OPTICAL MEASUREMENT TECHNIQUES FOR CLINICAL ASSESSMENT OF CORNEAL PHYSIOLOGY

By

Jesse Hart Shaver

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Molecular Physiology and Biophysics

December, 2007

Nashville, Tennessee

Approved:

Dr. David W. Piston, Ph.D. – Advisor

Dr. Albert H. Beth, Ph.D. – Chair

Dr. Frederick R. Haselton, Ph.D.

Dr. Paul Sternberg, Jr., M.D.

Dr. David H. Wasserman, Ph.D.

To my parents James and Adele, and my loving wife Kelley -

ACKNOWLEDGEMENTS

This work was made possible through financial support from several entities, primarily the Vanderbilt Medical Scientist Training Program (MSTP), and the Vanderbilt Vision Research Center Training Grant (VVRC). Other sources of support for materials, travel, and training include NIBIB R21, Molecular Probes grant, Research to Prevent Blindness, and the American Society for Lasers in Medicine and Surgery.

NIH/NIBIB 5R21EB002382-02
NIH/NEI 5T32EY007135-14 (Vanderbilt Vision Research Center training grant)
NIH/NIGMS 5T32GM007347-28 (Vanderbilt Medical Scientist Training Program grant)

I would like to thank my dissertation advisor, Dr. David W. Piston, for allowing me to pursue this research, and supporting me at all stages. Dave has been a great mentor and a great friend. I would like to thank the members of my dissertation committee for their patience and thoughtful advice that I will strive to carry forward.

Dr. Albert H. Beth, Ph.D. (Chair) – Molecular Physiology and Biophysics
Dr. Frederick R. Haselton, Ph.D. – Biomedical Engineering
Dr. Hassane S. Mchaourab. Ph.D. – Molecular Physiology and Biophysics
Dr. Paul Sternberg, Jr., M.D. – Vanderbilt Eye Institute
Dr. David H. Wasserman, Ph.D. – Molecular Physiology and Biophysics

I would like to thank several academic instructors and advisors, including Dr. Jin Shen, Dr. K. Sam Wells, Dr. Robert Campbell, Dr. Terry Dermody, Dr. Jeff Schall, and Dr. Denis O'Day. I would also like to thank my family for their kindness and support. Finally, I would like to thank my wife Kelley, for the love and encouragement she has given me.

TABLE OF CONTENTS

Page

DEDICATION……………………………………………………………………………………………iii

ACKNOWLEDGEMENTS………………………………………………………………………….....iv

LIST OF FIGURES……………………………………………………………………………………vii

LIST OF EQUATIONS……………………………………………………………………………..ix

LIST OF ABBREVIATIONS AND SYMBOLS………………………………………………………... x

LIST OF FUNDAMENTAL CONSTANTS AND PARAMETERS……………………………………….xi

Chapter

I. INTRODUCTION…………………………………………………………………………… 1
- a. Review of corneal anatomy and physiology………………………………………..… 3
- b. Normal variations in corneal thickness and hydration……………………………...…5
- c. Clinical relevance of corneal hydration, refractive index, and thickness……………..… 8
  - i. Glaucoma risk assessment and central corneal thickness……………………….8
  - ii. Corneal measurement for refractive surgical procedures………………………11

II. SYNTHETIC-APERTURE CONFOCAL PACHYMETRY…………………….……….....14
- a. Pachymetry as a major diagnostic modality…...………………………………….....14
- b. Review of clinical and research pachymetry systems……………………………16
  - i. Ultrasound………………………………………………………………… 16
  - ii. Optical-coherence tomography (OCT)…………………………………………17
  - iii. Optical low-coherence reflectometry (OLCR) and partial coherence interferometry (PCI)……………………………………………18
  - iv. Topography techniques applied to pachymetry………………………………..19
  - v. Scheimpflug pachymetry………………………………………………………19
  - vi. Confocal microscopy used for pachymetry……………..…………………..20
- c. Confocal optical systems for optical thickness measurement……………………………20
  - i. Review of traditional confocal optical systems…………………………………20
  - ii. Synthetic confocal systems: The hologram laser…………………………….... 22
- d. Synthetic-aperture confocal pachymeter system design…………………………………27
  - i. Optical system design………………………………………………………...27
  - ii. Scanning and drive mechanism design……………………………………..… 28
  - iii. Data processing system design…………………………………………….....34
  - iv. Performance and cost…………………………………………………….… 38
  - v. Integration with other diagnostic modalities………………………………….. 38

III. HYDRATION-REFRACTIVE INDEX-THICKNESS RELATIONSHIP………………....…40
- a. Review of previous literature and models……………………………………………..40
- b. Experimental methods………………………………………………………...… 51
  - i. Samples…………………………………………………………………… 51
  - ii. Hydration perturbation, drying protocol, and measurement…………………... 51
  - iii. Refractive index measurement………………………………………………...52
- c. Results and discussion……………………………………………………………………54
  - i. Corneal thickness as a function of hydration……………………………...… 54
  - ii. Refined Gladstone-Dale model of refractive index and hydration……………. 55

IV. RAMAN SPECTROSCOPY FOR HYDRATION MEASUREMENT……………….……..… 61
  a. Review of Raman scattering physics………………………………………………. 61
  b. Raman spectrometer design variations…………………………………………….. 63
    i. Excitation sources…………………………………………………………………63
    ii. Sampling and collection optics……………………………………………….... 64
    iii. Dispersion mechanisms………………………………………………………..64
    iv. Filters…………………………………………………………………...……65
    v. Detectors………………………………………………………………………..66
  c. Quantitative Raman spectrometry…………………………………………………..74
    i. Sources of variation and error in quantitative Raman spectroscopy………….... 74
    ii. Use of internal standards………………………………………………………75
    iii. Raman spectra of mixtures as linear superpositions of component spectra…...75
    iv. The OH-stretch Raman bands of water……………………………………….. 75
  d. Review of previous systems for Raman tissue hydration measurement………………....77
    i. Dermatologic systems……………………………………………………...… 77
    ii. Ophthalmic systems…………………………………………………………..... 77
  e. System design………………………………………………………………………78
    i. Excitation source selection……………………………………………………...78
    ii. Optical configuration……………………………………………………………...79
    iii. Detection system……………………………………………………………… 79
    iv. Signal recovery system………………………………………………………...80
    v. Performance………………………………………………………………………80
    vi. Areas for future improvement………………………………………………… 83

V. SYNTHETIC CHOPPING AND OPTIMAL PHASE-SENSITIVE DETECTION……………...85
  a. Noise, interference, cross-talk, offset, and drift…………………………………….… 85
    i. Photon and current shot noise………………………………………………….. 88
    ii. Thermal (Johnson) noise………………………………………………………90
    iii. Noise with a 1/f power spectral density in photodetection…………………… 92
    iv. Quantization noise and ADC errors……………………………………..…. 92
    v. System noise due to phase jitter and frequency instability……………………..93
    vi. Interference and cross-talk……………………………………………….… 93
  b. Overview of signal recovery techniques………………………………………………95
    i. Introduction…………………………………………………………..…… 95
    ii. Coherent averaging……………………………………………….……….. 95
    iii. Lock-in amplification…………………………………………….…………96
  c. Optimal methods for lock-in signal recovery…………………………………………… 97
    i. Wilmshurst analysis on optimal signal recovery………………………………. 97
    ii. Doering-Harvey optimal linear phase-sensitive algorithm…………………… 97
    iii. Zhang-Soukas rules for optimal phase-sensitive detection……………………98
  d. Amplitude-modulated signal waveforms…………………………………………...100
    i. Review of amplitude modulators and their performance limits……………….. 100
    ii. Random errors and deterministic errors……………………………….……101
    iii. Coherent interference and system crosstalk…………………………………...102
    iv. Waveform measurement techniques………………………………………….. 102
  e. Noise spectrum measurement…….…………..……………………………………… 102
  f. Synthetic chopping……………………………………………………………………. 103
    i. Description of the method……………………………………………………... 103
    ii. Implementation…………………………………………………………………110
    iii. Cost and performance of a complete system……………………….………110

VI. SUMMARY AND CONCLUSIONS…………………………………………………..... 112

BIBLIOGRAPHY AND REFERENCES

LIST OF FIGURES

Page

Figure 1 - The confocal optical system uses an aligned spatial filter to reject out-of-plane light……………. 21

Figure 2 - The hologram laser integrates many elements of a confocal system into a single integrated component……………………………………………………………...…23

Figure 3 - The hologram laser optical path is shown…………………………………………………..……24

Figure 4 - A micrograph of the detector photodiode array shows the five detector segments……………..… 25

Figure 5 - The focal error signal results from the subtraction of signals from two photodiodes in the array……………………………………………………………... 26

Figure 6 - The optical components used in the prototype synthetic aperture confocal pachymeter………..… 28

Figure 7 - A scanner and drive mechanism based on spiral flexures, which may be used in future clinical devices………………………………………………………...30

Figure 8 - A crank drive scanner system prototype……………………………………………………....32

Figure 9 - A Scotch yoke drive scanner system prototype……………………………………………….... 33

Figure 10 - The signal processing system for confocal pachymetry…………………………………….. 35

Figure 11 - The Cauchy parameters, taken from literature, for cornea, aqueous humor, and dry gelatin…..…48

Figure 12 - Dispersion curves for ocular media and Gladstone-Dale model components…………………... 49

Figure 13 - Detailed view of corneal dispersion and Gladstone-Dale models…………………………..…… 50

Figure 14 - Schematic of measurement of optical thickness, geometrical thickness, and mass during hydration……………………………………………………………………..…. 53

Figure 15 - The relationship between corneal thickness and hydration……………………………....… 54

Figure 16 - The refractive index of Rabbit cornea at 780nm as a function of tissue hydration H……………. 56

Figure 17 - The refractive index of Ox cornea, with a comparison of models………………………..……57

Figure 18 - Energy level diagram of Raman and Rayleigh scattering…………………………………………62

Figure 19 - Photoemission in a photocathode……………………………………………………………..…69

Figure 20 - Schematic of a photomultiplier tube……………………………………………………………70

Figure 21 - Modes of water molecular vibrations………………………………………………………….. 76

Figure 22 - A photodiode transimpedance amplifier………………………………………………………. 79

Figure 23 - The channel response matrix…………………………………………………………………... 81

Figure 24 - Signal channel vector analysis………………………………………………………………….81

Figure 25 - Optical bandpass filters and Raman spectral bands……………………………………………82

Figure 26 - System noise floor with increased measurement time………………………………………….... 83

Figure 27 - The Raman hydration measurement system………………………………………………..….. 84

Figure 28 - A typical noise power spectrum………………………………………………………………...87

Figure 29 - Photocurrent shot noise limits for signal to noise………………………………………………90

Figure 30 - A schematic lock-in amplifier system…………………………………………………………..96

Figure 31 - The synthetic reference signal allows direct measurement of the modulation waveform……….. 104

Figure 32 - Characteristics of the modulation waveform depend upon beam and aperture characteristics..… 104

Figure 33 - Measured modulation waveforms incorporate amplifier bandwidth and phase response………... 105

Figure 34 - Coherent averaging allows for the detection and removal of coherent interference…….…….. 106

Figure 35 - The Vernier principle of resolution enhancement…………………………………...……… 107

Figure 36 - The synthetic chopper…………………………………………………...………………… 110

LIST OF EQUATIONS

Page

Equation 1 - The definition of the hydration parameter H………………………………………..……….. 6

Equation 2 - The Hedbys-Mishima equation relating corneal thickness to hydration……..………………… 41

Equation 3 - The general Gladstone-Dale equation…………………………………………………..……..42

Equation 4 - The Gladstone-Dale equation rearranged in terms of a component index………………………42

Equation 5 - The Gladstone-Dale equation for corneal index with
hydrated fibrils and hydrated matrix components…………………………………………………… 43

Equation 6 - The Gladstone-Dale equation for hydrated fibril index
with water and dry collagen components…………………………………………………………… 43

Equation 7 - The Gladstone-Dale equation for hydrated matrix index
with water and dry matrix components………………………………………………...……… 43

Equation 8 - The Gladstone-Dale equation for hydrated matrix index,
in terms of water and corneal refractive indices……………………………….…………………. 43

Equation 9 - The Meek-Dennis-Khan Gladstone-Dale equation for hydration-dependent corneal index……. 44

Equation 10 - A Gladstone-Dale equation for hydration-dependent corneal index using
hydration-dependent volume fractions and indices for solvent,
dry collagen, and dry matrix components………………………………………………………… 46

Equation 11 - The form of the Cauchy dispersion equation……………………….………………………..47

Equation 12 - The form of the Sellmeier dispersion equation………………..………………………………47

Equation 13 - The intensity of Raman scattering…………………………….…………………………….. 63

Equation 14 - The grating equation………………………………………..…………………………….. 65

Equation 15 - The linear system of channel response………………………..…………………………….. 80

Equation 16 - The Schottky equation for shot noise…………………...……………………………………88

Equation 17 - The Poisson equation for uncorrelated discrete event probability density……………..……… 89

Equation 18 - The Johnson equation for thermal voltage noise………………………………….…………… 91

Equation 19 - The Johnson equation for equivalent thermal current noise…………………...……………… 91

LIST OF ABBREVIATIONS AND SYMBOLS

ADC	analog to digital converter
ADU	analog to digital unit
APD	avalanche photodiode
CCT	central corneal thickness
CPR	cycles per revolution, used to specify angular encoders
DSP	digital signal processor
HeNe	helium-neon laser
ECM	extracellular matrix
ENOB	effective number of bits
FWHM	full width at half-maximum
GAG	glycosaminoglycan
GAT	Goldmann applanation tonometer
IAPWS	International Association for the Properties of Water and Steam
IOP	intra-ocular pressure
IUPAC	International Union of Pure and Applied Chemistry
LASIK	laser-assisted in situ keratomileusis
MTTF	mean time to failure
NEA	negative electron affinity
NEP	noise equivalent power
OHTS	Ocular Hypertension Treatment Study
OLCR	Optical Low-Coherence Reflectometry
PCI	Partial Coherence Interferometry
PIN	positive-intrinsic-negative
PMT	photomultiplier tube
PRK	photorefractive keratectomy
SiPM	silicon photomultiplier

Q	resonator quality factor, equal to central frequency divided by bandwidth
a	polarizability of molecular bond electrons
I	intensity
P_{in}	incident power (W)
?	frequency
T_g	geometric corneal thickness
T_o	optical (apparent) corneal thickness
H	corneal hydration index
a	species-dependent hydration scale parameter in the Hedbys-Mishima eqn.
b	species-dependent hydration offset parameter in the Hedbys-Mishima eqn.
L	corneal thickness (in millimeters) from the Hedbys-Misima eqn.
n_c	refractive index of dry collagen
n_w	refractive index of solvent (water, under specified conditions)
n_g	refractive index of dry gelatin
n_p	refractive index of dry extrafibrillar matrix (proteins)
n_f	refractive index of hydrated fibrils
n_e	refractive index of hydrated extrafibrillar matrix
n_{stroma}	refractive index of stroma (at physiological hydration (H=3.2) if not otherwise specified)
$f_{c/f}$	volume fraction of dry collagen in a hydrated fibril, equal to $f_{c/s} / f_{f/s}$
$f_{c/s}$	volume fraction of dry fibrillar material (collagen) in the stroma
$f_{f/s}$	volume fraction of hydrated fibrillar material in the stroma
$f_{p/s}$	volume fraction of dry extrafibrillar material in the stroma
$f_{e/s}$	volume fraction of hydrated extrafibrillar material in the stroma
$f_{iw/s}$	volume fraction of intrafibrillar solvent (water) in the stroma
$f_{ew/s}$	volume fraction of extrafibrillar solvent (water) in the stroma
$f_{w/f}$	volume fraction of solvent (water) in a fibril

