Estimations of μ_s' by fitting the full TPSF to diffusion theory (method 4).

programme *diffit*, when requested to find the μ_s' of a data set collected at a spacing of 10 mm, with a μ_a of 0.0236 mm^{-1} and where all data in the data file is offset from a true zero by 3641

ps. The programme used a parameter *error_sum* to represent ξ.

```
blondie /home/duke/davek % diffit 3s5_10m.dat 0.0236 10 3641
Using musprime = 0.100000, error_sum=0.0235840001294931
Using musprime = 0.600000, error_sum=0.0132768144890196
Using musprime = 1.100000, error_sum=0.0388348147457337
Using musprime = 0.100000, error_sum=0.0235840001294931
Using musprime = 0.150000, error_sum=0.0188549218152532
Using musprime = 0.200000, error_sum=0.0154542629409038
Using musprime = 0.250000, error_sum=0.0130752476109878
Using musprime = 0.300000, error_sum=0.0115097338309048
Using musprime = 0.350000, error_sum=0.0106153775410548
Using musprime = 0.400000, error_sum=0.0102870748209949
Using musprime = 0.450000, error_sum=0.0104445959970356
Using musprime = 0.350000, error_sum=0.0106153775410548
Using musprime = 0.355000, error_sum=0.0105579823278363
Using musprime = 0.360000, error_sum=0.0105065105365413
Using musprime = 0.365000, error_sum=0.0104609321705039
Using musprime = 0.370000, error_sum=0.0104212187259482
Using musprime = 0.375000, error_sum=0.0103870894609759
Using musprime = 0.380000, error_sum=0.0103560246267107
Using musprime = 0.385000, error_sum=0.0103305158949623
Using musprime = 0.390000, error_sum=0.0103105374763479
Using musprime = 0.395000, error_sum=0.0102960648698435
Using musprime = 0.400000, error_sum=0.0102870748209949
Using musprime = 0.405000, error_sum=0.0102835452797423
Using musprime = 0.410000, error_sum=0.0102824625661996
Using musprime = 0.415000, error_sum=0.0102860082407449
mus'=0.410000
blondie /home/duke/davek %
```

In this case, the programme estimates a μ_s' of 0.410 mm^{-1}. The fitting procedure only takes about 15 seconds on a fast (300 MHz) Sun Ultra 10 workstation. Being essentially CPU bound, the execution time is very much dependant on the CPU speed of the computer, unlike the fitting method described in section 9.3.3, which is IO bound and so is largely unaffected by the CPU speed of the computer.

This method was not particularly accurate (mean error=+140%) at a μ_s' of 0.3 mm^{-1}, but improves (mean error=+57%) at a μ_s' of 1.0 mm^{-1} and improves further (mean error=+8.3%) when the μ_s' is increased further to 3.0 mm^{-1}. The errors are significantly worse than average at a source-detector spacing of 5 mm, which is to be expected since the diffusion equation is less accurate at close spacings and since any errors in the instrument calibration will cause significantly more error at close spacings.

Since the diffusion equation is an accurate indicator of photon arrival time at large distances and since the Monte Carlo method did not generate accurate data at a source-detector spacing of 20 mm, a fifth fitting method was developed, which is described in the next section. This uses the best model at close spacings (the Monte Carlo method) and combines this with the

best indicator at large spacings (the diffusion equation).

9.3.5 Method 5 - use of combined Monte Carlo and diffusion equation data.

Using the whole TPSF generated from the Monte Carlo data (section 9.3.3) was found to be an improvement over just using the peak of the TPSF and fitting this one parameter to the diffusion equation (sections 9.3.1 and 9.3.2), but at a source-detector spacing of 20 mm, the data from the Monte Carlo simulations was very noisy. However, at 20 mm spacing, the diffusion equation can be used, so a fifth fitting procedure was developed to use the diffusion equation at a source-detector spacing of 20 mm, but use the Monte Carlo data for spacing of 5, 10 and 15 mm. Figure 9.23 shows the results from a combined Monte Carlo and diffusion equation,
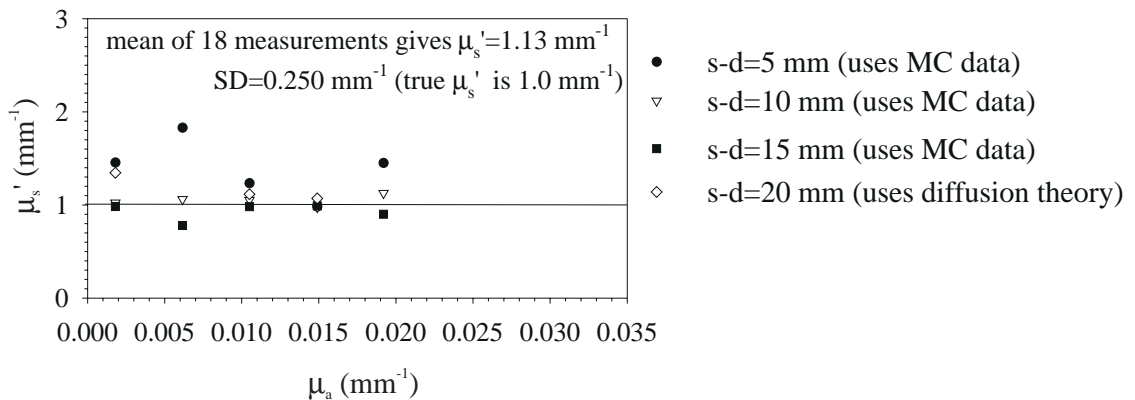
Figure 9.23 Results of estimating μ_s' from Monte Carlo simulations up to 15 mm, and diffusion theory beyond 15 mm. This is known as method 5.

where the Monte Carlo is used up to 15 mm, and the diffusion equation beyond 15 mm. As such this method combines methods 3 (section 9.3.3) and 4 (section 9.3.4), using each method at the most appropriate source-detector spacing. This method has a mean error of +13% at a μ_s' of 1.0 mm^{-1}, which is the only value of μ_s' used, for the reasons already stated in section 9.3.3. The error on the data at source-detector spacing of 5, 10 and 15 mm is the same as for method 3 and the error on the data at a spacing of 20 mm is the same as method 4. The 3 largest errors of the 15 measurements all occur at a source-detector spacing of 5 mm, which is again not surprising, as any inaccuracies in the instrument are most significant at close spacings where the duration of the TPSFs are short.

9.4 Comparison of the 5 different μ_s' fitting methods.

Five different methods have been attempted in order to derive μ_s' from experimental data

measured with the instrument data. It is helpful to compare these methods, in terms of absolute accuracy, standard deviation, execution time and the time required to generate data prior to attempting the fitting. This is shown in table 9.1.

Table 9.1 Comparison of 5 methods of estimating the transport scattering coefficient.