LIST OF FUNDAMENTAL CONSTANTS AND PARAMETERS

Planck constant h
$6.6260755 \cdot 10^{-34}$ J·s
$h / (2\text{p}) = 1.05457266 \cdot 10^{-34}$ J·s
Boltzmann constant k_B
$1.380658 \cdot 10^{-23}$ J/K = $8.617385 \cdot 10^{-5}$ eV/K
Elementary charge e
$1.60217733 \cdot 10^{-19}$ C
Avogadro (Loschmidt) number N_A
$6.0221367 \cdot 10^{23}$ particles/mol
Speed of light c
$2.99792458 \cdot 10^{8}$ m/s
Permeability of vacuum μ_0
$\mu_0 = 4\text{p} \cdot 10^{-7}$ $\text{T}^2 \cdot \text{m}^3/\text{J}$
$12.566370614 \cdot 10^{-7}$ $\text{T}^2 \cdot \text{m}^3/\text{J}$
Permittivity of vacuum e_0
$\text{e}_0 = 1 / (\mu_0 c^2)$
$8.854187817 \cdot 10^{-12}$ $\text{C}^2/\text{J} \cdot \text{m}$
Fine structure constant a
1 / 137.0359895
Electron rest mass m_e
$9.1093897 \cdot 10^{-31}$ kg
Proton rest mass m_p
$1.6726231 \cdot 10^{-27}$ kg
Neutron rest mass m_n
$1.6749286 \cdot 10^{-27}$ kg
Bohr magneton μ_B
$\mu_B = e h / (4\text{p} m_e)$
$9.2740154 \cdot 10^{-24}$ J/T
Nuclear magneton μ_N
$\mu_N = e h / (4\text{p} m_p)$
$5.0507866 \cdot 10^{-27}$ J/T
Free electron g factor g_e
2.002319304386
Free electron gyromagnetic ratio $?_e$
$?_e = 2\text{p} g_e \mu_B / h$
$1.7608592 \cdot 10^{11}$ 1/s·T
$?_e / (2\text{p}) = 28.024944$ GHz/T

Electron magnetic moment μ_e
$\mu_e = -(1/2)\, g_e \mu_B$
$-9.2847701 \cdot 10^{-24}$ J/T
Proton gyromagnetic ratio (H_2O) $?_p$
$2.67515255 \cdot 10^{8}$ 1/s·T
$?_p / (2\text{p}) = 42.576375$ MHz/T
Proton magnetic moment μ_p
$1.41060761 \cdot 10^{-26}$ J/T
Proton-electron ratios
$m_p / m_e = 1836.152701$
$\mu_e / \mu_p = 658.2106881$
$?_e / ?_p = 658.2275841$ (protons in water)
Charge-to-mass ratio for the electron e / m_e
$1.75880 \cdot 10^{11}$ C/kg
Atomic mass unit amu
$1.66054 \cdot 10^{-27}$ kg
Bohr radius a_0
$5.29177 \cdot 10^{-11}$ m
Electron radius r_e
$2.81792 \cdot 10^{-15}$ m
Gas constant R
$R = N_A k_B$
8.31451 $\text{m}^2 \cdot \text{kg}/\text{s}^2 \cdot \text{K} \cdot \text{mol}$
Molar volume V_{mol}
22.41383 m^3/kmol
Faraday constant F
$F = N_A e$
$9.64846 \cdot 10^{4}$ C/mol
Proton g factor (Landé factor) g_H
5.585
Gravitational constant G
$(6.673 +- 0.010) \cdot 10^{-11}$ $\text{m}^3/\text{kg} \cdot \text{s}^2$ (CODATA)
Acceleration due to gravity g
9.80665 m/s^2
Compton wavelength of the electron $?_c$
$?_c = h / (m_e c)$

CHAPTER I

INTRODUCTION

The cornea is a transparent avascular tissue window that functions as the main focusing element of the eye, providing 70% of the ocular refractive power. The biophysical properties of the cornea are of tremendous importance to normal vision. These properties include mechanical strength, resilience to deformation, and optical transparency. Surgical manipulation of the refractive properties of the cornea is now common. Refractive surgery has created a new demand for accurate clinical measurements of parameters such as hydration and transparency, as well as corneal thickness. Additionally, a growing body of clinical research has determined that accurate knowledge of central corneal thickness is useful in predicting the risk of vision loss from glaucoma, particularly in patient populations with ocular hypertension. This dissertation addresses new and refined application of optical measurement techniques to meet these clinical demands, with an emphasis on techniques that are clinically feasible in terms of cost, reliability, and safety.

Presented are new biophysical data, including the results of optical measurements of the hydration-refractive index-thickness relationship over a wider range of hydration than previously-published data, taken at 780nm. This data is used to refine existing Gladstone-Dale models of corneal refractive index as a function of hydration. New consideration is given to the wavelength-dependence of model components. The revised Gladstone-Dale models are able to approximate previously-published Cauchy equations for dispersion in the cornea at physiological hydration, as well as previously-published measurements of refractive index as a function of hydration.

The new models, built expressly with ocular dispersion in mind, are significant for several reasons. Accurate optical measurement of corneal thickness is predicated on accurate knowledge of the corneal refractive index at the wavelength of measurement. Biophysical models of corneal scattering and transparency use refractive index parameters of stromal constituents as inputs, which must be accurate. Finally, the refractive index of the cornea is used in optical eye models, which are used clinically in corneal refractive surgery and cataract surgery.

In addition to improving basic biophysical models, another goal of this research is practical innovation in the area of clinical metrology, specifically the measurement of several biophysical parameters of the cornea. A new implementation of optical pachymetry is described, based on synthetic-aperture confocal microscopy. A new system for corneal hydration measurement by Raman spectral analysis is also presented, using low excitation laser power levels and low-cost room-temperature detectors. Low excitation power and low cost are both important for clinical applicability. This system uses a new optical signal recovery technique, synthetic chopping. While the prototype system is not currently capable of measuring corneal hydration in a reasonable amount of time, straightforward improvements in the optics and detectors are possible, which may enable eventual clinical use of a similar system.

Some of the principles, methods, techniques, and devices developed in the course of this research are of broad utility, particularly those dealing with low-cost, high-performance signal recovery using modulation and synchronous detection. Thus, this technical portion of the research may find practical use outside of medical instrumentation, especially in low-cost devices.

Because of the focus on instrumentation development, a fair amount of the material presented is technical and practical in nature. This work is a blend of basic and applied science. Well-designed modern laboratory and clinical instruments will produce reliable results, independent of the user. Indeed, the exceptional performance of the many instruments now available can tend to diminish the importance of instrument development and characterization for the typical user. However, experimental data is difficult to interpret without a working knowledge of the measurement principle used to gather the data, and its limitations. The instrument is an important part of the experiment, and instrument development is an important part of science.

The escalating costs of healthcare provide a basic motivation for this research. Technological innovation impacts the cost of healthcare directly and indirectly. Expensive diagnostic and therapeutic technologies offer demonstrably-superior performance to low-cost alternatives, but the incremental increase in performance is often accompanied by an order-of-magnitude increase in cost. Each stage of improved technology also brings a transition cost associated with training, regulatory compliance, marketing (often to the patient as well as the provider), sales, and product liability insurance. At the same time, failure to adopt a new technology can lead to malpractice liability cost for the provider. Despite all of these costs, a given technology may still achieve a favorable system-wide cost-benefit ratio if the benefits to the patient are substantial (Bodenheimer, 2005). In light of these observations, a concerted effort to reduce the costs of instrumentation and improve diagnostic efficiency should be a very high priority for medical device developers.

a. Review of corneal anatomy and physiology

As the main refractive optical element of the eye, the cornea must maintain its curvature and transparency over an entire lifespan. The human cornea is described as having five histological layers. This classical description tends to ignore the vital importance of the three-layer fluid tear film on the corneal surface, which is removed during histological preparation.

The three tear film layers are the lipid layer (secreted by meibomian glands), the aqueous layer (secreted by lacrimal glands), and the mucous layer (produced by goblet cells in the conjunctiva). The thickness of the tear film is approximately 3 to 7 microns (Fogt et al, 1998, Wang et al, 2003), and may vary by species. The tear film is optically important, as it forms a smooth refractive surface and slows epithelial cellular dehydration. The tear film also serves to protect against infectious agents through the action of antimicrobial enzymes, such as lactoferrin and lysozyme. The tonicity of the tear film fluctuates diurnally, and is higher when the eye is open (Friedman, 1972). The tear film is thus a very important structure for corneal physiology.

The human corneal epithelium is 40 to 50 microns thick, composed of approximately six layers of cells located beneath the tear film (Freund et al, 1986). Based on some clinical evidence, the cells are thought to originate in the periphery of the cornea, the corneal limbus. The cells in the basal periphery are columnar, and the cells of the central surface are squamous. The cells eventually die and slough away under the constant sweep of the eyelid, with a lifespan measured in days. The smoothness of the epithelial layer is crucial to clear imaging, as it forms the substrate for the air-tear film interface, the primary refractive surface. Injuries, infections, or disruption of epithelial hydration can compromise surface quality. The epithelial cells have a sub-micron basement membrane, distinct from the more prominent Bowman's layer.

The second histological layer, Bowman's layer, is a dense collagen and ground-substance layer, 10 microns thick, which protects the stroma. It is not present in all mammals. Compromise of this layer has been linked to scarring and opacities in the stroma, perhaps as a result of exposure to epithelial chemokines. Bowman's membrane expresses genes COL1A1, COL1A2, COL6A1, COL6A2, COL6A3, and COL12A1 (Ihanamäki et al, 2004).

The corneal stroma is the thickest layer, comprising 90% of the full thickness. It is comprised of highly-ordered, uniform-diameter collagen fibrils embedded in a proteoglycan and glycoprotein ground substance (Ihanamäki et al, 2004). A group of flat quiescent cells, the keratocytes, populate the stroma. The cells are interconnected, forming a sparse network. These cells make up approximately 10-15% of the stromal volume and 10-15% of the water mass (Kaye, 1969, Leonard & Meek, 1997, Meek et al, 2003), although some estimates place the keratocyte volume at only 3-5% (Freund et al, 1986). The fibrils are arranged into sub-layers called lamellae, each about 2 microns thick (Freund et al, 1986). The fibrils in each lamella are parallel to each other. The fibrils in adjacent lamellae run in different directions, similar to the alternating wood grain in layers of plywood. This gives a distinctive appearance in cross section EM micrographs.

The visual line-of-sight is perpendicular to the fibers of all lamellae. The majority of stromal collagen is type I in fibrillar form, but there are smaller amounts of type V, as well as some type III and type IV, which has extensive globular domains (Leonard & Meek, 1997). The full survey of expressed stromal collagen genes include COL1A1, COL1A2, COL2A1, COL3A1, COL5A1, COL5A2, COL5A3, COL5A4, COL6A1, COL6A2, COL6A3, COL7A1, COL12A1, COL13A1 and COL24A1 (Ihanamäki et al, 2004).

The precise arrangement of collagen fibrils in the stroma is a major topic in corneal biophysics. Since the work of Maurice in the 1950's, several theories have been advanced to explain the transparency of the stroma, summarized by Meek and coworkers in *Eye* (Meek et al, 2003). At physiological hydration, the stroma is very transparent, with very little scattering. Since the components of the stroma (collagen, water, proteoglycans, and glycoproteins) have different refractive indices, this lack of scattering is surprising. To explain this effect, it is postulated that individual fibrils scatter, but spatial correlation between fibrils causes destructive optical interference for light scattered away from the forward direction, and constructive interference for light scattered in the forward direction, resulting in minimal deleterious effects for vision (Freund et al, 1986, Freegard, 1997). Shifts in stromal hydration alter the interfibril spacing, and disrupt these interference effects. This is thought to be the most important cause for the increase in scatter seen in corneal edema (Meek et al, 2003).

Descemet's membrane is a thin layer found beneath the stroma, and serves as a basement membrane for the monolayer of endothelial cells. Expressed collagen genes include COL4A1, COL4A2, COL4A3, COL4A4, COL4A5, COL4A6, COL6A1, COL6A2, COL6A3, COL8A1, and COL8A2 (Ihanamäki et al, 2004).

The corneal endothelium is the posterior and final layer of the cornea, and serves the important function of regulating water and solute transport from the aqueous humor into the stroma. The endothelial layer is a hexagonal cellular monolayer, and is less than 10 microns thick. The cells maintain the stroma in a slightly-dehydrated state, actively transporting water against the osmotic pressure gradient, while allowing nutrients such as glucose to cross into the stroma. Maintenance of homeostatic corneal deturgescence requires significant energy and feedback, as exogenous environmental factors can affect aspects of the water flux rates (Friedman, 1972). The corneal endothelium is derived from the neural crest. The endothelial cell density is 2500-3200 cells per square millimeter. Injury to the endothelium can occur during cataract surgery and other intraocular surgeries. The cells are not replaced to any great extent, and injury to the endothelium can result in irreversible edematous swelling of the cornea. These patients may be treated with corneal transplant surgery.

b. Normal variations in corneal thickness and hydration

Central corneal thickness (CCT) is defined as the geometrical thickness (perpendicular to the outer corneal surface) at a defined central reference point on the corneal, typically the line of sight (Ehlers & Hjortdal, 2004). The tissue hydration parameter (H) is defined as the mass ratio of the water mass divided by the dry cornea mass.

$$H = \frac{mass_{water}}{mass_{dry}} = \frac{mass_{wet} - mass_{dry}}{mass_{dry}}$$

Equation 1

A meta-analysis of over 300 published studies on CCT (Doughty & Zaman, 2000) found mean reported value for central corneal thickness in normal adults measured during daytime is 534 microns, with a standard deviation of 31 microns and a nearly Gaussian distribution. The range within one standard deviation of the mean is therefore 503 to 565 microns, while two standard deviations cover the range from 472 and 596 microns. Mean corneal thickness values are thought to differ significantly by race (Kniestedt et al, 2006). It is clear that some ophthalmic conditions are also associated with thicker and thinner mean values.

The measured normal mean value for the central corneal thickness also depends upon the measurement modality. Different classes of pachymeters, based on different measurement principles, have been shown to give different values for the same patient in side-by-side comparisons. Modality-dependence is also seen in meta-analysis, which combine many studies that use the same modality. Doughty and Zaman noted that there is a significant difference in the population mean corneal thickness as measured with ultrasound (544 microns +/- 34 microns SD) compared to slit-lamp-based optical pachymetry (530 microns +/- 29 microns SD). In other words, the mean corneal thickness had a modality-dependent difference of about 15 microns. While this observation may be a consequence of actual bias in the instruments, it may also reflect differing characteristics of the patient populations in each group of studies in the meta-analysis. Most comparative studies find that ultrasound measures a larger value than the observer-dependent slit-lamp optical pachymetry method. A more detailed discussion of pachymetry technologies is presented in Chapter II of this thesis, which describes a new optical pachymeter.

In general, corneal thickness is thought to be relatively stable in an individual after two years of age, with some researchers claiming a slight thinning over the individual’s lifetime

(Ehlers & Hjortdal, 2004). In the absence of surgery, pathology, or degeneration, corneal thickness is regarded as a fixed biometric entity (Doughty & Zaman, 2000). This view is based on a belief that significant changes to the collagen and interfibrillar matrix do not occur in the absence of disease. However, many studies have determined that there are normal physiological fluctuations in individual corneas measured at different times of day. Specifically, the cornea is thought to thicken during sleep, and thin upon awakening, due mainly to hydration changes. Although some researchers have estimated the magnitude of the fluctuation is as much as 11%, a more common finding is a diurnal fluctuation with peak-to-peak magnitude of ~6% (Doughty & Zaman, 2000, Manchester, 1970, Harper et al, 1996). Thus, a 534-micron "average" cornea might oscillate from ~517 microns (near bedtime) to ~548 microns (at awakening). The cause of this thickness change is thought to be normal fluctuations in corneal hydration that occur during the sleep-wake cycle. These diurnal fluctuations are not necessarily easy to detect, and are not universally reported in all pachymetry studies. Presumably, since the magnitude of hydration-driven fluctuations is of the same order as the standard deviation of a normal population of individuals, changes due to hydration alone could easily cause an individual near the margin of the normal CCT range to be categorized as normal or abnormal. Likewise, the ability to accurately measure corneal hydration could be used to more accurately gauge (and account for) hydration-dependent effects on corneal thickness. The realistic meaningful limit of corneal thickness measurement precision and reporting is about 1 micron (Ehlers & Hjortdal, 2004), based on the observation that blinking can change the tear film by a few microns, and that the posterior corneal limit is defined by a textured endothelial surface with at least a few microns of surface variation.

Iatrogenic mechanisms of central corneal thickness alteration, both permanent and temporary, include laser ablation, mechanical compression during surgeries, drying due to prolonged eye opening (Dayanir et al, 2004), atropine sulphate ointment administration, saline or complex tropicamide eye drop administration (Gao et al, 2006), topical oxybuprocain anesthetic administration (Asensio et al, 2003), topical proparacaine administration, and long-term contact lens wear. Saline eye drops are thought to influence corneal thickness and hydration by altering the tonicity of the tear film, while various drugs may have an effect on the endothelial cells and their ability to pump fluid.

Since anesthetic eyedrops are administered prior to use of the contact ultrasound pachymeter, the effects of the anesthetic (typically reported as 10-15 microns of hydration-related thickness increase) add a complex source of variability to the measurement. Furthermore, mechanical effects, such as probe indentation, are also problematic when multiple measurement attempts are necessary.

c. Clinical relevance of corneal hydration, refractive index, and thickness

Corneal hydration, refractive index, and thickness are three interrelated biophysical parameters that have direct relevance in clinical ophthalmology. They are important in the development of general models of normal and pathological physiology. They are also important parameters in the care of individual patients.

i. Glaucoma risk assessment and central corneal thickness

The interest in accurate measurement of central corneal thickness has recently increased, due to findings of major clinical studies dealing with glaucoma and ocular hypertension. The Ocular Hypertension Treatment Study (OHTS) was a large multi-center clinical trial which recruited patients with normal visual fields, normal optic disc exams, and untreated elevated intraocular pressure as measured by Goldman applanation tonometry (GAT). The patients were randomized into treatment and control groups, and followed for several years. A major finding of this study was that IOP-lowering medications reduced the incidence of glaucomatous visual field loss in this patient group. A surprise finding was that central corneal thickness was a strong prospective predictor of glaucoma risk in this patient population, representing a large independent risk factor in a multivariate risk model (Gordon et al, 2002).

Several ocular factors were examined for their value in predicting glaucoma risk in patients with ocular hypertension. In the group that did not develop glaucoma (N=1493), the mean central corneal thickness was 574.3 microns (SD 37.8 microns). In the group that did develop glaucoma (N=125), the mean central corneal thickness was 553 microns (SD 38.8 microns).

Having a CCT in the bottom third of the measured OHTS dataset (< 555 microns) increased a patient's risk of progression to glaucoma by a factor of three compared to those with a CCT of 588 microns or more, independent of intraocular pressure and other study variables. This increased risk has been extensively analyzed, and verified in further studies (Gordon et al, 2002, Brandt, 2004, Brandt et al, 2001, Medeiros et al, 2003). The mean central corneal thickness values for non-progressing ocular hypertension patients are much higher than the accepted average CCT for a normal population. The mechanism linking CCT to increased glaucoma risk remains unclear.

Glaucoma is a group of eye disease characterized by an increased in intraocular pressure (IOP) which causes pathological changes in the optic disk over time, leading to visual field deficits. It is estimated that 3 million people in the United States have glaucoma, with over 120,000 blind as a result. The term "glaukoma" has been associated with a variety of visual disorders of various description and etiology, including cataract, since the time of the ancient Greeks. It was recognized by Richard Bannister in 1622 that a certain pattern of blindness was associated with a firm eye, measured at that early date by manual digital tonometry. However, the term "glaucoma" was applied to various inflammatory visual disturbances until the early 19th century, when the work of Dr. Antoine-Pierre Demours and Dr. G.J. Guthrie linked the characteristic "firm eye" to the diagnosis. Dr. William McKenzie, author of a widely-read ophthalmology textbook, cited increased ocular tension in his descriptions of both chronic and acute glaucoma. The eminent Dutch ophthalmologist and visual physiologist Franciscus Cornelius Donders described a process of increased ocular tension without inflammatory symptoms as "simple glaucoma" in 1862. While the imprecision of medical terms and definitions seem quaint from the modern perspective, there is still room for confusion in current diagnostic terminology.

Glaucoma is now defined as a diverse group of eye conditions which cause characteristic progressive damage to the optic nerve. In most cases, this damage is associated with increased intraocular pressure, although a definite pressure-dependent mechanism and causal relationship is not universally agreed upon for all types of glaucoma. The diagnosis of Normal Tension Glaucoma reflects the notion that some patients have optic nerve damage and visual losses similar to other types of glaucoma, but without the characteristic increase in intraocular pressure above the threshold limit of 21 mmHg. One theory is that NTG patients have some trait which makes the optic nerve susceptible to damage at normal pressure levels, but under this theory, the linking relationship between pressure and damage seems unnecessary. Such a structural defect may result in damage irrespective of pressure. In that case, the definition need not refer to abnormal pressure once the precise nature of the defect is identified. The details of the link between pressure and damage are relevant when discussing central corneal thickness, because some research indicates that a thin cornea is indicative of a thin (compliant) lamina cribrosa, and thus, patients with thin corneas may experience proportionately greater tissue deflections and resultant nerve damage for a given IOP elevation. An inverse correlation between CCT and optic disc area and deformation was recently observed (Pakravan et al, 2007). The authors argue that thin corneas indicate larger, more deformable discs, and thus a greater propensity for nerve injury.

The authors of the OHTS initially theorized that CCT is important in glaucoma risk assessment because corneal thickness affects the accuracy of IOP measurement by Goldmann applanation tonometry. The Goldmann applanation tonometer measures pressure indirectly, by measuring the force necessary to flatten a 3.06 mm-diameter circular zone of the cornea. Goldman tonometry was introduced in the 1950's and was rapidly accepted for measurement of IOP. The instrument was inexpensive, simple to understand, and was compatible with existing slit lamp examination practices. The results of GAT are reproducible, but there are several potential sources of error, highlighted by Goldmann himself (Goldmann & Schmidt, 1957). Among these, it was noted that the accuracy would be reduced for corneas with mean thicknesses that differed from the average value, which Goldmann estimated as 500 microns. At the time, there was little published information on the variability of corneal thickness among individuals, and this factor was deemed unimportant to overall accuracy. Numerous studies have found a significant relationship between corneal thickness and measured IOP using the Goldmann apparatus. Corneal hydration is also important for some types of tonometry, as it influences the mechanical rigidity of the tissue (Kniestedt et al, 2005). However, a widely-accepted algorithm for correction of IOP values based on a central corneal thickness measurement has not yet been developed. In a meta-analysis of 133 datasets including IOP and CCT measurements, it was determined that a 10% difference in CCT

(~50 microns) could cause a 3.4 +/- .9 mmHg difference in IOP (Doughty & Zaman, 2000). One early study related GAT measurements to direct cannulation measurements, and determined that IOP estimates were in error by as much as 7 mmHg for every 100 micron deviation from a mean value of 520 microns (Ehlers et al, 1975). In the OHTS, 25% of the recruits had CCT values greater than 600 microns. Using the Ehlers correction formula, it is possible that as many as half of the OHTS recruits had true IOP values of less than 21 mmHg. The OHTS dataset has been reexamined for correlation of glaucoma risk with corrected IOP, using the measured thickness. These efforts have not shown a strong correlation between corrected IOP and glaucoma progression risk, which the authors attribute to the narrow range of IOP in the inclusion criteria. Central corneal thickness may be a truly independent risk factor for vision loss.

Regardless of the mechanism linking CCT and glaucoma risk, improvements are needed in order to translate research findings into standard clinical practice. One major problem with current ultrasound pachymetry technology is that it shows a moderate degree of both user-dependence and instrument-dependence. In other words, two users of the same instrument may obtain measurements with different mean values and different variability, depending upon the specific technique applied. Two users of different instruments may obtain different results depending upon the details of the instrument probe construction and ultrasound signal processing algorithms. Additionally, an individual's central corneal thickness varies physiologically. Therefore, the measurement should be repeated over time during several office visits and averaged to obtain the most representative result (Wickham et al, 2005). In order to apply published diagnostic guidelines, the user-dependent and instrument errors must be reduced, and the measurement procedure must be streamlined so that pachymetry can be integrated into a standard eye examination without substantial discomfort and difficulty.

ii. Corneal measurement for refractive surgical procedures

In addition to the newly-recognized application in glaucoma diagnostics, pachymetry is also useful in the context of refractive surgery. Over 1.3 million refractive surgical procedures were performed in the United States in 2001 (Fisher et al, 2003). Corneal refractive surgery has renewed interest in clinical measurement of corneal biophysical properties, such as hydration, refractive index, and thickness. These parameters are important for the development of generalized optical and mechanical models of the cornea, to determine the effects of surgical alteration (Ehlers & Hjortdal, 2004). Clinical measurement of the parameters is also useful for individualized patient care and treatment planning.

The measurement of corneal thickness is crucial as a screening tool for refractive surgery patients. Pachymetry is used to establish that a given patient's stroma will have adequate mechanical strength to maintain its curvature following creation of a stromal flap and ablation of mid-stromal tissue (in LASIK) or ablation of Bowman's membrane and portions of the anterior stroma (PRK). Biomechanical models of the cornea are used to predict the safety factor of a given surgical alteration. If either the model or the measurement is wrong, iatrogenic ectasia or keratoconus may result, causing serious high-order visual aberrations which are difficult to treat.

Accurate knowledge of corneal refractive index is useful in several contexts. Corneal refractive index is a key parameter in optical models of the normal and surgically-altered cornea, but it is also a key parameter for optical pachymetry as well. Optical measurement of geometric corneal thickness requires multiplication of the assumed refractive index and the measured optical thickness. The refractive index of the cornea is known to vary with hydration, and the refractive index is also a function of wavelength. Both of these effects should be taken into account for accurate measurements.

Corneal hydration is important for ablative surgical procedures. The corneal stroma dries rapidly during LASIK, due to the absence of the tear film, epithelium, and Bowman's layer. Room air is drawn across the surface of the cornea at a high rate of flow to collect the plume of ablated material, and this increases the rate of drying. The rate of stromal ablation strongly depends upon the hydration state of the tissue (Dougherty et al, 1994, Kim & Jo, 2001, Fisher et al, 2003). If a cornea is particularly dry, due to drug effects, ocular pathology, delays during surgery, or other anomalous conditions, the ablation rate may be unexpectedly large, resulting in greater removal of tissue than intended. Simply measuring the corneal thickness during ablation might help to correct some of this uncertainty, but without knowing both the thickness and hydration during surgery, it is very difficult to predict the final stromal thickness following surgery, after normal tissue hydration is re-established. Given the importance of these biophysical parameters, there is evident need for low-cost instruments capable of measuring corneal hydration and thickness quickly and safely.

CHAPTER II

SYNTHETIC-APERTURE CONFOCAL PACHYMETRY

In order to improve the accuracy and reliability of CCT measurement for purposes of individualized patient glaucoma risk assessment, it is desirable to reduce the variability of individual measurements using better instrumentation. Important goals are lower user-dependence and lower intrinsic variability. It is also necessary to improve the overall ease and simplicity of the measurement process, so that repeated measurements can be incorporated as a convenient part of routine eye care delivery. Repeated measurements taken over the course of several eye examination visits are the best way to account for the normal physiological range of CCT values for an individual patient. Repeat measurements allow averaging of physiological variation.

In this chapter, existing pachymeter instrumentation will be reviewed, and a new pachymeter, based on synthetic-aperture scanning confocal technology, will be presented. This pachymeter is based on an optical measurement principle that has advantages over ultrasound technology, including the possibility of noncontact operation, while affording the possibility of a low-cost instrument, unlike existing optical pachymeters.

a. Pachymetry as a major diagnostic modality

The accurate measurement of corneal thickness was once a low priority in clinical ophthalmology. CCT is now seen as an important part of individualized glaucoma risk assessment, a part of general eye care. With the findings of the OHTS and later studies, corneal thickness measurement is now a major focus of clinical research.

Several authors have highlighted the difficulty of applying the findings of studies such as the OHTS in a regular clinical practice (Wickham et al, 2005, Shildkrot et al, 2005). This is partly due to the normal physiological variability in CCT, and partly because ultrasound pachymetry is subject to user-dependent and instrument-dependent sources of error.

A measurement repeatability study for ultrasound pachymetry showed large variations (greater than 20 microns) in measurements taken several weeks apart. These variations occurred in greater than 20% of patients (Shildkrot et al, 2005). Another repeatability study showed similar results, with a mean difference of 9.6 +/- 26.9 microns in the right eye and 19.0 +/- 29.2 microns in the left eye, with a systematic bias toward thicker readings on the second measurement (Wickham et al, 2005), likely due to operator factors.

The OHTS investigators took several precautions to assure quality in the ultrasound pachymetry measurement protocol (Brandt et al, 2001). Ultrasound pachymeters of the same make and model (DGH-500 Pachette) were issued to each clinical trial center. Study personnel were uniformly trained and certified on the instrument, and the training included thorough instructions in proper use and calibration protocols. Instruments were calibrated on a monthly basis, and instruments which failed calibration or other periodic quality control tests were immediately replaced. Each reported measurement in the study consisted of the average of five separate measurements taken from each eye, with the right eye measured first. These five measurements were averaged separately, maintaining an independent measurement for each eye.