Method	μ_s'=0.3 mm^{-1}			μ_s'=1.0 mm^{-1}			μ_s'=3.0 mm^{-1}		
	mean	error(%)	(sd)	mean	error(%)	(sd)	mean	error(%)	(sd)
1	0.58	+93%	(0.21)	2.01	+101%	(1.01)	4.25	+42%	(1.14)
2	0.28	-6.7%	(0.10)	1.30	+30%	(0.48)	3.02	+0.7%	(0.49)
3	----	-----	(-----)	1.12	+12%	(0.27)	----	-----	(----)
4	0.72	+140%	(0.50)	1.57	+57%	(0.61)	3.25	+8.3%	(1.02)
5	----	-----	(-----)	1.13	+13%	(0.25)	----	-----	(----)

Method 1, described in section 9.3.1, whereby μ_s' was derived from the time to the peak of the measured TPSF (t_{peak}), ignoring any contribution of t_{peak} due to the instrument, was not very accurate, with a mean error of +93% at μ_s'=0.3 mm^{-1}, +101% at μ_s'=1.0 mm^{-1} and +46% at μ_s'=3.0 mm^{-1} and had a large standard deviation on the results. The data collected at a source-detector spacing of 5 mm (solid round circles on figure 9.12) shows the greatest error, with the data generally improving as the source to detector spacing was increased. This trend was to be expected, as those measurements at small spacing produced the shortest duration TPSFs, which were most affected by the 380 ps FWHM system response. It was possible that there were some changes in the optical properties of the phantom with increasing μ_a at a μ_s' 0.3 mm^{-1}, although it was not easy to see the changes with method 1 due to its poor accuracy. This method, while the fastest, was not very accurate nor significantly simpler than some of the other methods.

Reasonable accuracy, with mean errors ranging from -6.7% to +30% depending on μ_s', is seen to be achieved with method 2, described in section 9.3.2. In this method, the peak time of the TPSF was used, after taking into account the system response. This method was also very fast to implement. It required no database of Monte Carlo data, and little processing power to calculate. Although the calculations were performed on a Sun workstation, there is no reason why a PC could not have been used. In fact, the old Sun IPC with a CPU speed of 25 MHz used for many calculations is considerably slower than a modern PC. It's probably quite practical to

use a microcontroller, such as the Intel 8051 to control the instrument and perform this fitting procedure.

Methods 1 and 2, described in sections 9.3.1 and 9.3.2, have the great advantage that it is in principle very fast to acquire the experimental data they needed. Since only the peak position of the TPSF was required it wasn't necessary to acquire data on the full TPSF, which was necessary with the methods 3, 4 and 5. For these 3 methods, not only must the full TPSF be acquired, but some data either side must also be acquired so the baseline correction can be applied. Using methods 1 and 2 and some approximate algorithm (for example diffusion theory) for estimating the peak time, it would be possible to measure the peak position very fast. If the approximate time of the peak could be estimated to +/- 50 ps, which seems a perfectly reasonable assumption at modest source to detector spacings, then it would only be necessary to scan a 100 ps width of the TPSF to find the peak. Using steps of 15.3 ps, this is only 7 steps. At the fastest scan time of 8 Hz, taking 10 measurements at each position, would require only 1.125 seconds per time step, or 8.05 s. With some optimisation, it should be possible to perform the measurement and estimate μ_s' in less than 12 s.

Method 3, described in section 9.3.3, where the Monte Carlo data was used to fit to the experimental TPSF, generated slightly better accuracy (mean error of +12 %) than all the other methods for a μ_s' of 1.0 mm^{-1}, and also generated the lowest standard deviation (0.27 mm^{-1}). Since the Monte Carlo data had only been collected at a range of transport scattering coefficients from 0.2 to 3.0 mm^{-1}, virtually the same range as the experimental data, it was not really practical to use the data from the 4th order polynomial fit to estimate the μ_s' of the medium when the actual μ_s' was 0.3 or 3.0 mm^{-1}. For the data at μ_s'=3.0 mm^{-1}, it would be necessary to extrapolate the polynomial - a risky business at best. With the benefit of hindsight it would have been better to also have also generated Monte Carlo data for a μ_s' of 0.1, 5 and 10 mm^{-1}. As such, while this method does appear to work better than all the other methods, it has only been possible to make comparisons with one third of the experimental data collected.

Method 4, described in section 9.3.4, and using diffusion theory to fit the whole of the TPSF, was not very accurate (mean error=+140%) at a μ_s' of 0.3 mm^{-1}, but improved (mean error=+57%) at a μ_s' of 1.0 mm^{-1} and improved further (mean error=+8.3%) when the μ_s' was increased further to 3.0 mm^{-1}. The error would be expected to decrease at higher scattering coefficients for two reasons. Firstly the diffusion equation is more closely approximated when μ_s' is larger. Secondly, any random errors in calibration are less significant at higher values of μ_s' when the TPSFs are longer and so the calibration errors contribute a smaller fractional error.

An inspection of figure 9.22 shows by far the largest errors occur when the source-detector spacing is 5 mm, when the method overestimates by a considerable amount. A similar pattern occurs when μ_s' is 1.0 mm^{-1}, with all the estimates for data taken at a source-detector spacing of 5 mm giving large over-estimations, but all those at 10 mm or more giving much close estimates.

Method 5, described in section 9.3.5, which used both the Monte Carlo data at source-detector spacings of 15 mm and less and diffusion theory beyond this, had a mean error of +13% at a μ_s' of 1.0 mm^{-1}. The error the data at source-detector spacing of 5, 10 and 15 mm is the same as for method 3. The 3 largest errors of the 15 measurements all occured at a source-detector spacing of 5 mm, which is again not surprising since the measured TPSF will have been be most dominated by the instrument response.

Chapter 10.
Discussion and Conclusions.

The aims of the research were to develop an inexpensive time-domain system to measure the TPSF in tissue via implanted optical fibres from which the changes in the tissue scattering coefficient could be monitored. A range of different algorithms to determine from the TPSF changes in the tissue scattering properties has also been derived and tested. Both of these tasks have been completed successfully.

The cross-correlator (see chapter 6) is quite inexpensive (ignoring the laser cost, which was already in our laboratory) and which could be replaced by a pulsed semiconductor laser diode. Several expensive components (mainly optical) were purchased during development of the cross-correlator and subsequently found to be unnecessary, as the design changed, but the final system probably contains components totalling less than £10,000. The most expensive components being the commercial lock-in amplifier (£3,500), SRD module (£500), bias-tee (£500), phase shifter (£500), the various RF amplifiers and RF switch (probably another £700). Power supplies (1 x 24 V, 1 x 500 V, 1 x multi-output) probably total another £500 and the IO cards for the PC another £500. The PC used was already available in the laboratory. In contrast, the system shown in figure 1.23 cost several hundred thousand pounds in components alone, although this is admittedly for a 32-channel detection system and not a single channel like the current system. The performance of the system in figure 1.23 in terms of sensitivity and temporal resolution is also much better than the system described here. The use of much larger detection area (approximately 50 mm^2 per channel) of the microchannel plate PMT's gives improved sensitivity over the 0.03 mm^2 area of the APD used here. However, the 4 detectors (8 channels each) cost around £15,000 each - somewhat more than the £80 for a APD detector! The temporal response for the complete system shown in figure 1.23 is around 100 ps.