To help ensure data quality, repeat measurements were performed any time the difference between left and right eyes exceeded 40 microns, since left and right CCT values tend to be very similar. A subset of patients at one clinic (N=77) underwent repeat measurements on a follow-up visit (35 to 560 days later, 384 +/- 75 days on average) to gauge the reproducibility. The mean difference between repeated measurements for these patients was 12.1 +/- 17.2 microns, with a lower difference (7.3 +/- 12.3 microns) when the operator was the same person and a higher difference (12.8 +/- 17.7 microns) when the operator was a different person. The OHTS pachymetry protocol helped to reduce error. Instrument errors were reduced by using a single instrument and following rigid calibration and testing procedures. User errors were reduced by providing highly-uniform training across study centers. Despite the care taken in the protocol design, observer-dependence was manifested in the results, and this level of control may be difficult to implement in a regular clinic.

b. Review of clinical and research pachymetry systems

i. Ultrasound

Ultrasound pachymetry is the current gold standard in clinical practice. Ultrasound pachymetry is based upon a simple measurement principle, and offers the best precision for a given cost of any available technology (Bechmann et al, 2001, McLaren et al, 2004). A pulse of high-frequency sound (typically in the 10's to 100's of MHz) is emitted by a piezoelectric transducer probe, and this ultrasound pulse is transmitted into the cornea while the probe is in physical contact with the corneal surface. The ultrasound pulse propagates through the cornea, and is reflected near the endothelium-aqueous interface due to an acoustic impedance mismatch. The pulse then returns to the probe, where the sound energy couples back into the piezoelectric material, creating a voltage which may be sensed. The pulse transit time, along with an assumed value for the speed of sound in the cornea, are then used to calculate the corneal thickness. The limit of precision is therefore dictated by the narrowness of the initial impulse and the precision with which the propagation can be timed. These intrinsic limits to precision are below 1 micron using current technology. Other factors determine the absolute accuracy and precision of these instruments in practice, including user-dependent errors and physiological sources of uncertainty (McLaren et al, 2004).

Ultrasound pachymeters are available as stand-alone devices from many manufacturers, with a retail price in the range of $2000 to $6000. The agreement and reproducibility of ultrasound instruments among manufacturers has been compared (reviewed by McLaren et al, 2004). In addition to the details of ultrasound probe construction and electronic circuitry design, each manufacturer uses a unique algorithm for obtaining a thickness value from the raw data, and the details of this algorithm are often proprietary. One major source of variability among ultrasound algorithms is the assumed speed of sound propagation in the cornea, known as the acoustic index. The acoustic index of the cornea is given by various authors as 1620 m/s (Atta, 1999), 1640 m/s, and 1664 m/s (Nemeth 2006). The manual for one popular pachymeter, the DGH Pachette 500, states that 1640 m/s is the universally-accepted value, but the manual also explains how the user can choose another value. In addition to the acoustic index uncertainty, the exact structure which causes the acoustic reflection is still a matter of some debate (Ehlers & Hjortdal, 2004).

Several studies show that ultrasound pachymetry is highly reproducible in the hands of a skilled user (Myrowitz et al, 2007, Miglior et al, 2004), generally giving a test-retest difference less than 5 microns for most instruments. However, interobserver reproducibility is usually somewhat less impressive. Miglior and coworkers found a clinically-significant difference (>15 microns) between test and re-test measurements in 52 out of 459 (11.3%) of intraobserver measurements, and 304 out of 1377 (22%) of interobserver measurements. The specific reasons for user-dependence have been linked to probe placement, probe tilting, anesthetic drop administration, and corneal indentation (Miglior et al, 2004, Sabetti et al, 2001, Bechmann et al, 2001, Rainer et al, 2002, Oshika et al, 1997).

While ultrasound pachymetry can yield accurate and reproducible results if sufficient care is taken, one unavoidable shortcoming always remains; the need for corneal contact. In order to couple an ultrasound pulse into the cornea and sense the echo, the piezoelectric transducer head must be in contact with the cornea through a solid or liquid medium. Because contact is required, ultrasound measurement protocols require administration of topical anesthetic. Anesthetic administration increases the examination time, as the anesthetic must be administered and then allowed to take effect. Anesthetics have also been shown to transiently alter corneal hydration and thickness in a dose-dependent manner, which represents a difficult confounding factor (Gao et al, 2006, Asensio et al, 2003, Nam et al, 2005). A non-contact system for pachymetry could potentially reduce the time required for the measurement, and eliminate errors due to effects of the anesthetic. Non-contact pachymetry could also be integrated with other instrumental evaluations.

ii. Optical-coherence tomography (OCT)

Optical-coherence tomography is an imaging modality that has been applied to pachymetry with impressive results. Clinical OCT systems have been designed for retinal 2D tomographic imaging, as well as 2D anterior segment imaging.

Bechmann and coworkers compared pachymetry using a modified retinal OCT system with ultrasound pachymetry, and found good correlation between the two modalities (Bechmann et al, 2001). However, the OCT system gave consistently lower absolute values than ultrasound. The magnitude of this difference was increased in measurements of edematous corneas. The authors theorized that the speed of sound was influenced by edema to a proportionally greater degree than the refractive index. The Zeiss OCT3 has also been used for pachymetry (Fishmann et al, 2005), with a similar result.

OCT, like some other optical methods, has the advantage of noncontact operation and reliable positioning through use of fixation lights. However, like other advanced optical diagnostic instruments, ophthalmic OCT units are expensive compared to ultrasound, and the technology has remained as a secondary diagnostic testing modality. OCT instruments are not available in many eye-care settings, due to reasons of cost.

iii. Optical low-coherence reflectometry (OLCR) and partial coherence interferometry (PCI)

Optical low-coherence reflectometry (OLCR) and partial coherence interferometry (PCI) are advanced optical pachymetry techniques that have shown superior repeatability and accuracy in realistic clinical usage tests (Gillis et al, 2004, Schmid 2003, Gaujoux et al, 2007, Rainer et al, 2002, Rainer et al, 2004). The OLCR system was developed and marketed for clinical use by Schwind and Haag-Streit. The system is based on a fiber-optic low-coherence interferometer which is scanned over a 3mm range of optical path differences by a rotating 30mm glass cube retroreflector in the reference arm (Walti et al, 1998). The cube is rotated at approximately 3.7 Hz, which scans the cornea at approximately 15 Hz. The rotational position of the cube is measured using a 512 CPR optical rotary encoder disk, with a 256X interpolation giving a 50-microradian effective angular resolution. The light source is a fiber-pigtailed super-luminescent diode with a central wavelength of 850nm, which delivers a preset optical power in the range of 8-200 microwatts to the cornea. The detector module is a silicon photodiode with a preamplifier, filter, rectifier, and 60 kHz low-pass filter.

Due to rapid scanning ability and the optical resolution of the device, the OLCR is capable of submicron reproducibility and has very low user-dependence. The OLCR is also able to measure without any patient contact, and the time required for a measurement ranged from 10 to

30 seconds (positioning and measurement time), compared to 60 to 115 seconds (average 80 seconds) per measurement with ultrasound.

While the OLCR system has demonstrated superior measurement repeatability, lower interobserver variation, and non-contact measurement, it has not replaced ultrasound in practice. This is largely due to the price difference. The OLCR has a retail price of ~$10,000 which is almost four times the price of an ultrasound pachymeter. This extreme price difference has prevented widespread adoption of a demonstrably-superior instrument. The price of the OLCR is difficult to reduce by re-engineering, since several critical components, such as the superluminescent diode, cost several hundreds of dollars. A mechanically-scanned low-coherence interferometer requires an alignment process during the initial setup that is somewhat tedious and labor intensive. Further engineering may reduce the cost and price of OLCR instrumentation.

iv. Topography techniques applied to pachymetry

Topographers are instruments that use slit-scanners, projected grids, or reflected images of Placido disk structures to map the corneal surface curvature, usually for the purpose of treatment planning in refractive surgery. Certain topographers, such as the Orbscan, have been demonstrated for pachymetry (McLaren et al, 2004, Chakrabarti et al, 2001, Rainer et al, 2004, Marsich et al, 2000, Modis et al, 2001). The topographer pachymetry measurements tend to show lower reproducibility than ultrasound, and a variable direction of bias. While topography is an automated non-contact modality, it is not well-suited to perform highly accurate pachymetry. Topographers are also expensive. As a consequence, they are not available in all eye care clinics.

v. Scheimpflug pachymetry

Scheimpflug optical pachymeters are an older type of optical pachymeter. While these instruments are accurate and low-cost, they are fairly user-dependent (Ehlers & Hjortdal, 2004). Intraobserver standard deviations are as low as 5 microns, but interobserver variations may be 20 microns or more. Newer Scheimpflug systems, such as the Oculus Pentacam and Pachycam, use a CCD imager to make the measurement process more user-independent and more precise. One recent study found similar repeatability between the Pentacam and ultrasound pachymetry

(O'Donnell et al, 2005), with the Scheimpflug system giving systematically lower results. These systems have integrated keratometry function.

vi. Confocal microscopy used for pachymetry

Confocal microscopes designed specifically for the cornea have been developed in order to examine corneal structures such as the epithelium and endothelium, as well as nerve fibers and keratocytes. Confocal microscopy is also used to detect pathogen microorganisms, such as acanthamoeba cysts. For pachymetry, the confocal microscope has shown impressive performance, achieving micron-resolution and the ability to measure sublayer thickness (Li et al, 1997, Ivarsen et al, 2002). The accuracy is limited mainly by motion artifacts and slow Z scanning. Unfortunately, corneal confocal microscopes are expensive ($40,000 to $100,000), and require an index-matching gel to minimize the required detector dynamic range. Furthermore, these are three-dimensional scanning instruments, which scan X and Y rapidly, and scan Z slowly (McLaren et al, 2004). The confocal optical arrangement is well suited for accurate pachymetry, but current ophthalmic confocal imaging instruments are optimized for a different purpose.

c. Confocal optical systems for optical thickness measurement

i. Review of traditional confocal optical systems

The confocal microscope was invented by Marvin Minsky during his junior fellowship at Harvard in 1955, and was patented in 1957. The confocal microscope is a scanning system which builds a 3D image by scanning an object point-by-point. Illumination is achieved by focusing a collimated beam to a small focal spot with an objective lens. Reflected, scattered, or fluorescent light from the focal spot is collected (typically by the same objective lens), and re-collimated. This beam is then directed by a beam splitter to a second lens, which focuses it to a small focal spot. A pinhole, placed at the focal spot of the second lens, serves as a spatial filter, allowing light that originated at the objective lens focal volume to pass to the detector, while rejecting most of the other light scattered from other focal planes.

The term "confocal" refers to the placement of the point source and detector spatial filter aperture in conjugate focal planes of the lens system. The confocal principle allows for high-contrast imaging, even in thick samples with significant scattering. The pinhole size, in conjunction with the objective lens numerical aperture, determines the depth resolution of the system. A smaller pinhole gives greater depth resolution, but at the price of decreased signal. Pinholes smaller than the Airy diameter significantly limit the signal without appreciably increasing resolution. The depth resolution of confocal optics enables three-dimensional microscopic imaging and accurate thickness measurements of transparent layers.

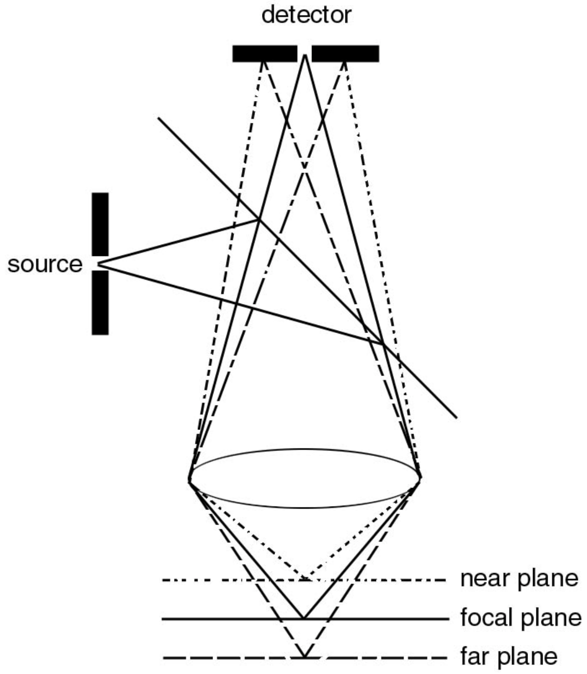

Fig. 1 - In a confocal optical system, a point source and point detector are in aligned in conjugate focal planes. Light originating from the source is focused by an objective lens. Light emitted from the objective focal volume is collected by the objective lens, and is transmitted through a pinhole spatial filter to the detector. Light scattered from other planes is largely blocked by the spatial filter.

Several implementations of confocal imaging systems have been used, with variations on the type of spatial filter employed. The Nipkow disk, with many pinholes arranged in a spiral pattern on a spinning disk, is used to achieve rapid raster scanning of the X and Y dimensions. Slit-shaped scanning apertures have also been used, with slightly-different confocal imaging properties. Fiber optic confocal microscopes have also been developed, where the tip of a single fiber (one arm of a fiber beam splitter) serves as both a light source and detector spatial filter. The beam may also be scanned in X and Y using galvanometer drives or piezoelectric drives. The Z axis is the slowest scanning axis, since it is difficult to move the objective rapidly and precisely.

ii. Synthetic confocal systems: The hologram laser

The hologram laser is a type of integrated optoelectronic component developed in the late 1980's for use in optical disc drives. The purpose of the hologram laser was to increase reliability of optical disc drive mechanisms, while at the same time reducing size and cost, particularly the cost of assembly (Yoshida et al, 2000). A traditional optical disc pickup uses separate components for the laser, grating, beam splitter, and detector. The hologram laser reduces the component count in an optical disc pickup, because the laser diode, grating, holographic optical element (HOE) beam splitter, and photodiode array are all incorporated into a single integrated component. A major benefit of this design is that alignment between the laser source, hologram beam splitter, and detector array is set during component assembly, and maintained afterward in a stable, compact configuration. The components are produced in very high volumes and are low-cost.

Hologram lasers are offered by several manufacturers (SHARP, Panasonic, Toshiba, and Sony) at several wavelengths and laser output power levels, depending upon the intended use. The original CD format used 780nm light, while newer disc formats such as DVD use 650nm light. The new Blu-ray and HD-DVD disc formats use 405nm light. Another key difference between these disc formats is the required NA of the objective lens. The specified NA of a CD pickup objective is .45, while DVD's use an NA of .60. The Blu-ray format uses an NA of .85, while the HD-DVD uses an NA of .65. The hologram laser component is very compact, typically less than 10mm along its longest dimension. Functionally, the hologram laser component comprises most of the elements necessary to form a confocal measurement system. With the addition of an external collimating lens, scanning objective lens, and lens position transducer, a complete confocal system is formed. Because of the small component dimensions, light weight, and unitized construction, the confocal system is rugged and maintains alignment when subjected to physical shocks and environmental fluctuations. Many hologram lasers have an integrated amplifier for each detector.

Hologram lasers are different from typical confocal optical systems, because they utilize a synthetic-aperture system for spatial filtering. The synthetic-aperture system is formed by a multi-element detector and a specially-patterned holographic beam splitter. Light originating at the laser diode source is first split by a diffraction grating on the lower surface of the glass HOE cube, and then collimated by an external collimating lens. The NA of this lens is such that the central beam (zero order) and positive and negative first diffraction orders are collimated. The collimated beam is focused by an objective lens to a very small spot size, near the diffraction limit. Some of the light reflected or scattered from the focal volume region is re-collimated by the objective lens, and returns through the collimating lens. At this point, the light interacts with the holographic beam splitter on the top surface of the glass cube, and is directed toward a multi-element photodiode detector. The amplified signals from individual photodiode elements are then processed to achieve a function equivalent to confocal spatial filtering.

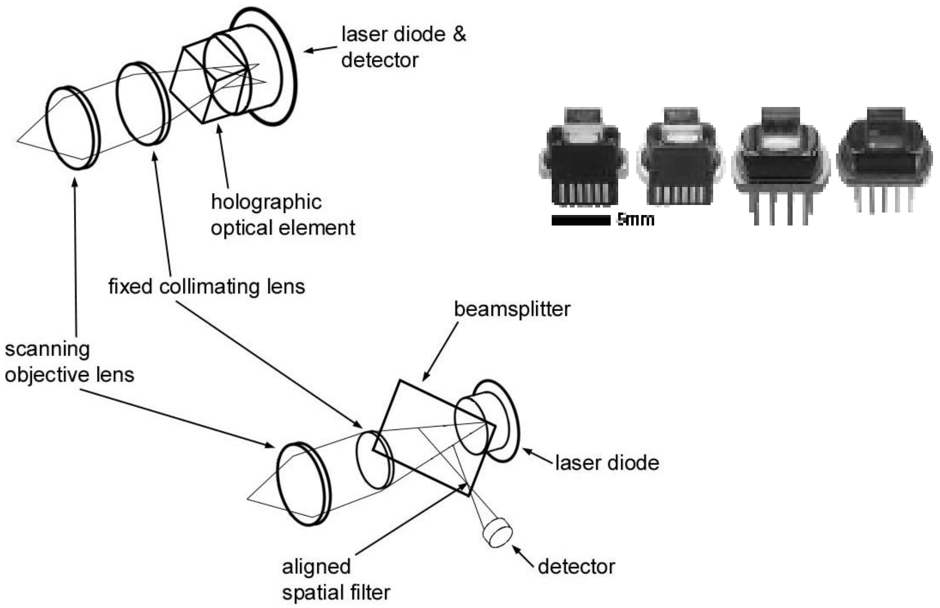

Fig. 2 - A hologram laser is an integrated optoelectronic component that incorporates a laser diode, grating, holographic beam splitter, and multi-element photodiode detector. The component count is lower than a traditional optical pickup assembly, and the resulting device is more rugged and compact. The entire hologram laser component occupies less than one cubic centimeter, which enables very compact device designs.

One such spatial filtering operation is used to calculate the focal error signal (FES), which is used to drive an electromagnetic servo loop that maintains precise adjustment of the focused beam on the surface of an optical data disc. This process is necessary because the polycarbonate or vinyl used to manufacture disks is not perfectly flat, and the clamping mechanism used to mount the disk to the motorized spindle is also of limited precision. Several different versions of the FES signal generation process have been developed and implemented.

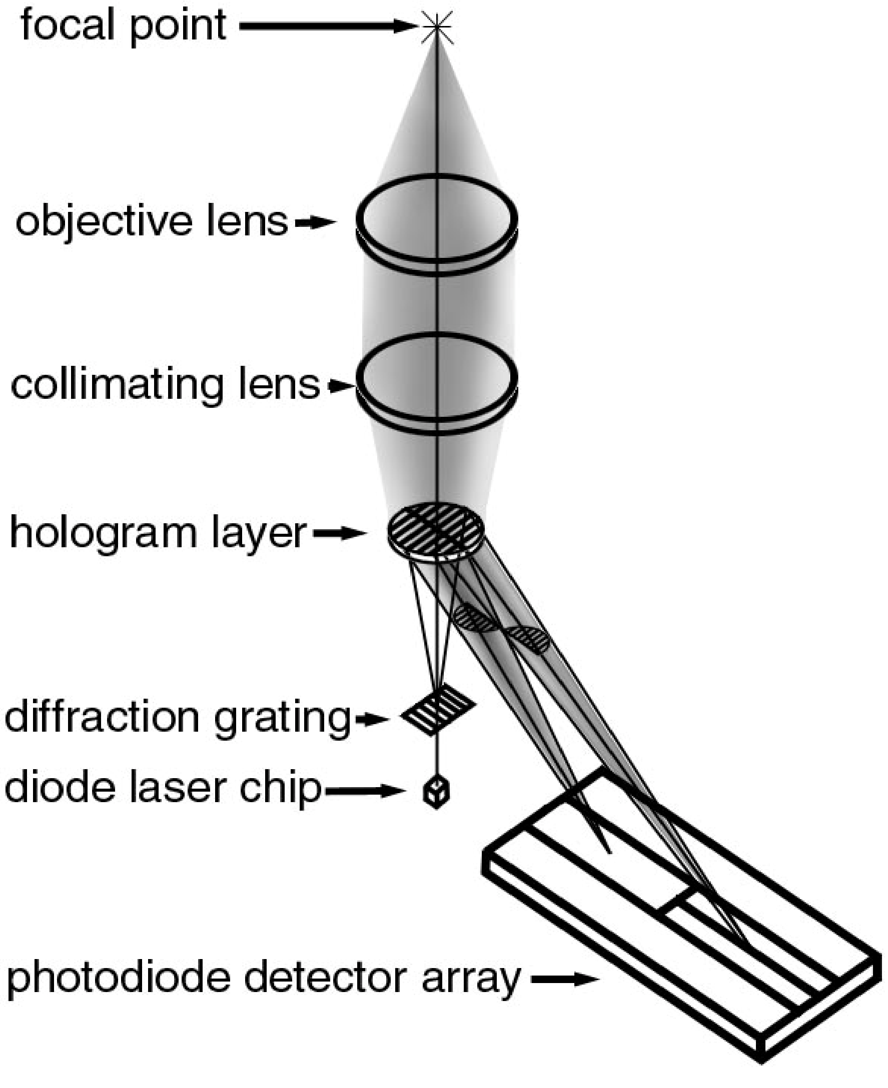

Fig. 3 - The hologram laser includes a diode laser chip, diffraction grating, hologram layer, and photodiode detector array. Also depicted are an external collimating lens and objective lens, which complete the confocal system. A diode laser chip inside the hologram laser emits a divergent laser beam, which passes through a diffraction grating. The purpose of this grating is to generate 1st order beams, which are used to provide a tracking signal when the device is used to read optical disc data. The primary and 1st order beams pass through the hologram layer, and are focused into a parallel beam by the collimating lens. This collimated beam is then focused to a point by the objective lens. A portion of the reflected or scattered light from the focal point is collected by the objective lens, which again focuses it into a collimated beam, propagating back toward the hologram laser. When this light reaches the hologram layer, it is deflected to the photodiode detector array, with a special image pattern arising from differences between zones of the hologram. The image pattern is comprised of six signal spots, which fall onto different areas of the photodiode detector array. Two of the signal spots are from the zero-order beam, while the remaining four arise from the 1st order beams. The four reflected signal spots from the 1st order diffracted beams are not depicted in this diagram for clarity, but these signal spots are shown in another diagram, Fig. 5.

The focal error system used in one type of hologram laser uses a five-element detector. The focal error signal is formed by subtraction of the measured signals from two smaller central photodiode elements (numbered 2 and 3 in Fig. 5). As the reflective disc surface nears the focal point, the signal on one of the elements increases. When the reflective disc surface is at the focal point, the light is tightly focused onto the border between the two detector elements. As the focal spot moves past the reflective surface, the imaged light crosses the border, and registers primarily on the second photodiode element. Subtraction of the signals from the two photodiodes gives an S-shaped curve with a sharp zero crossing at the point where the reflective surface and focal point coincide. Near the zero-crossing, the focal error signal amplitude changes 20% of its maximum for each micron of focal depth change, allowing submicron measurements of surface position in most instances. Interpolation and fitting can further improve resolution.

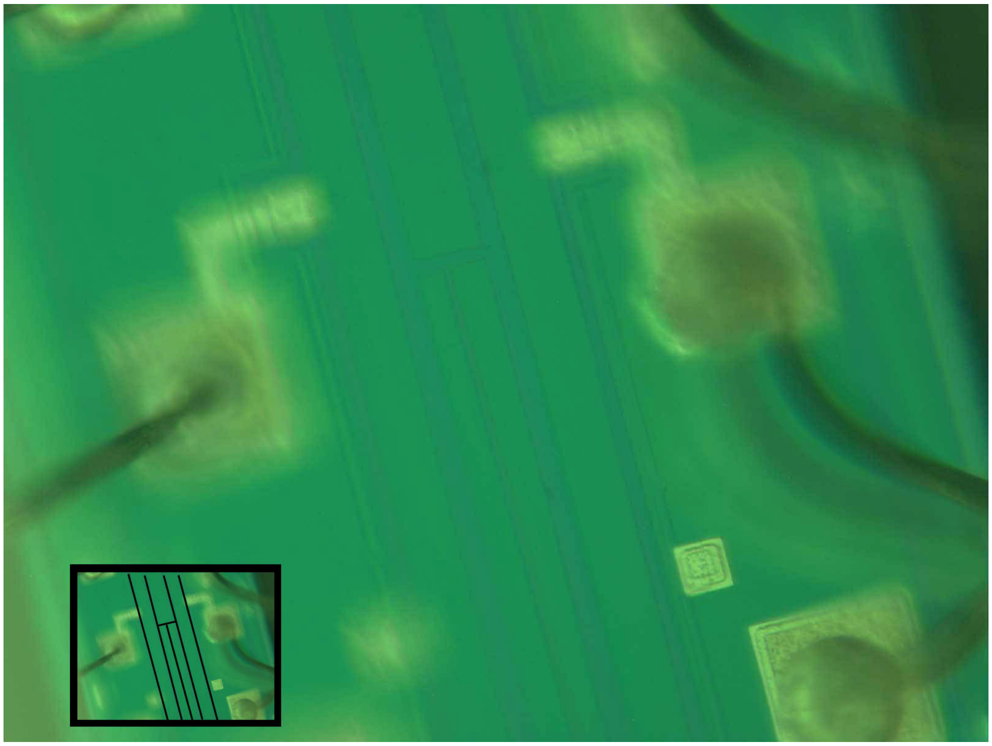

Fig. 4 - A microscopic image of the silicon photodiode array from a SHARP Corporation GH6C605B series hologram laser is shown, with an inset highlighting the borders between the five photosensitive regions. The focal error signal (FES) is formed by subtracting the signals from the two smaller photodiode elements. The GH6C605B incorporates integrated transimpedance amplifiers for current-to-voltage conversion for each photodiode. The border region between the two small photodiode elements may be regarded as forming the synthetic confocal spatial filter, in conjunction with the holographic and electronic signal processing steps.

In an optical disk drive, this signal is used to adjust the position of the objective at a very high bandwidth, which is necessary to maintain a small focal spot on an imperfectly-flat disc as it spins at several thousand revolutions per minute. However, this closed-loop FES-driven system is not sufficient for pachymetry, since there is no scanner/encoder system, which is necessary for measurement of the objective lens position. By adding a scanner and encoder system, the focal error signal from a hologram laser can be used to measure the thickness of a thin transparent object, by measuring the objective lens movement between zero-crossings from the anterior and posterior surface of the object. Thickness is calculated using the difference in these lens positions.

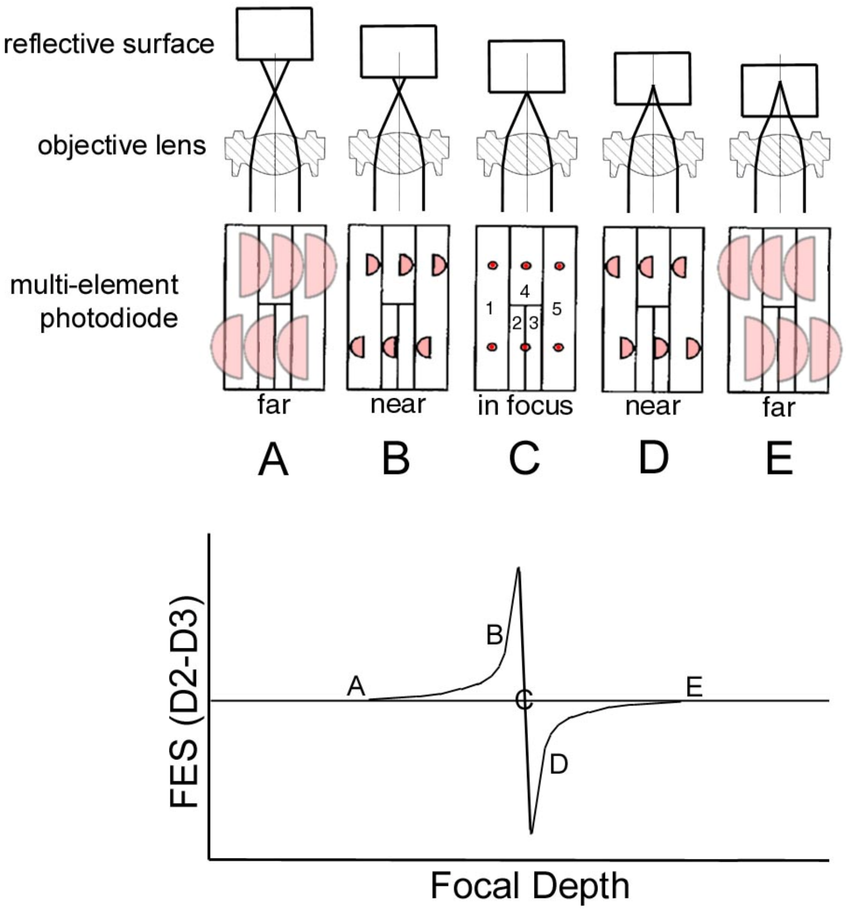

Fig. 5 - When a reflective surface is distant to the focal volume (A), a small amount of diffuse light reaches the detector elements. As the focal volume nears the surface (B), reflected light is focused onto one of a pair of photodiodes (element 2). When the focal volume reaches the surface (C), light is evenly focused onto the photodiode pair (elements 2 and 3). As the focal volume passes the surface (D), the reflected light is imaged onto element 3. Finally, when the surface is far, the imaged light is again diffuse. The focal error signal (FES) is generated by subtracting the signals from photodiode elements 2 and 3. Also shown in the diagram are the spots from the 2nd order diffracted beams. These beams are used in optical disk applications to drive a servo which keeps the main beam centered on a data track.

Compact disk pickup devices have been used in many scientific, industrial, and biomedical applications. The BioCD is a microanalytical application of CD technology, which allows sensitive measurements of extinction coefficient and refractive index by spinning disk interferometry (Nolte & Regnier, 2004). CD technology has also been used for 3D scanning confocal microscopy (Benschop & van Rosmalen, 1991, Windecker et al, 1998, Ehrmann et al, 1998), as well as 2D profilometry (Bartoli et al, 2001, Fan et al, 2002). Finally, CD technology has been applied to the noncontact measurement of transparent object thickness, in LCD screen manufacturing (Ryu et al, 2001).