The cross-correlator is quite small (apart from the laser). Most items are in three 19" rack-mount modules, the front panels of which are 483 mm wide. The APD bias supply, is 2U (88 mm) high, the lock-in amplifier 3U high (132 mm) and the rest of the electronics also 3U high (132 mm). The PC controlling the system was a rather old design, and not very compact. However, there is no reason why a compact laptop machine could not be used. The system is very rugged, and no amount of abuse seems able to destroy the avalanche photodiode detectors.

Even about 500 mW of optical power focused onto the APD has not done it any harm - something that would certainly have destroyed a PMT. In 5 years of research, an APD was never destroyed.

For a time domain system, a pulsed laser was necessary. For the system described here, a large mode-locked Ti:sapphire laser pumped by an argon ion laser was originally used. Although this laser was large, it can in principle be replaced by a smaller semiconductor laser, so its size was not considered a concern, even though the aim was to design an instrument that could be made portable. Pulsed lasers are an area of rapid development, so it would be unwise to buy one until its portability was immediately required. IMRA America Inc (1044 Woodbridge Avenue, Ann Arbor, MI 48105, USA) produce a fibre doped laser, called the Femtolite™ with a 10 mW average power and a pulse width of 180 fs. Unfortunately, its 500 kHz maximum repetition rate is not suitable for generating pulses with a SRD. Clark-MXR GmBH (Weisenstr, 4, D-64625, Benscheim, Germany) produce a 775 nm laser with a pulse width of less than 100 fs, but a repetition rate of 37 MHz. Although this is somewhat lower than desired, it is just feasible for an SRD to be driven at 37 MHz to produce pulses. Pico Quant GmbH (Uternehmen für optoelektronische Forschung und Entwicklung, 12484, Berlin, Rudower Chaussee 5 (IG), Germany) produce a model PDL-800 laser with a choice of wavelengths between 635 and 910 nm, <40 ps FWHM pulse width and a peak power of 50 to 200 mW. Coherent produce the Vitesse-800 diode-pumped laser which has a very respectable 200 mW average power at 800 nm and produces pulses of less than 100 fs pulse width. Its repetition rate of 80 MHz is almost identical to the 82 MHz laser used for this project and would make an ideal portable laser source. At close source to detector spacings the reduction in output power from using a smaller laser would not increase the acquisition time, as in many experiments at close spacings, the 700 mW laser needed to be attenuated by several optical densities to prevent overloading of the system. However, at larger spacings, the time constant of the lock-in amplifiers filters would need be increased, with a resultant increase in the time necessary to acquire the data.

The main contributor to the width of the impulse response is the capacitance of the APD, which limits the rate at which the voltage across the diode can be changed. Unfortunately, there is apparently now little commercial interest in making smaller area APDs with a smaller capacitance. APDs with a 50 μm diameter collection area that were previously commercially available and which would no doubt have resulted in better performance, are no longer available. For most users who need a fast amplified photodetector, the combination of a normal PIN diode and a wide bandwidth amplifier gives better performance. Unfortunately, since the cross-

correlator needs a device with a gain that is rapidly variable, this approach is not possible here. Although variable-gain amplifiers are available (the Mincircuits ZFL-1000GH variable-gain amplifier was used in this project for an AGC system), the gain of these amplifiers can not be altered sufficiently fast to replace the APD. The system can not be made more sensitive to light without increasing the area of the APD, which will increase the capacitance and so further slow the temporal response.

One of the problems with the system is that it is necessary to take a reference or calibration measurement, to calibrate the time axis. Without this calibration, only relative, not absolute timings are possible, which have a much more limited use. This presents a problem, as to make the calibration measurement, the source and detector need first to be moved together, and a reference measurement taken, then moved apart for the actual measurement. Any drift of the system between taking these two measurements produces an error in the timings measured.

Figure 10.1 A possible method of obtaining a faster, more accurate time calibration.

There are several ways that this inconvenience could be avoided. Firstly, the detection APD could be switched between the output from the tissue and a direct sample from the laser (via some appropriate attenuation). This would require two optical switches, both of which would need to be 2-way. A simpler and probably superior system, would have been to split the laser beam into two paths, the first path going through the tissue, then to the APD, the second direct to the APD, as shown in figure 10.1. Hence two signals would fall on the APD at different times. If the fibres going through the tissue were pushed together, a calibration would tell the time delay between the reference path and the tissue path. This calibration should not change, and so only need to be done once. The system could then measure the relative time between these two signals, and use the earlier calibration data to get the true time. This would be inexpensive to implement and should make a large improvement in system accuracy and convenience.

The use of the SRD, whilst a novel method of pulse generation that allows for a totally electronic cross-correlator, does not produce the clean delta pulses really required. The Schottky
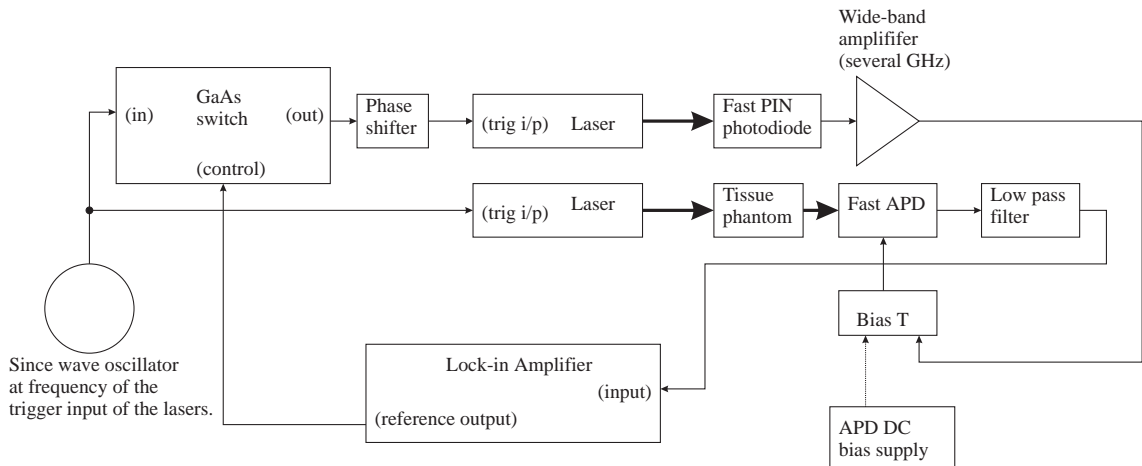
Figure 10.2 A possible method of improving the system response by dispensing with the SRD pulse generator, but still achieving a fully electronic system with no moving parts. Optical paths are shown as thick lines.

diodes added in series with the output improved this considerably, although it did degrade the temporal response from 275 ps to 380 ps FWHM. Figure 10.2 shows one possible way of implementing a totally electronic cross-correlator, without the SRD, but using a fast photodiode to generate the reference pulses. In the system outlined, a sine wave oscillator would drive two pulsed lasers - one directly and the other through a phase shifter. The laser driven through the phase shifter (a low power inexpensive model) would be used to generate reference pulses using a photodiode which is followed by a wide-band amplifier. The sine wave drive to the laser producing the references pulses would be switched on and off to turn the pulses on and off, as in the existing system. The phase shifter and lock-in amplifier would also be under computer control in the same way as in the current system. Several of the commercial lasers mentioned above can be locked to an external oscillator, including the Pico-Quant PDL-800, which we now have in our laboratory.