Compared to a traditional confocal system, the synthetic aperture system has several key advantages. The physical alignment between the point source, holographic optical element, and multi-element detector spatial is less troublesome than alignment with a pinhole spatial filter. This makes the device easier to produce, and more rugged. The focal error signal system also allows a very effective method for automatic peak detection. The zero-crossing of the focal error signal is well defined and amenable to algorithmic identification.

d. Synthetic-aperture confocal pachymeter system design

i. Optical system design

The optical design of the synthetic aperture confocal pachymeter is based upon components designed for optical disc pickups. These include the hologram laser, collimation lens, and objective lens. The hologram laser used in prototyping is a SHARP GH6C605B3B (SHARP Corporation). This device has a laser wavelength of 780nm, and is typically operated at 3mW output power. For this system, the laser is operated at a lower power (<400 microwatts) using an iC-WK laser driver circuit (iC Haus GmbH), with power regulated using the feedback photodiode.

The collimating lens is a Kodak A101 plastic plano-aspheric singlet lens with a numerical aperture of .11, a focal length of 20.12mm, and a design wavelength of 780nm. The transmission at this wavelength is greater than 90%. The lens material is polycarbonate plastic, with a refractive index of 1.5737 at this wavelength. The objective lens is a Philips CAY046 plastic bi-aspheric singlet lens with a numerical aperture of .40, a focal length of 4.64 mm, and a free working distance of 3.1mm. The lens material is acrylic. The Kodak A445 lens was also used as the objective lens in some prototype iterations. It has similar specifications to the CAY046, but is currently out of production. The important lens system specifications are NA, working distance, mass, and diameter.

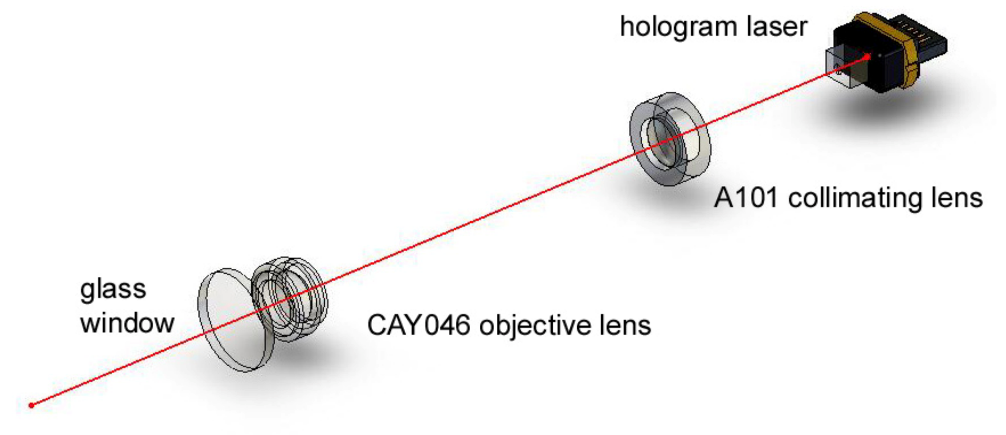

Fig. 6 - The optical components of the synthetic aperture confocal pachymeter are simple and low-cost. The 780nm light output from the hologram laser, a SHARP GH6C605B3B, is collimated by a Kodak A101 .11NA lens. The collimated beam is focused using a Philips CAY046 objective lens, through a glass window, and into the space in front of the window. During Z-scanning, the CAY046 oscillates backward and forward along the collimated beam axis, which causes the location of the focal spot to oscillate as well.

The optical system described is adequate for prototyping purposes, but an eventual clinical system for pachymetry would benefit from a custom design for the objective lens, in order to maximize the free working distance. Opinions vary as to the optimal working distance, but most ophthalmic devices are designed to allow several millimeters between the corneal surface and the instrument. One guideline states that this distance should be at least equal to the length of eyelashes, in order to avoid contact with the lashes during a blink. On the other hand, a shorter working distance may afford greater measurement ease, since the instrument position must be adjusted such that the corneal surfaces are inside the instrument scanning range. The difficulty of finding the correct position may be greater, if it is too far from the front surface of the instrument. A long scanning range or an automated position adjustment system would minimize this difficulty.

ii. Scanning and drive mechanism design

The scanning and drive mechanism are crucial to the overall system performance. The scanning mechanism is the mechanical system which allows smooth, repeatable translation of the objective lens along the system optical axis, with suitable constraints which minimize tilting or translation off of the axis. The drive mechanism is the system which creates and couples the forces necessary to operate the scanning mechanism. The scanning and drive mechanisms are often considered as a unit, because there is considerable functional interplay between the systems.

The key performance metrics for the scanning and drive mechanism are maximum scan rate, maximum scan range (scan amplitude), linearity and repeatability of motion, required power, mean time to failure (MTTF), noise, and vibration. The scanning and drive mechanism must be capable of scanning fast enough so that relative motion between the patient and instrument is small on the timescale of a single scan pass. This is necessary for accurate non-contact pachymetry. Additionally, rapid scanning allows for the averaging of several measurement scans in a short period of time, which can improve measurement accuracy. The range of the scan must be large enough to allow a large measurement window, so that the absolute distance between the probe and the cornea is not critical. If, for example, the scan range is only 1mm, a typical cornea must be precisely positioned within about +/- 500 microns, relative to the pachymeter, in order to complete the measurement. A scan range of 5mm is more forgiving, allowing positioning within about +/- 2.25 mm, a much easier target acquisition range for handheld use. The linearity and repeatability of the motion are also critical. This is especially true if the objective lens position measurement system depends upon assumed characteristics of the motion. Straightness of the translational motion is also key, to assure that effects such as lens tilt or deviation from the optical axis do not modulate or distort the detected optical signal.

Practical considerations dictate the requirements for power consumption, mean time to failure, and noise output. A portable system must be very efficient in order to run on a battery supply for typical daily usage. Mean time to failure for a medical device is significant, since downtime for equipment service or replacement is unacceptable for health care delivery. Most diagnostic ophthalmic equipment is designed to function for several years with minimal service requirements. Finally, noise output must be very low, since the instrument will be positioned near the patient’s eye, and noise can cause anxiety. Noise can also startle the patient, leading to injury.

Flexures are compliant mechanisms that allow constrained, geometrically well-defined motion through elastic deformation of a region of a structure. Flexures are used in several precision optical mechanisms, including scanning interferometers and optical disk drives, as well as precision lithography equipment and EM/AFM stages. A flexure-based scanner mechanism has several key advantages. Flexures are inexpensive and durable, exhibiting a long life of wear-free backlash-free operation as long as the elastic limit of deformation is not exceeded. Flexures allow a smooth and continuous motion, and generate minimal noise during motion. However, flexures cannot tolerate large loads, and the absolute force-displacement relationship is difficult to predict or control, even with very tight manufacturing tolerances (Sydenham & Thorn, 2005).

Spiral flexures are a type of flexure geometry especially well-suited to Z-axis lens scanners. A spiral flexure is created by etching or machining material from a flat substrate (usually metal) to create a disk with a symmetric spiral pattern of relief cuts. These cuts allow motion in the axial direction under a moderate load, but are extremely rigid along radial directions. Using two or more of these flexures, spaced along the axis of desired motion, a very precise translation motion results, with no friction and very minimal tilt or rotation of the moving member. Notably, spiral flexures have been used to guide the pistons of pumps in Stirling cycle cryocoolers for space applications, with no measurable wear after twenty thousand continuous hours of operation.

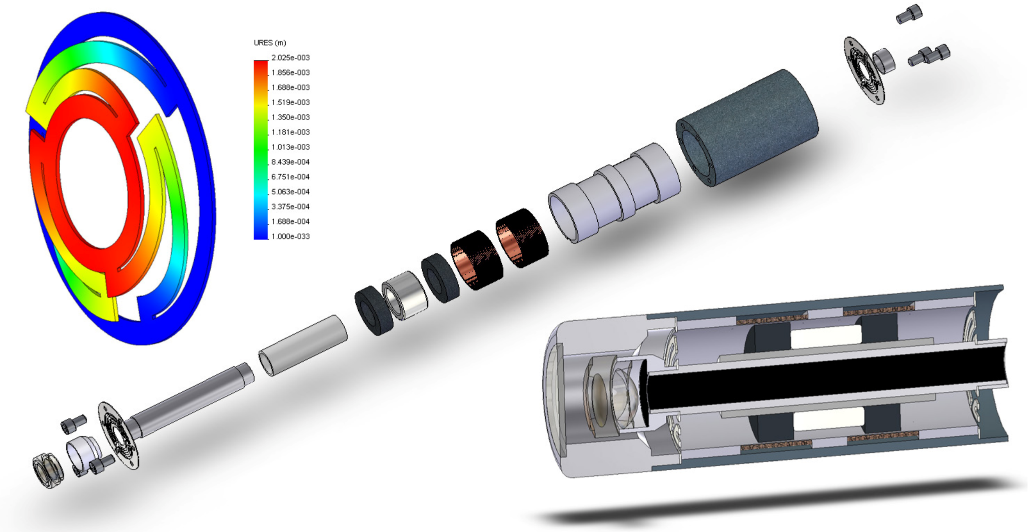

Fig. 7 - A scanner mechanism based on spiral flexure motion guides is shown. The drive mechanism is a moving-magnet linear motor with hollow cylindrical magnets and a pair of counter-wound coils. This scanning mechanism is capable of rapid scanning rates with very low drive power requirements, if used in a resonant mode. Finite-element modeling predicts that this flexure is capable of a 4mm scan range without exceeding the elastic deformation limits of the material. However, the exact amplitude of the scan (scan range) is difficult to predict, because slight differences in the material properties or geometry of the flexures will change the force-deflection properties considerably. This type of scanner and drive system is well-suited for non-contact pachymetry, because it is rapid, compact, durable, efficient, and quiet.

The linear electromagnetic motor drive is well-suited for use with a spiral flexure scanning mechanism. The linear electromagnetic motor is a moving-magnet design that uses a tubular rare-earth magnet and pole pieces, and a pair of stationary counter-wound coils wired in series. The coils are driven by an oscillating current near the mechanical resonant frequency. The drive signal can be commutated using an H-bridge FET circuit. The drive may be open or closed loop (with feedback) A feedback signal can be generated optically, magnetically (using a Hall-effect sensor), or by using the drive coil to sense the moving magnetic field during an idle portion of the cycle. The amplitude of the motion may be adjusted through feedback, using information from the confocal signal arising from the instrument cover window, or another transducer type.

The moving-magnet design is free from durability issues that plague moving-coil designs if large scan amplitudes are necessary. By using hollow cylindrical pole pieces to contain the magnetic flux lines, as well as a ferromagnetic material for the outer housing tube, a very compact and efficient drive with a small air gap can be built. This design with has a free aperture equal to the objective lens clear aperture diameter, a scan amplitude of 4mm, and an outside diameter of less than 15mm. While lower-mass moving-coil designs operate at higher frequency, the moving magnet designs are more reliable because there are no flexible electrical connections.

The spiral flexure scanning mechanism with a linear electromagnetic motor drive has been reviewed and determined to be very well-suited for the synthetic aperture confocal pachymeter. An early design has been developed and simulated. However, this design was not pursued in prototyping, due to time constraints. Improved versions of this design with longer working distance will be investigated for clinical versions of the instrument.

The bushing-based scanner mechanism is a very simple arrangement that can give acceptably precise motion, but with increased noise and wear potential compared to flexures. The prototype design utilizes IGUS Iglide polymer bushings, as well as custom-machined Teflon bushings, in contact with a machined aluminum lens tube with a high-quality surface finish. Because a bushing-based scanner mechanism has significant friction, the lifetime for acceptable precision under rapid scanning is limited. For optimally-designed Iglide bushings subjected to the scanning rates and loads required in this application, the mean time to failure will be less than 5000 hours of operation, which is low for ophthalmic diagnostic devices. The noise produced by a bushing joint may also be objectionable in an ophthalmic device. In addition to the potential for wear, the significant friction of a bushing-based scanning mechanism precludes the use of a resonant linear electromagnetic drive, since the damping and energy losses are so high. Accordingly, a more powerful drive system is necessary, using a rotary electric motor.

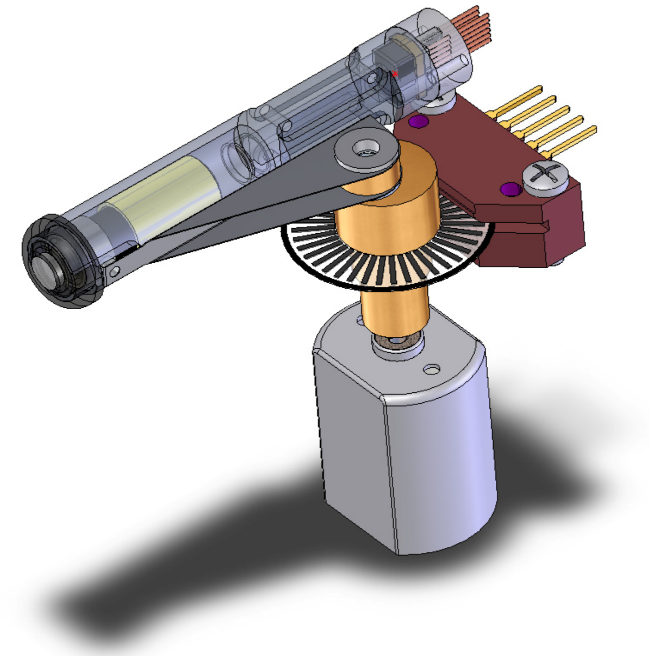

Fig. 8 - A bushing-based rotary drive scanner mechanism with a crank drive coupling is shown above. In this variation, a connecting rod is formed from sheet steel, and the rotational joint between the piston and connecting rod is replaced by a flexure. The flexure is formed by the bend pattern of the steel, so that out-of-plane motion is restricted to the region near the end of the connecting rod. While this type of connecting rod is quiet and compact, two main flaws are present. The need for a clear aperture in the center of the piston restricts the design to an off-center coupling, which causes moment loading of the piston and uneven wear. Second, because of the non-localized bending, the motion of the piston was significantly non-sinusoidal. The non-sinusoidal motion requires a much more involved calibration procedure to ensure accurate measurement.

This rotary motor system is less optimal for the application, because of the added noise and potential for failure at rotary joints, as well as the significantly-higher power consumption. Nonetheless, these rotary-drive busing-based scanner mechanisms are very suitable for research purposes, because they may be fabricated easily, and because they are simple to control and adjust. Two basic rotary drive mechanisms were investigated, with crank or Scotch yoke linkages.

Crank drive mechanisms use a crankshaft, piston, and connecting rod to transform rotary motion into quasi-sinusoidal reciprocating motion. The crank is one of the oldest mechanisms, and records show that cranks were used during the Han dynasty in China 2000 years ago. The connecting rod is often attributed to Watt, but appears much earlier in writings of the Arab scholar

Al-Jazari (1136-1206), whose important text on mechanisms including pumps, locks, and gears is regarded as one of the most significant works in mechanical engineering.