Measurements at source detector spacing of 5, 10, 15, 20 and 30 mm were made with the cross-correlator in a phantom (see chapter 7) consisting of a mixture of distilled water, Intralipid^{TM} and an infra-red absorbing dye, at a range of different optical properties. Full TPSFs were measured for each spacing, μ_a and μ_s'. The transport scattering coefficients used were μ_s'=0.3, 1.0 and 3.0 mm^{-1}. The absorption coefficients used were μ_a = 0.0018, 0.00616, 0.0105, 0.0149, 0.0192, 0.0236, 0.0279 and 0.0323 mm^{-1}. Several different algorithms were considered to enable μ_s' to be determined from the measured TPSF. The first, (called method 1 and described in section 9.3.1) did not work well due to the temporal resolution of the system. It has a mean error of +93% at a μ_s' of 0.3 mm^{-1}, +101% at a μ_s' of 1.0 mm^{-1} and +46% at a μ_s' of

3.0 mm^{-1}. The errors are largest at close source-detector spacings, which is to be expected. The second method (called method 2 and described in section 9.3.2), improved considerably on these results, and required little additional computational effort. The mean error was between -6.7% and +30% depending on tissue optical properties. Both of these methods involved comparing the peak time of the TPSF measured experimentally (the largest of which was 1560 ps), to that expected from diffusion theory. The first method ignored the effect of the system response, whereas the second, which worked much better, took this into account. Both methods were quite poor at close source-detector spacing, particularly 5 mm, where the errors were typically twice the error at 10 to 30 mm spacing.

Several other methods were attempted based on fitting to the whole TPSF - not just the peak time. The first of the methods fitting the whole TPSF (called method 3 and described in section 9.3.3) used data from a Monte Carlo model (see chapter 8) that needed to be simulated first. Due to an insufficient range of transport scattering coefficients being simulated with this model, which would be too time-consuming to rectify quickly, method 3 was only tested when the experimental μ_s' was 1.0 mm^{-1}. The mean error was 12%, although most of this error was again due to measurements made with a source-detector spacing of 5 mm, where all the measurements were more than 25% in error 3 of the 4 measurements made at 5 mm spacing had an error of at least 40%. In contrast all the measurements made with a source-detector spacing of 10 mm or more were predicted within 25% of the correct value.

The Monte Carlo model was developed to allow the effect of local heteogeneities in the optical properties of the tissue. The general structure of the model resulted in long computation times, as most of the photons leaving the source were never detected. To reduce the time required to obtain results of useful statistical significance, the programme was subsequently modified to allow execution on up to 24 different processors on a computer network, even though the computers were of different manufacturers, using different operating systems and processors using different byte ordering. This parallel programme resulted in a considerable decrease in execution time. It is believed to be the first time a Monte Carlo computer programme has been written to run on such a diverse network of computers[145]. Once a data base of Monte Carlo data is obtained, the time required to fit the experimental data to that of the Monte Carlo data, was about 2 minutes. Since the process is essentially IO bound, it does not depend greatly on the CPU speed of the computer. However, this time could be reduced very considerably (perhaps eighty times) by reading from disk only the section of Monte Carlo data required, rather than the whole data file, most of which is not needed.

Another algorithm (section 9.3.4) for estimating μ_s' from the experimental measurements, used the diffusion equation for an infinite volume (equation 1.34). Using the Monte Carlo data resulted in a slight improvement in the estimations of μ_s' over all other methods at a μ_s' of 1 mm^{-1}, although since Monte Carlo data was not generated for values of μ_s' above 3.0 mm^{-1} it was not possible to estimate μ_s' at the highest μ_s' used in the Intralipid^{TM} tank.

10.1   Further work.

One piece of work that will definitely be performed in the fairly short term is to run the Monte Carlo programme for a wider range of μ_s' than has been done to date. Simulations at a μ_s' of 0.1 mm^{-1}, 5 mm^{-1} and 10 mm^{-1} will allow the fitting of the experimental cross-correlator data collected at a μ_s' of 0.3 mm^{-1} and 3 mm^{-1} to the Monte Carlo data using the methods described in sections 9.3.3 and 9.3.5. The simulations will be time consuming, but once set running require no human intervention until they are finished. The fitting procedures as discussed earlier, do require human intervention. Fortunately, this work will cost nothing in terms of hardware or software costs.

A second piece of work, that may be performed, time permitting, is to analyse the data simulated in the Monte Carlo model for a detection fibre at 90° to the source fibre. It will be useful to determine if the TPSFs generated are longer for the same radial distance between source and detector than when the source and detector fibres are in line.

Most of the other improvements that are possible and have been discussed in the previous chapter, unfortunately require more money than is likely to be available.

Publications arising from this work.

The following is a list of publications describing this work.

Kirkby D. R. and Delpy D. T., "Measurement of the tissue temporal point spread function (TPSF) by use of a cross-correlation technique with an avalanche photodiode detector", Proc. SPIE, **2389**, pp 190-197, (1995).

Kirkby D. R. and Delpy D. T., "Measurement of tissue point spread function by a gain-modulated avalanche photodiode detector", Phys. in Med. and Biol., **41**, pp 939-949, (1996).

Kirkby D. R. and Delpy D. T., "An optoelectronic cross-correlator using a gain-modulated avalanche photodiode for measurement of the tissue temporal point spread function", Proc OSA, **TOPS 2**, pp 108-112, (1996).

Kirkby D. R. and Delpy D. T., "Parallel operation of Monte Carlo simulations on a diverse network of computers", Phys. Med. Biol., **42**, pp 1203-1208, (1997).

Appendix 1.
Abbreviated data sheets for some electronic devices used.

This Appendix gives the most important data, taken from manufacturers data sheets, of the main electronic devices used. All the RF components were designed for use in a 50 Ω system.

Hamamatsu S2381 APD.

Parameter	Typical	Maximum	Units
Effective Area	0.0314		mm^2
Spectral Response	350-1050		nm
Quantum efficiency (λ=830 nm, M=1)	70		%
Peak response	830		nm
Breakdown voltage	120	300	V
Junction capacitance	2		pF
Fall time	200		ps
Gain	100		-
Excess Noise index	0.30	0.42	-

Minicircuits ZHL-32A power amplifier.

Freq range	Gain (min)	Maximum output power (1 dB compression)	Maximum input power without damage
50 kHz to 130 MHz	25 dB	+29 dBm	+10 dBm

Minicircuits ZFL-1000GH Variable Gain Amplifier.

Freq range	Gain (min)	Control range (0-5 V)	Maximum output power (1 dB compression)	Maximum input power without damage
10 MHz-1.2 GHz	24 dB	≥30 dB	+13 dBm	+10 dBm

Minicircuits ZYSWA-2-50DR GaAs RF switch.

Frequency range (GHz)	Insertion Loss (dB)	1 dB compression point	In-out isolation (dB)
DC-5 GHz	0.9 dB (typical) 1.5 dB (maximum)	20 dB (typical) 15 dB (minimum)	50 dB (typical) 41 dB (minimum)

Manufacturers test data for the Herotek GC082-142 SRD harmonic generator.