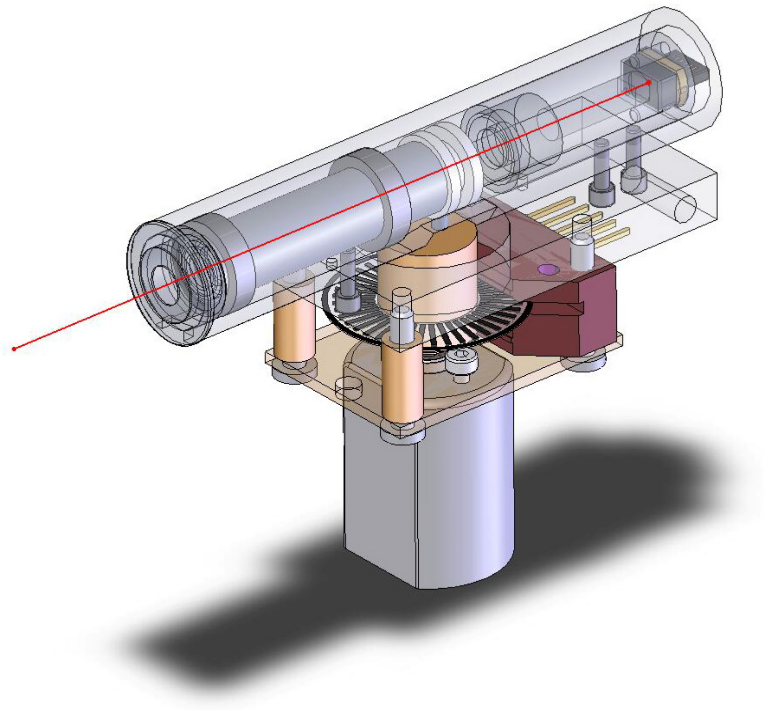

Fig. 9 - A bushing-based rotary drive scanner mechanism with a Scotch yoke coupling is shown above. The tube guide bushings are Teflon or a similar polymer. In this variation, a Scotch yoke is formed by an eccentric steel pin in the motor flywheel, engaged with a slot milled into a Tecamid bushing (glass-reinforced extruded Nylon 6-6), which forms a durable cross-slide bearing surface. The Tecamid cross-slide bushing and aluminum slide tube have a very low mass, and the mechanism is fairly quiet. The motion achieved is sinusoidal within 2 microns across the five millimeter stroke amplitude, and this residual departure is likely due to tolerances in the motor bushings. The mean-time-to-failure for this mechanism is unknown.

Because the crank mechanism has a connecting rod, two rotational joints are needed, in addition to the rotary bearings of the motor. However, full rotational freedom is not necessary on the piston end of the connecting rod, and so a flexure may be substituted for the wrist pin. This variation was pursued in prototyping, as it afforded compact probe tip geometry and a potential for noise reduction. However, this design was deemed unsuitable, due to moment loading and binding of the piston, and non-sinusoidal motion which deviated by more than 50 microns from a 5mm peak-to-peak sinusoid amplitude profile.

The Scotch yoke is another common mechanism for converting rotational motion into reciprocating motion. The Scotch yoke requires no connecting rod, and reduces the side loading and binding of the sliding piston. The primary advantage of the Scotch yoke for rapid Z-scanning of the lens is the very precise and pure sinusoidal motion of the piston. This motion is almost perfectly sinusoidal, which allows a very simple calibration procedure. Objective lens position determination is also simple, using data from a rotary encoder on the motor shaft. Because of the well-characterized motion, this scanner mechanism was used in further prototyping.

iii. Data processing system design

The data processing system for the synthetic aperture confocal pachymeter may be subdivided according to the tasks that must be performed. The system must simultaneously measure the optical signals from the hologram laser photodiode elements and the position of the objective lens. These measurements must be spatially transformed to map the optical signals into a linear coordinate system that describes the objective lens motion, with the correct dimensions. The data processing system must then identify portions of the signal which correspond to anterior and posterior corneal surfaces, and calculate the distance between them as determined by objective lens motion. Finally, this optical thickness value must be multiplied by an assumed refractive index, to arrive at a geometric thickness value.

Several strategies exist for measurement of the objective lens position during scanning. These strategies are determined by the type of transducer employed, and certain transducer types are naturally more suited to a rotary or linear drive system. Several transducer types were evaluated for suitability in this application.

Rotary encoders are transducers which enable direct high-resolution digital determination of rotary position. These encoders may use optical or magnetic information on an encoder disk to indicate the position, and the encoding system may be absolute or indexed. Absolute rotary encoders give the current rotary position at all times, using a multi-track radial bit code, such as Gray code. However, these encoder systems have a lower resolution and higher cost, and are usually needed only when the system must instantly find the rotary position after a power interruption. Indexed rotary encoders use only two tracks, a position track and an index pulse track, to measure rotary position information with very high resolution. The encoder module typically uses a quadrature system to obtain rotation direction information.

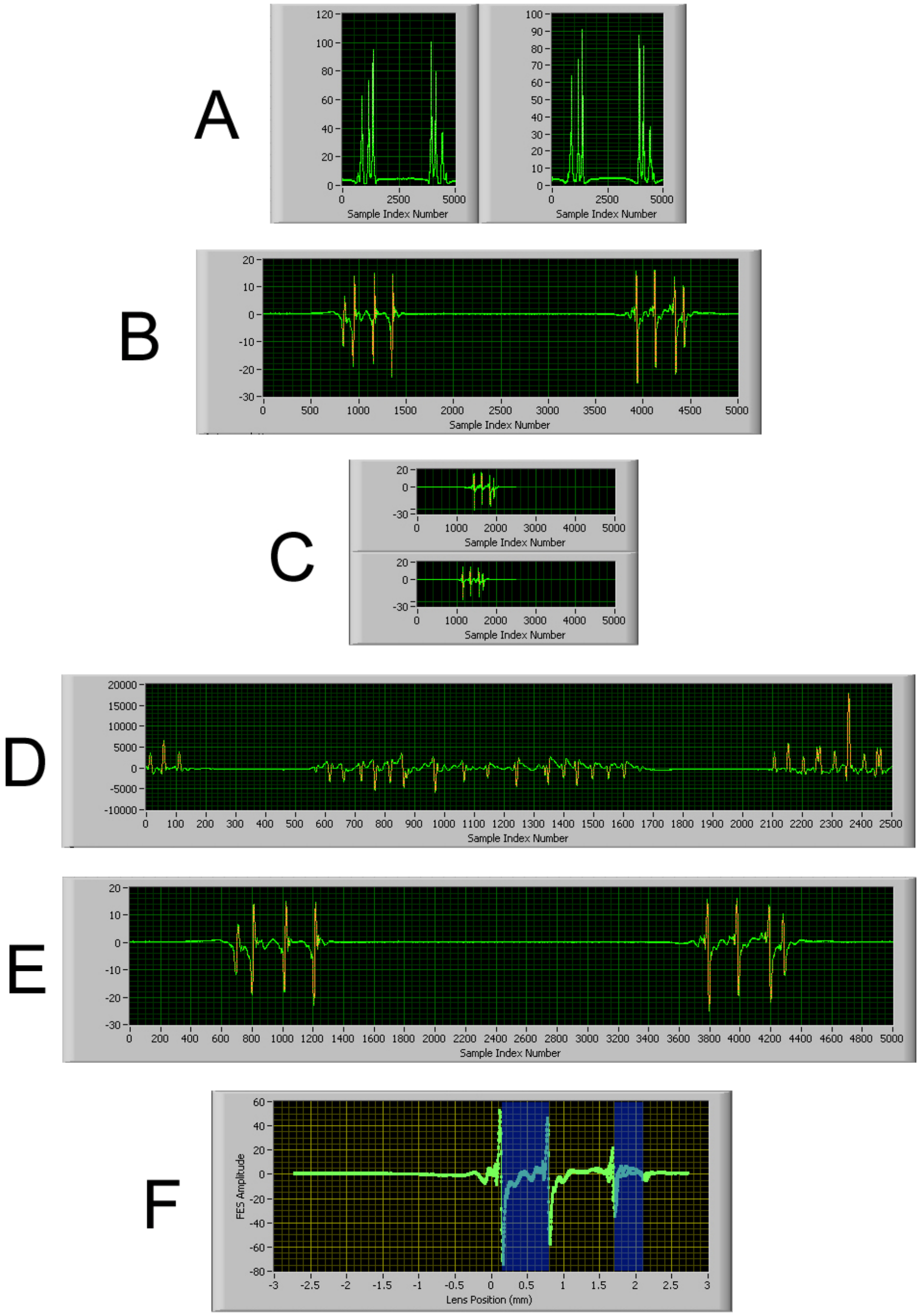

Fig. 10 - The signal processing system used for synthetic-aperture confocal pachymetry is shown graphically. A 5000 CPR rotary encoder is used to trigger data sampling, giving a data block of 5000 samples comprising one full scan period. The two raw photodiode signals from the hologram laser (A) are subtracted, giving the focal error signal (B). This signal block is divided into two blocks of 2500 samples, and the order of one block is reversed (C). The cross-correlation of these two blocks is computed as a function of the folding point (D). The maximum value of the cross-correlation gives the index value of the minimum scan stroke amplitude (in this example, position 2350). The phase of the focal error signal can be aligned using this value (E). Using the known thickness and refractive index of the glass window, and assuming that the waveform of the scanning motion is known, the data can be mapped into a linear scale showing the signal as a function of objective lens position (F). The thickness of a second transparent layer, such as the cornea, can then be computed if the refractive index is known. To determine the precise location of a reflective surface, the zero-crossing location must be determined. A simple approach uses slope and high-low thresholds to window the data, and then looks for the index of the nearest value to zero. A more complex approach fits a line to the windowed zero-crossing data, and calculates the location of the zero-crossing on the scan motion curve.

Hall-effect sensors are based on the interaction of a magnetic field with an electrical conductor through which an electric current is flowing. The magnetic field creates a voltage between opposite sides of the conductor, referred to as the Hall voltage. Hall-effect sensors have been incorporated into sophisticated integrated circuits which give a highly-linear output voltage signal in response to linear displacement of a bar magnet. The linear range of these sensors is limited to one-half the length of the magnet. A non-contact Hall-effect sensor is suitable for direct position transduction for scanners with a linear electromagnetic drive. However, there is some potential for crosstalk between the Hall-effect sensor and the magnetic drive system.

Various processing strategies may be used to recover the scanning lens position from the optical signal itself, particularly if the scan path of the focal volume passes through one or both surfaces of the cover window. If the ratio of the sampling frequency and the scanning frequency is known and the scanning motion is known to be sinusoidal and symmetrical in the forward and reverse direction, an autocorrelation method may be used to determine both the scanning amplitude and phase. This method makes use of an internal calibration signal which may be derived from the window, which has a stable known thickness and refractive index. The motion of many flexure-based scanners is very nearly sinusoidal. A system of this type eliminates some of the complexity and cost of position transducers, but may require more processing power to implement.

If the sampling frequency is F_{sample} and the scanning frequency is F_{scan} , the ratio of F_{sample} to F_{scan} gives the number of samples per scan period, N. For example, if the scanning frequency is 1 Hz, and the sampling frequency is 5000 Hz, then a block of N=5000 samples will comprise one full scan period. The block of N samples can be regarded as a circular array. The N samples can be further subdivided into two blocks of $N/2$ (with a zero pad on the shorter array if N is odd). The data order of one of these blocks is reversed, and the cross-correlation of the two blocks is computed for each possible folding point ($N/2$ possible points). The folding point with the highest computed cross correlation value will be either the maximum or minimum of the scanning stroke, depending upon the phase. This folding point can be used to align the phase of the sampled data, which will now begin at either 0 degrees or 180 degrees relative to the scan waveform.

Once the sampled data block is aligned in phase to the scanning period, the next problem is to determine the scan amplitude. The optical signals from the surfaces of the window, which are present in all scans, can be used to determine both the sign (+/-) and amplitude of the scan, if the waveform of the scanning motion is known. In the case of both Scotch yoke and resonant flexure scanners, the motion waveform is a nearly pure sinusoid. Since the physical thickness and refractive index of the window are known stable quantities, they can be used to determine the scan amplitude, and serve as an internal calibration system for flexure-based scanning mechanisms which may drift if the driving force changes or if external forces (such as gravity) change direction as a result of changes in instrument orientation, unavoidable in a handheld medical device.

The optical signal from the hologram laser requires little external processing before digitization. A pair of adjustable-gain voltage amplifiers based upon the Burr Brown OPA657 op amps is used as a buffer between the hologram laser and the data acquisition board. Two channels of a 4 channel 16-bit data acquisition board (National Instruments PCI-6122) are used to digitize the amplified signals from photodiode elements 2 and 3. The acquisition is triggered using the output of an indexed incremental digital rotary position encoder (US Digital EM1-1-1250). The output from the encoder is processed using an LS7184 ASIC (LSI Computer Systems) to enable 4X quadrature demodulation, and to condition the TTL pulses to ensure a uniform pulse width to trigger the data acquisition. The encoder index pulse is used to start a data acquisition block, and 5000 samples are then taken using the clock signal from the LS7184 IC, operated in 4X mode.

The amplified, digitized signals from the hologram laser photodiode detectors are digitally low-pass filtered, and the focal error signal is computed by subtraction. The autocorrelation algorithm is used to adjust the scan phase to the data block phase. The calibration algorithm is used to determine the scan amplitude. While each of these values is computed for each measurement scan, experiments have shown that they are stable. For this type of scanning mechanism, continuous calibration is unnecessary, but may be necessary for a flexure-based scanning mechanism in the future. The scanning mechanism and encoder mechanism currently limit the prototype system resolution to 1 micron. It may be possible to improve this resolution and reduce device cost at the same time.

iv. Performance and cost

The prototype system exhibits performance equal to more-sophisticated optical pachymeter systems, at a much lower cost. The prototype gives +/-1 micron repeatability when measuring standard test objects, with accuracy determined the refractive index uncertainty. The intrinsic optical Z-resolution of hologram lasers is submicron, so it is possible that future engineering of the scanning system will improve precision even further. The total cost of the optical components is less than \$20. The control electronics and analog signal conditioning circuit cost less than \$10. The rotary optical encoder is the most expensive component, at \$33 in single-unit quantities. The cost of the unit at 100-unit quantities is \$22. The laboratory data acquisition and digital signal processing for this device can be performed on a low-cost microprocessor, although external ADC chips may be necessary. The ADC chips used on the data acquisition board are obsolete, but the next generation in the product line is priced at less than \$20 per channel. Allowing for custom manufactured parts and assembly, the total unit cost for this device should be below \$250. A clinical version of the described system will likely employ a flexure-based scanner and linear electromagnetic drive system for reasons outlined earlier. The signal processing and automatic calibration algorithms demonstrated in this prototype will allow very accurate operation of this type of scanner.

v. Integration with other diagnostic modalities

Two major goals guide this instrument development effort. The first goal is to reduce the inherent uncertainty (due to user-dependent errors) that is present in a single pachymetry measurement. The second goal is to reduce the time and effort necessary to perform pachymetry, so that it can be seamlessly incorporated into part of the routine eye exam. This second goal is very important. Because of the physiological variability of CCT, the measurand is a moving target. It must be measured several times to accurately determine the range of thicknesses, regardless of technical or user-dependent sources of error that may also be present. Non-contact optical methods such as OLCR and PCI have shown very good repeatability. This is most likely due to the absence of factors such as corneal indentation and eye-drop interactions. The potential for tilt and positioning errors is also lower for optical methods which use visual fixation by the patient in order to achieve repeatable measurement position, along the line of sight. The synthetic aperture confocal pachymeter is also non-contact, and uses an optical measurement principle that is intrinsically precise, so it is reasonable to assume that it can achieve similarly-impressive repeatability in clinical use. A non-contact measurement has a definite speed advantage over a contact instrument, but in order to maximize efficiency, it may be desirable to integrate the synthetic confocal transducer with existing diagnostic instrumentation, such as the automated pneumotonometer or automated refractor/keratometer.