RF drive required	Drive frequency	RF output level	Input VSWR
+27 dBm	82 MHz	≥ -2 dBm (up to 4 GHz) ≥ -17 dBm (4 - 8 GHz) ≥ -27 dBm (8 - 12.4 GHz) ≥ -32 dBm (12.4 - 18 GHz)	1.92:1

R&K model PS-3 Phase shifter.

Input frequency	Phase shift range	Insertion loss	Maximum RF input	Phase control linearity
82 MHz	0° to 360°	7 dB maximum	+20 dBm	15%

References.

1. Vasil'ev P.,"Ultrafast Diode Lasers - Fundamentals and Applications", Artech House, Boston, USA, (1995).

2. Delpy D. T., Cope M., Van der Zee P., Arridge S., Wray S. and Wyatt J., "Estimation of optical pathlength through tissue from direct time of flight measurement", Phys. Med. Biol., **33**, pp 1433-1442, (1988).

3. Henderson B. W. Dougherty T. J., "How does photodynamic therapy work?", Photochem. Photobiol., **55**, pp 145-157, (1992).

4. Wilson B. C., Patterson M. S. and Lilge L., "Implicit and Explicit Dosimetry in Photodynamic Therapy: a New Paradigm", Lasers in Medical Science, **12**, pp 182-199, (1997).

5. Bown S. G., "Photodynamic Therapy to Scientist and Clinicians - One World or Two?", Journal of Photochemistry and Photobiology, B. Biology, **6**, pp 1-12, (1990).

6. Wyman D. R., "Selecting source locations in multi-fibre interstitial laser photocoagulation", Lasers in Surgery and Medicine, **13**, pp 656-663, (1993).

7. Wyman D. R. and Whelan W. M., "Basic optothermal diffusion theory for laser photocoagulation", Med. Phys, **21**, pp 1651-1656 (1994).

8. Amin Z., Buonaccorsi G., Mills T., Harries S., Lees W. R. and Bown S. G., "Interstitial laser photocoagulation - evaluation of a 1320 nm Nd-Yag and an 805 nm diode-laser - the significance of charring and the value of precharring the fiber tip", Lasers in Medical Science, **8**, pp. 113-120, (1993).

9. Stegar A. C., Lees W. R., Walmsley K. and Bown S. G., "Interstitial laser hyperthermia: a new approach to local destruction of tumours", British Medical Journal, **299**, pp 362-365, (1989).

10. Blier A. R., Jolez F. A., Cohen M. S., Weisskoff R. M.,Dalcanton J. J., Higuchi N., Feinberg D. A., Rosen B. R., McKinstry R. C. and Hushek S. G.,"Real-time magnetic resonance imaging of laser heat deposition in tissue", Mag. Res. in Med., **21**, pp 132-137, (1991).

11. Essenpreis M., "Thermally induced changes in optical properties of biological tissues", Ph.D. thesis, University of London, (1992).

12. Benaron D. A., Ho D. C., Rubinsky B. and Shannon M., "Imaging (NIRI) and Quantitation (NIRS) in Tissue Using Time-Resolved Spectrophotometry: The Impact of Statically and Dynamically Variable Optical Path Length", Proc. SPIE, **1888**, pp 10-21, (1993).

13. Jöbsis F. F., "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters", Science, **198**, pp 1264-1267, (1997).

14. Ashton N., Ward B. and Serpell G. "Role of oxygen in the genesis of retrolental fibreplasia: a preliminary report", Br. J. Opthalmol., **37**, pp 513-520, (1953).

15. Kinsey, V. E., Jacobus J. T., and Hemphill F. M., "Retrolental fibroplasia: cooperative study of retrolental fibroplasia and the use of oxygen", Arch. Opthalmol.,**56**, pp 481-487, (1956).

16. Horecker B. L.,"The absorption spectra of haemoglobin and its derivatives in the visible and near infra-red regions", J. Biol. Chem., **148**, pp 173-183, (1943).

17. Cope M., "The development of a near infrared spectroscopy system and its application for the non invasive monitoring of cerebral blood and tissue oxygenation in the newborn infant", Ph.D thesis, University of London, (1991).

18. Drain L. E., "Coherent and non-coherent methods in Doppler optical beat velocity measurement", Journal of Physics D: Applied Physics, **5**, pp 481-495, (1972).

19. Drain L. E., "The LASER Doppler technique", John Wiley & Sons, London, (1980).

20. Riva C., Ross B. and Benedek G. B., "Laser Doppler measurements of blood flow in capillary tubes and retinal arteries", Investigative Ophthalmology, **11**, pp 936-944, (1972)

21. Stern M. D., "*In vivo* evaluation of microcirculation by coherent light scattering", Nature, **254**, pp 56-58, (1975).

22. Doppler C., "Ueber fas farbige licht der dopplersterne und einiger anderer Gestirne des Himmels", Abhandl Konigl Bohm Ges, **2**, pp 465-482, (1893).

23. Obeid A. N., Boggett D. M., Barnet N. J. and Dougherty G., "Depth discrimination in laser Doppler skin blood flow measurement using different lasers", Med. & Biol. Eng. & Comput., **26**, pp 415-419, (1988).

24. Dougherty G. "A laser Doppler flowmeter using a variable coherence length to effect depth discrimination", Review of Scientific Instruments, **65**, pp 3220-3221, (1992).

25. Essex T.J.H. and Byrne P. O., "A laser Doppler scanner for imaging blood flow in skin", J. Biomedical Engineering, **13**, pp 189-194, (1991).

26. Harrison D. K., Evans S. D., Abbot N. C., Beck J. S. and McCollum P. T., "Spectrophotometric Measurements of Haemoglobin Saturation an Concentration in Skin during the Tuberculin Reaction in Normal Human Subjects", Clinical Physics and Physiological Measurement, **13**, pp 349-363, (1992).

27. Feather J. W., Hajizadeh-Safer M., Leslie G. and Dawson J. B, "A Portable Scanning Reflectance Spectrophotometer using Visible Wavelengths for the Rapid Measurement of Skin Pigments", Phys. Med. Biol., **34**, pp 1883-1900, (1989).

28. Patterson M. S., Wilson C. and Whman D. R., "The propagation of optical radiation in tissue I. Models of radiation transport and their application", Lasers in Medical Science, **6**, pp 155-167, (1991).

29. Bohren C. F. and Huffman D. R., "Absorption and scattering of light by small particles", Wiley Interscience, New-York, (1983).

30. Bolin F. P., Luther L. E., Taylor R. C. and Ference R. J., "Refractive index of some mammalian tissues using a fibre optic cladding method", Applied Optics, **28**, pp 2297-2303, (1989).

31. Ishimaru A. and Kuga Y., "Attenuation constant of a coherent field in a dense distribution of particles", Journal of the Optical Society of America, **72**, pp 1317-1320, (1982).

32. West R., Gibbs D., Tsang L. and Fung A.K., "Comparison of optical scattering and the quasi-crystalline approximation for dense media", J. Opt. Soc. Am., **11**, pp 1854-1858, (1994).

33. Patterson M. S., Wilson C. and Wyman D. R., "The propagation of optical radiation in tissue. II: Optical properties of tissues and resulting fluence distributions", Lasers in Medical Science, **6**, pp 379-390, (1991).

34. Cheong W., Prahl S. A. and Ashley J. W. ,"A review of the optical properties of biological tissues", IEEE J. Quantum Electronics, **26**, pp 2166-2184, (1990).

35. Mitic G., Kölzer, Otto J., Piles E. Sölkner S. and Zinth W., "Time-gated transillumination of the biological tissues and tissuelike phantoms", Applied Optics, **33**, pp 6699-6710, (1994).

36. Kubelka P. and Munk F., "Ein Beitrag zur Optic der Farbanstriche", Z Tech Phys, **12**, 593-601, (1931).

37. Chandrasekhar S., "Stochastic problems in physics and astronomy", Rev. Mod. Phys., **15**, pp 1-88, (1943).

38. Arridge S. R., Cope M. and Delpy D. T., "The theoretical basis for the determination of optical pathlength in tissue: temporal and frequency analysis", Phys. Med. Biol., **37**, pp 1531-1560, (1992).

39. Patterson M. S., Chance B., and Wilson B. C., "Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties", Applied Optics, **28**, pp 2331-2336, (1989).

40. Livesley R. K., "Finite elements: an introduction for engineers", Cambridge University Press, Cambridge, UK, (1983).

41. Silvester P. P. and Ferrari R. L., "Finite elements for electrical engineers", 2nd edition, Cambridge University Press, Cambridge, UK, (1990).

42. Itoh T., "Numerical Techniques for microwave and millimetre-wave passive structures", Chapter 2 "The finite element Method", pp 33 - 132, Wiley Interscience, New York, (1989).

43. Kirkby D. R., "Finding the characteristics of Arbitrary Transmission Lines", QEX, pp 3-10, (December 1996).

44. Ramo S., Whinnery J. R., and Van Duzer T.,"Fields and Waves in Communication Electronics", 2nd edition, Wiley, (1984).

45. Director L. B., Frid S. E., Mendeleev V. Y. and Scovorodko S. N., "Computer simulation of heat and mass transfer in tissue during high-intensity long-range laser irradiation", Anal. New York Acad. Sci., **858**, pp 56-65, (1998).

46. Yao Y. Q., Wang Y., Pei Y. L., Zhu W. W. and Barbour R. L., "Frequency-domain optical imaging of absorption and scattering distributions by a Born iterative method", J. Opt. Soc. Am. A, **14**, pp 325-342, (1997).

47. Arridge S. R.,"A finite element approach for modelling photon transport in tissue", Med. Phys, **20**, pp 299-309, (1993).

48. Schweiger M., Arridge S. R. and Delpy D. T., "Application of the finite-element method for the forward and inverse models in optical tomography", Journal of Mathematical Imaging and Vision, **3**, pp 263-283, (1993).

49. Schweiger M., Arridge S. R., Hiraoka M. and Delpy D. T., "The finite element method for the propagation of light in scattering media: Boundary and source conditions", Med. Phys, **22**, pp 1-13, (1995).

50. Schweiger M., "Application of the finite element method in infrared image reconstruction of scattering media", Ph.D. thesis, University of London, (1994).

51. Lux I. and Koblinger L., "Monte Carlo particle transport methods: neutron and photon calculations", CRC Press, Florida, USA, (1990).

52. G. Mie, "Beitrage zur Optic trüber Medien spieziell kolloidaler Methallösungen", Ann. Phys., **25**, pp 377-445, (1908).

53. Henyey L. C. and Greenstein J. L., "Diffuse radiation in the galaxy", Astrophys. J., **93**, pp 70-83, (1941).

54. Van der Zee P., "Measurement and modelling of the optical properties of human tissue in the near infrared", Ph.D. thesis, University of London, (1992).

55. Farrell T. J., Wilson B. C. and Patterson M. S., "The use of a neural network to determine tissue optical properties from spatially resolved diffuse reflectance measurements", Physics in Medicine and Biology, **37**, pp 2281-2286, (1992).

56. Marquet P., Bevilicqua F., Depeuringe C. and de Haller E. B., "Determination of reduced scattering and absorption coefficients by a single charge coupled-device array measurement, part 1: comparison between experiments and simulations", Optical Engineering, **34**, pp 2055-2063, (1995)

57. Bevilacqua F., Marquet P., Depeusinge C and de Haller E. B., "Determination of reduced scattering and absorption coefficients by a single charge-coupled-device array, Part II: Measurements on biological tissues", Optical Engineering, **34**, pp 2064-2069, (1995).

58. Patterson M. S., Schwartz E. and Wilson B. C., "Quantitave reflectance spectrophotometry for the non-invasive measurement of photosensitizer concentration in tissue during photodynamic therapy", Proc. SPIE, **1056**, pp 115-122, (1989).

59. Matcher S. J., Kirkpatrick P., Nahid K., Cope M. and Delpy D. T., "Absolute quantification methods in tissue near infra red spectroscopy", Proc. SPIE, **2389**, pp 486-495, (1995).

60. Berndt K., Klose E., Schwarz P., Feller K. H. and Fabler, "Time resolved fluorescence spectroscopy of cyanine dyes", Z. Phys., Chemie, Leipzig, **265**, pp 1079-1086, (1984).

61. Duncan A., Whitlock T. L., Cope M. and Delpy D.T. "A multiwavelength, wideband, intensity modulated optical spectrometer for near infrared spectroscopy and imaging", Proc. SPIE, **1888**, pp 248-257, (1993).

62. Fishkin J. B. and Gratton E., "Propagation of photon-density waves in strongly scattering media containing an absorbing semi-infinite plane bounded by a straight edge", J. Opt. Soc. Am. A, pp 127-140, (1993).

63. Fishkin J. B., So P. T. C., Cerussi A. E., Fantini S., Franceschini M. A. and Gratton E., "Frequency-domain method for measuring spectral properties in multiple-scattering media: methaemoglobin absorption spectrum in a tissuelike phantom", Applied Optics, **34**, pp 1143-1155, (1995).

64. Gratton E., Jameson D. M., Rosato N. and Weber G., "Multifrequency cross-correlation phase fluometer using synchrotron radiation", Rev. Sci. Instrum., **55**, pp 486-494, (1984).

65. Berndt K., "Application of a gain-modulated avalanche photodiodes in phase-sensitive fluorescence spectroscopy", Optics Communication, **56**, pp 30-35, (1985).

66. Berndt K., Durr H. and Palme D., "Picosecond fluorescence lifetime detector", Optics Communication, **55**, pp 271-276, (1985).

67. Chance B., Maris M., Sorge J. and Zhang M. Z., "A phase modulation system for dual wavelength difference spectroscopy of haemoglobin deoxygenation in tissues", Proc. SPIE., **1204**, pp481-491, (1990).

68. "Photomultiplier tube - principle to application", Hamamatsu Photonics K. K., 1126-1, Ichino-cho, Hamamatsu City, 435, Japan, (1994).

69. Wilson J. and Hawkes J. F. B., "Optoelectronics - an introduction", Prentice-Hall, London, Great Britain, (1983).

70. Gower J., "Optical communication systems", Prentice-Hall, London, Great Britain, (1984).

71. Jenkins T. E., "Optical sensing techniques and signal processing", Prentice-Hall, Hertfordshire, Great Britain, (1987).

72. Fishkin J. B. and Gratton E., "Propagation of photon-density waves in strongly scattering media containing and absorbing semi-infinite plane bounded by a straight edge", J. Opt. Soc. Am. A, **10**, pp 127-140, (1993).

73. Kohl M., Watson R. and Cope M., "Optical properties of highly scattering media determined from changes in attenuation, phase and modulation depth", Applied Optics, **36**, pp 105-114, (1997).

74. Yoo K. M. and Alfano R. R., "Imaging object hidden in highly scattering media using femtosecond second-harmonic-generation cross-correlation time gating, Opt. Lett, **16**, pp 1019-1021, (1991).

75. Choi N. K. and Taylor H. F., "Novel cross-correlation technique for characterisation of subpicosecond pulses from mode-locked semiconductor lasers", Appl. Phys. Lett., **62**, pp 1875-1877, (1993).

76. J. R. Lakowicz (editor), "Topics in fluorescence spectroscopy", volume 1: techniques, Plenum Press, New York, USA, (1991).

77. Fichtner W. and Häcker W., "Time resolution of Ge avalanche photodiodes operating as photon counters in delayed coincidence", Rev. Sci. Instrum., **47**, 3, pp 374-377, (1976).

78. Häecker W., Groezinger O. and Pilkuhn M. H., "Infrared photon counting by Ge avalanche photodiodes", Appl. Phys. Lett., **19**, 4, pp 113-115, (1971).

79. Buller G. S., Fancy S. J., Massa J. S., Walker A. C., Cova S. and Lacaita A., " Time-resolved photoluminescence measurements of InGaAs/InP multiple-quantum-well structures at 1.3 μm wavelengths by use of germanium single-photon avalanche photodiodes", Applied Optics, **35**, pp 916-921, (1996).

80. Cova S., Lacaita A., Zappa F. and Lovati P., "Avalanche photodiodes for near-infrared photon-counting", Proc SPIE, **2388**, pp 56-66, (1995).

81. Cova S., Lacaita A., Ghioni M., Ripamonti G. amd Townsend P. D., "20 ps timing accuracy resolution with single photon avalanche photodiodes", Rev. Sci. Instrum., **60**, pp 1104-1110, (1989).

82. "Instrumentation for Fluorescence Lifetime Spectroscopy", EG&G application note, AN50; EG&G, 100 Midland Road, Oak Ridge, TN 37831-0895, USA.

83. Hebden J. C., Schmidt F. E. W., Fry M. E., Schweiger M., Hillman E. M. C., Delpy D. T. and Arridge S., "Simultaneous reconstruction of absorption and scattering images using multi-channel measurement of purely temporal data", Opt. Lett., **24**, pp 534-536, (1999).

84. Brigham E. O., "The fast Fourier transform", Prentice Hall, (1988).

85. Cooper G. R. and McGillem C. D., "Methods of signal and system analysis", Holt Reinholt and Winston Inc, New Yory, USA, (1967).

86. Cruz J. B. and Van Valkenburg M. E., "Signals in linear circuits", Houghton Mifflin Company, Boston, USA, (1974).

87. Panter P. F., "Modulation noise and spectral analysis", McGraw Hill, New York, USA, (1965)

88. Berndt K., "Picosecond opto-electronic cross-correlator with single photon sensitivity", Optics Communication, **61**, pp 33-37, (1987).

89. Hebden J. C., Arridge S. R. and Delpy D. T., "Optical imaging in medicine I. Experimental techniques", Phy. Med. Biol., **42**, pp 825-840, (1997).

90. Hieftje G. M. and Haugen G. R., "Correlation-based approaches to time-resolved fluorimetry", Analytical Chemistry, **53**, pp 755A - 763A, (1981).

91. Svelto O., "Principles of Lasers", Plenum Press, New-York, (1982).

92. Wilson J. and Hawkes J.F.B., "Lasers - principles and Applications", Prentice Hall, (1987).

93. Agrawal G. P. and Dutta N. K., "Long-Wavelength Semiconductor Lasers", Van Nostrand, (1986).

94. Petermann K., "Laser Diode Modulation and Noise", Kluwer Academic Publishers, Netherlands, (1988).

95. Yariv A., "Optical Electronics", 3rd Edition, Holt-Saunders, Japan, (1985).

96. Senior J. M., "Optical Fibre Communications - Principles and Practice", Prentice-Hall, USA, (1985).

97. Gowar J., "Optical Communication Systems", Prentice-Hall, USA, (1984).

98. Van de Ziel, J. P., Logan, R. A. and Mikulyak R. M., "Generation of Subpicosecond pulses from an actively Mode-Locked GaAs Laser in an External Ring Cavity", Applied Physics Letters, **39**, pp 867-869, (1981).

99. Bowers, J. E., Morton P. A., Mar A., and Corzine S. W., "Actively Mode Locked Semiconductor Lasers", IEEE Journal of Quantum Electronics, **60**, pp 1426-1439, (1989).

100. Van der Ziel, J. P., Tsang, W. T., Logan R. A., Mikulyak R. M., and Augustyniak W. M., "Sub Picosecond Pulses for Passively Mode-Locked GaAs Buried Optical Guide Semiconductor Laser", Applied Physics Letters, **39**, pp 525-527, (1981).

101. Clayton G. B., "Operational Amplifiers", 2nd edition, Butterworths, (1980).

102. Sze S. M., "Physics of Semiconductor Devices", 2nd edition, John Wiley and Sons, New York, (1981).

103. Johnson J. B., "Thermal agitation of electricity in conductors", Physical Review, **32**, pp 97-109, (1928).

104. Nyquist H., "Thermal agitation of electric charge in conductors", Physics Review, **32**, pp 110-113, (1928).

105. McKay K. G. and McAfee K. B., "Electron multiplication in silicon and germanium", Physical Review, **91**, pp 1079-1084, (1953).

106. McKay K. G., "Avalanche multiplication in silicon and germanium", Physical Review, **94**, pp 877, (1954).

107. Anderson L. K., McMullin P. G., D'Asaro L. A. and Goetzberger A., "Microwave photodiodes exhibiting microplasma-free carrier multiplication", App. Phys. Lett., **6**, pp 62-64, (1965).

108. Miller S. L., "Avalanche breakdown in germanium", Physics Review, **99**, pp 1234-1241, (1955).

109. McIntyre R. J., "Multiplication noise in uniform avalanche photodiodes", IEEE Trans. Electron Devices, **ED-13**, pp 164-168, (1966).

110. Cova S., Longoni A., Andreoni A. and Cubeddu R., "A Semiconductor Detector for Measuring Ultraweak Fluorescence Decays with 70 ps resolution", IEEE Journal of Quantum Electronics, **QE-19**, pp 630-634, (1983).

111. Cova S., Longoni A. and Ripamonti G., "Active Quenching and Gating Circuits for Single-Photon Avalanche Photodiodes (SPADs)", IEEE Transactions on Nuclear Science, **NS-29**, pp 599-601, (1982).

112. Cova S., Ghioni M., Lacaita A., Samori C. and Zappa F., "Avalanche photodiodes and quenching circuits for single-photon detection", Applied Optics, **35**, pp 1956-1976, (1996).

113. Chang K., "Handbook of microwave and optical components - Volume 2: Microwave solid state components", John-Wiley, New York, pp 240-241, (1990).

114. Krakauer S. M., "Harmonic generation, rectification, and lifetime evaluation with the step recovery diode", Proceedings of the IRE, **50**, pp 1665-1676, (1962).

115. "Comb generator simplifies multiplier design", Hewlett Packard application note, Number 983, Hewlett-Packard, Palo Alto, CA, USA, (1981).

116. "Harmonic generation using step recovery diodes and SRD modules", Hewlett Packard application note, Number 920, Hewlett-Packard, Palo Alto, CA, USA, (1969).

117. Hall S. H. R., "Shunt-mode harmonic generation using step recovery diodes", Microwave Journal, **10**, pp 69-78, (1967).

118. "How to get more output power from a comb generator module with the right bias resistance", Hewlett Packard application note, number 984, Hewlett-Packard, Palo Alto, CA, USA, (1981).

119. "Step Recovery diode doubler", Hewlett Packard application note, number 989, Hewlett-Packard, Palo Alto, CA, USA, (1981).

120. "Pulse and waveform generation with step recovery diodes", Hewlett Packard application note, AN 918, Hewlett-Packard, Palo Alto, CA, USA.

121. Bogart T. F., "Laplace Transforms - Theory and Experiments", John Wiley and Sons, New York, USA, (1983).

122. "HP Communications components, GaAs & silicon, designers catalogue", Hewlett Packard, USA, (1993).

123. "Harmonic Generation using Step Recovery Diodes and SRD modules", Hewlett Packard application note, AN 920, Hewlett-Packard, Palo Alto, CA, USA.

124. Firbank M., "The design, calibration and usage of a solid scattering and absorbing phantom for near infra-red spectroscopy", Ph.D. thesis, University of London, (1994).

125. Firbank M., Oda M. and Delpy D. T., "An improved design for a stable and reproducible phantom material for the use in near-infrared spectroscopy and imaging", Phys. Med. Biol., **40**, pp 955-961, (1995).

126. Jessop G. R., "VHF UHF manual", 4th edition, Chapter 4 - Receivers, Radio Society of Great Britain, pp 4.13, (1980).

127. Kirkby D. R. and Delpy D. T., "Measurement of tissue temporal point spread function (TPSF) by use of a cross-correlator with an avalanche photodiode detector", Proc. SPIE, **2389**, pp 190-197, (1995).

128. Edminister J. A., "Electric circuits - Schaum's outline series", McGraw-Hill book company, Ney York, UA, (1965).

129. Williams, A. B., "Electronic Filter Design Handbook", McGraw-Hill, USA, (1981).

130. Blinchhikoff H. J., "Low-Transient High-Pass Filters", IEEE Trans. Circuit Theory, pp 663-667, (November 1970).

131. Baxandall P. J., "Noise in Transistor Circuits: 1 - Mainly on fundamental noise concepts", Wireless World, pp 388-392, (November 1968).

132. Baxandall P. J., "Noise in Transistor Circuits: 2 Noise figure; Negative Feedback: Measurements", Wireless World, pp 454-459, (December 1968).

133. Yoo K. M., Das B. B. and Alfarno R. R., "Imaging of a translucent object hidden in a highly scattering medium from the early portion of the diffuse component of a transmitted ultrafast laser pulse", Optics Letters, **17**, pp 958-960, (1992).

134. van Staveren H. J., Moes C. J. M., van Marle J., Prahl S. A. and van Gemert J.C., "Light scattering of Intralipid-10% in the wavelength range of 400-1100 nm", Applied Optics, **30**, pp 4507-4514, (1991).

135. Flock S. T., Jacques S. L., Wilson B., Star W. M. and van Gemert M. J. C., "Optical Properties of Intralipid: A phantom medium for light propagation studies", Lasers in Surgery and Medicine, **12**, pp 510-519, (1992).

136. van Staveren H. J., Moes C. J. M., van Marle J., Prahl S. A. and van Gemert J.C., "Light scattering of Intralipid-10% in the wavelength range of 400-1100 nm", Applied Optics, **30**, pp 4507-4514, (1991).

137. Lovell A. T., Hebden J. C., Goldstone J. C. and Cope M., "Determination of the transport scattering coefficient of red blood cells", Proc. SPIE (in press).

138. Matcher S. J., Cope M., Delpy D. T., "Use of water absorption spectrum to quantify chromophore concentration changes in near-infrared spectroscopy", Phy. Med. Biol, **38**, pp 177-196, (1993).

139. Patterson M. S., Schwartz E. and Wilson B. C., "Quantitative reflectance spectrophotometry for the noninvasive measurement of photosensitizer concentration in tissue during photodynamic therapy", Proc. SPIE, **1065**, pp 115-122, (1989).

140. Madsen S. J., Patterson M. S. and Wilson B. C., "The use of india ink as an optical absorber in tissue simulating phantoms", Phys. Med. Biol., **37**, pp 985-993, (1992).

141. Hale G. M. and Querry M. R., "Optical constants of water in the 200-nm to 200-μm Wavelength Region", Applied Optics, **12**, pp 555-563, (1973).

142. Press W. H., Teukolsky S. A., Vetterling W. T. and Flannery B. P., "Numerical Recipes in C", 2nd edition, Cambridge University Press, (1992).

143. Kirkby D. R., "Monte Carlo Simulation of Laser Doppler Blood Flow Measurement through Thick Tissue Sections", M.Sc. thesis, University of London, (1993).

144. "i486 Microprocessor", data sheet, Intel Corporation, (1989).

145. Kirkby D. R. and Delpy D. T., "Parallel operation of Monte Carlo simulations on a diverse network of computers", Phys. Med. Biol., **42**, pp 1203-1208, (1997).

146. Press W. H., Teukolsky S. A., Vetterling W. T. and Flannery B. P., "Numerical Recipes in C", 1st edition, Cambridge University Press, (1988).

147. Wood D., "Random, Or Just Confused?", Transactor for the AMIGA, **1**, Issue 2, pp 38-46, (June 1988).

148. Turbo Pascal Compiler version 4.0, Borland International (UK) Ltd, 8 Pavilions, Ruscombe Business Park, Twyford, Berkshire, RG10 9NN.

149. Spiegel M. R., "Theory and Problems of Statistics", Schaums Outline Series, McGraw-Hill Book Company, New York, (1961).

150. Wolfram S., "The Mathematica book", 3rd edition, Wolfram Media/Cambridge University Press, (1996). See http://www.wolfram.com

151. Kirkby D. R. and Delpy D. T., "An optoelectronic cross-correlator using a gain modulated avalanche photodiode for measurement of the tissue temporal point spread function", Proc OSA, **TOPS 2**, pp 108-112, (1996).

152. Kirkby D. R. and Delpy D. T., " Measurement of tissue temporal point spread function (TPSF) by use of a gain-modulated avalanche photodiode detector", Phys. Med. Biol., **41**, pp 939-949, (1996).

153. David K. Harrison, private communication, (1999).