An integration of noncontact pachymetry with an existing diagnostic platform will allow the measurement to be performed with very little additional measurement time, since the patient is already positioned. Integrated pachymetry with routine readings would improve the reliability of the measurement, and thereby improve the value for glaucoma risk assessment according to published clinical guidelines. Reliability and availability of evidence is a precondition for evidence-based medicine.

CHAPTER III

HYDRATION-REFRACTIVE INDEX-THICKNESS RELATIONSHIP

The relationship between corneal hydration, thickness, and refractive index has been investigated by many authors. The corneal thickness increases as the tissue hydration increases, due to swelling. The refractive index of the cornea decreases with increased hydration. This is because the refractive index of water is lower than the refractive index of hydrated fibrils and ground substance components. For many purposes, simply appreciating these two general trends is satisfactory. However, a more detailed model is necessary in some cases. In particular, accurate optical measurements of corneal thickness require knowledge of the corneal refractive index. Optical models of the cornea also require accurate refractive index input parameters.

All optical methods of pachymetry, whether based on interferometry, confocal microscopy, or another method, are indirect. They involve multiplication of a measured apparent thickness by an assumed refractive index value, either phase or group. This is no different from ultrasound methods, which assume an acoustic index value that is unknown for an individual. Early developers of optical pachymeters suggested that the effect of refractive index variation on the measurement is negligible (Maurice & Giardini, 1951). Other authors have offered analysis that suggests that refractive index variation could cause a substantial error, on the order of +/- 3% (Arner & Rengstorff, 1972). These results were controversial at the time, and were immediately disputed based on application of the Gladstone-Dale equation and the Hedbys-Mishima equation (Fatt & Harris, 1973).

a. Review of previous literature and models

Hydration status is known to play a determining role in corneal thickness, in both normal and pathological physiology. It is believed that hydration-induced swelling occurs almost exclusively along the thickness dimension. An enduring model of the relationship between corneal hydration (H) and corneal thickness in millimeters (L) was developed by Hedbys and Mishima (Hedbys & Mishima, 1966).

$$H = (a \times L) - b$$

Equation 2

The Hedbys-Mishima equation makes use of two empirical parameters that are species-specific. These values incorporate the normal species baseline thickness and expansion coefficient. In humans, the suggested values are a=7.0 and b=0.64, while in New Zealand rabbits of 2 to 3.5 kg body mass, the suggested values are a=10.0 and b=0.426. These values hold few significant figures, and the form of the equation requires some algebraic manipulation to calculate hydration for corneas with a known (baseline) thickness at a known (baseline) hydration. This equation assumes that at a normal hydration (H=3.2, ~76.2% water by mass), a "normal" human cornea has a baseline thickness of ~549 microns. This is roughly 15 microns above the modern accepted average from pachymetry studies. From this observation alone, one may conclude that the equation structure, human empirical parameter values, accepted normal hydration value, or accepted normal corneal thickness, or some combination of these, is in error.

A review of the original Hedbys-Mishima papers shows that, in fact, these authors were well-aware of baseline differences within species. Hedbys and Mishima found that New Zealand rabbits with 3.5 to 5 kg body mass had hydration parameters a=8.65 and b=0.626. The utility of the equation for absolute modeling of individual corneas is thus called into question. It is not well-suited to model a particular cornea with known baseline thickness, as it does not make use of this information. Hedbys and Mishima used the Maurice-Giardini optical pachometer to measure thickness. No effort to assess or correct a hydration-dependent index shift effect was made.

The biophysical hydration-thickness model has remained virtually unchanged since the 1960's, and has been applied quantitatively in later analysis without verification of the empirical parameters. Later studies of the refractive index-hydration relationship use pachymetry to gauge hydration in individual corneas, which assumes the applicability of particular thickness-hydration relationship parameters. The Hedbys-Mishima equation gives a linear relationship between corneal thickness and hydration which has been verified by many authors, but the absolute parameter values clearly depend upon individual cornea characteristics.

The Gladstone-Dale equation gives a first-order mathematical relationship for the refractive index of mixtures. The equation was first put forth in the 1850's as a tool for gaseous mixture analysis, and was first applied to analysis of corneal refractive index in a classic paper (Maurice, 1957). According to the Gladstone-Dale equation, the refractive index of a mixture may be calculated as a partial sum of the component indices, weighted according to the volume fractions (of the total volume) that they occupy.

$$n_{total} = n_1 f_1 + n_2 f_2 + n_3 f_3 + n_4 f_4 + \cdots n_i f_i$$

Equation 3

Since the sum total of the component volume fractions is unity, the Gladstone-Dale equation may be rewritten in terms relative to a particular component's refractive index. This rewritten form of the equation is often used in the corneal models, written in terms of the water or aqueous humor component index.

$$n_{total} = n_1 + f_2(n_2 - n_1) + f_3(n_3 - n_1) + \cdots f_i(n_i - n_1)$$

Equation 4

The difficulty of direct application of the Gladstone-Dale law is that it requires accurate knowledge of component refractive indices, densities, and relative composition. The components of the cornea are not easily separable in the native physiological state, and therefore, the refractive indices of the native components cannot be measured individually. Gladstone-Dale models of the cornea are built using a variety of measurements from modalities such as x-ray diffraction, electron microscopy, and Abbe refractometry. Several assumptions are typically necessary. These measurements are then used to accurately calculate component volume fractions in the normal physiological state. However, the indices for the separate components must also be accurately-known, in order to build a good model.

The Gladstone-Dale equation for the hydrated corneal stroma may be stated in terms of hydrated-fibril and hydrated-extrafibrillar matrix components.

$$n_{stroma} = f_{f/s} n_f + f_{e/s} n_e$$

Equation 5

The indices of the hydrated fibrillar and hydrated extrafibrillar components of the stromal Gladstone-Dale model can be calculated from a second sub-set of Gladstone-Dale models, using the refractive indices of water, dry collagen, and dry extrafibrillar material as input parameters.

$$n_f = n_w + \frac{f_{c/s}}{f_{f/s}}(n_c - n_w)$$

Equation 6

$$n_e = n_w + \frac{f_{p/s}(n_p - n_w)}{1 - f_{f/s}}$$

Equation 7

The equation for the extrafibrillar material refractive index can be rewritten in terms of the dry collagen fibril/stroma volume fraction ($f_{c/s}$) and the hydrated fibril/stroma volume fraction ($f_{f/s}$). With this rearrangement, only two volume fractions are needed to calculate both fibrillar and extrafibrillar indices. These volume fraction parameters have been determined through x-ray diffraction studies and chemical analysis of corneal composition for many species.

$$n_e = n_w + \frac{(n_{stroma} - n_w) - f_{c/s}(n_c - n_w)}{1 - f_{f/s}}$$

Equation 8

Meek, Dennis, and Khan added hydration-dependence to the basic Gladstone-Dale model of the swollen corneal stroma (Meek, Dennis, & Khan, 2003). This model was developed to predict both the composite (stromal) refractive index, as well as the mismatch which may occur between fibrillar and extrafibrillar components, leading to increased scattering. The model indicated that the increase in mismatch alone could not account for the observed loss of transparency. The authors concluded that scatter arising from fibril spacing effects was dominant, because the relative increase in index mismatch was small (Meek, Dennis, & Khan, 2003).

$$n_{stroma}(H) = n_w + \frac{f_{f/s} n_f + f_{p/s} n_p - (f_{f/s} + f_{p/s}) n_w}{1 + f_{w/s}\left(\frac{H}{3.2} - 1\right)}$$

Equation 9

In the Meek-Dennis-Khan model, the hydration-dependent refractive indices of stroma, hydrated fibrils, and hydrated extrafibrillar-material can be calculated from three refractive indices (water, dry collagen, and dry matrix) and four volume fractions (dry collagen in stroma, hydrated fibrils in stroma, dry matrix in stroma, and water in stroma). However, some changes to this model may be necessary. In particular, little attention was paid to the wavelength-dependence of input parameters to the model, such as the refractive indices of collagen and water.

Maurice used n_c=1.547 for the refractive index of dry collagen, and calculated n_f=1.47 and n_e=1.345, values which were later revised to n_f=1.51 and n_e=1.345. The value n_c=1.547 matches the 1953 measurement by Aurell and Holmgren (Smith, 1969). Maurice's measurements used the Becke line method or an Abbe refractometer, but no measurement wavelength was stated in either case. The Becke line method involves the use of index matching fluids and a microscope. The sample is immersed in a set of fluids of known refractive index, and the microscope stage is lowered. If the sample has a higher index than the liquid, a faint line moving into the sample is observed. If the sample has a lower index, a faint line moving out is observed. A more extensive discussion of the technique and its limitations is given by Faust (Faust, 1955).

Leonard and Meek calculated the average component refractive indices from 40 species as n_f=1.416 and n_e=1.359, with very small differences between species (Leonard & Meek, 1997). Leonard and Meek used Maurice's original value for the refractive index of dry collagen, and n_{stroma}=1.375 (citing Sivak) in their calculations. Leonard and Meek used n_w=1.333 (approximating water), while Meek, Dennis, and Khan used n_w=1.335 (approximating saline) in a later paper. In each instance, the wavelengths for the calculation input parameters (refractive indices of dry collagen and water) were not well-specified. Additionally, component indices derived from these parameters may be flawed. The models may suffer from several input parameter errors, which cause them to poorly describe measurements taken at a given wavelength.

Material chromatic dispersion is the variation in refractive index as a function of wavelength (Born & Wolf, 1999, Smith, 2000). The dispersion of glass is responsible for the functional behavior of prisms, which separate an incident beam into a spectrum. According to Snell's law, the angle of refraction for each component color depends upon the index of refraction of the prism, and so different wavelengths will be bent at different angles. Dispersion in imaging systems is generally undesirable, responsible for chromatic aberration.

The chromatic dispersion properties of ocular media are important for models of ocular optics, and dispersion curves have been measured directly for several ocular media using various techniques. Equations for the chromatic dispersion of ocular media at physiological hydration (including the cornea and aqueous humor) have recently been published (Atchison & Smith, 2005). These equations are Cauchy equations with parameters based on previous measurements of ocular media dispersion by several authors, including Sivak and Mandelman (Sivak & Mandelman, 1982). Water has a well-studied chromatic dispersion curve. Pure water at 21.5°C has a refractive index of n=1.328821 at 780.237nm, and a refractive index of 1.333260 at 587.725nm (Daimon & Masamura, 2007). Salt water at the salinity of the sea (34.998%) is also dispersive. At 20°C, this salt water has a refractive index of n=1.33644 at 700nm and an index of 1.33972 at 579.1nm (Quan & Fry, 1995). The indices of salt water and aqueous humor are both higher than the index of pure water at a given wavelength, temperature, and pressure.

The existing Gladstone-Dale model of the hydration-refractive index relationship cites values for the refractive indices of collagen and water at unspecified wavelengths. Based on existing corneal chromatic dispersion data alone, the argument can be made that new models of corneal refractive index should be built (and tested) at a specified wavelength.

$$n_{stroma}(H,\lambda) = f_{w/s(H)} n_{w(\lambda)} + f_{c/s(H)} n_{c(\lambda)} + f_{p/s(H)} n_{p(\lambda)}$$

Equation 10

A simple Gladstone-Dale model for corneal refractive index as a function of hydration can be written in terms of the refractive indices (at a specified wavelength) of water (or aqueous humor), dry collagen, and dry extrafibrillar matrix, multiplied by hydration-dependent volume fraction terms. The volume fractions of water and dry components at each hydration can be calculated from existing experimental data on the densities of water, collagen, and extrafibrillar matrix, and the known mass proportion of collagen to matrix based on chemical composition and x-ray diffraction studies (Leonard & Meek, 1997). All that remains is to determine appropriate values for the component refractive indices at a given wavelength. The refractive index of water at a given visible wavelength can be readily calculated from the IAPWS equation (Harvey et al, 1998). However, the refractive indices of dry collagen and dry extrafibrillar matrix are more difficult parameters to obtain. Existing Gladstone-Dale models of the cornea use the 1957 Maurice value (n_c=1.547) for dry collagen, and proceed to calculate the refractive index of dry extrafibrillar matrix using a Gladstone-Dale model and an assumed value for the refractive index of the cornea at physiological hydration (H=3.2, n_{stroma}=1.375), again with an unspecified wavelength. Thus, the calculated values for dry matrix index depend upon several input parameters. The values for n_c and n_{stroma} and the component volume fractions may all be in error. One such calculation places the refractive index of dry extrafibrillar matrix at n_p=1.485 (Meek, Dennis, & Khan, 2003).

Very little published data is available on the visible-wavelength chromatic dispersion of dry collagen, and only limited data is available on the dispersion of collagen in its physiological form (Bashkatov et al, 2000). However, the optical properties of dry gelatin have been extensively studied and modeled, including its chromatic dispersion curve. Gelatin is a substance that is used in several areas related to optical engineering, including photographic emulsions, gratings, and holographic optical elements. Gelatin is also used in pill capsules and drug vehicles. The IR and Raman spectral properties of gelatin have also been studied extensively. Gelatin is produced by partial (irreversible) hydrolysis of animal collagen, and it has a very similar chemical structure to collagen. The hydrogen bonds between collagen molecules are broken, and some intermolecular bonds are broken as well, which makes the protein more soluble. A set of Cauchy equation parameters to model the chromatic dispersion of dry gelatin has been published (Martínez-Antón & Bernabeu, 1998), along with measurement data that show the dispersion. If it is assumed that the chromatic dispersion of dry gelatin and dry collagen are similar, a substitution can be made in Gladstone-Dale models of corneal refractive index. It will be shown that the chromatic dispersion of the cornea at physiological hydration is well-described by a Gladstone-Dale combination of gelatin and aqueous humor chromatic dispersion curves. Based on this observation, it is likely that the chromatic dispersion of dry collagen and dry gelatin are very similar.

Cauchy equations, described by Augustin Louis Cauchy in 1836, are parametric equations that relate wavelength to refractive index for a particular transparent substance. These equations are based upon a theoretically-incorrect foundation, and become very inaccurate in the infrared region. The later Sellmeier equation (1871) is used to model anomalous dispersion in the UV and IR, but still gives non-physical results near each absorption peak.

$$n_{Cauchy}(\lambda) = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \frac{D}{\lambda^6} \cdots$$

Equation 11

$$n^2{}_{Sellmeier}(\lambda) = 1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3} \cdots$$

Equation 12

While both equations are theoretically flawed, they are still used extensively to describe optical material dispersion at visible wavelengths. Other dispersion equation forms include

Hartmann, Cornu, Herzberger, and Conrady (Smith, 2000). These equation forms have very little theoretical basis, and are used simply as fitting structures for measured data. A comparative discussion of these equation forms is presented by Atchison and Smith, in the context of ocular dispersion modeling. The Cauchy parameters for ocular media, including the cornea and aqueous humor, are given by Atchison and Smith (Atchison & Smith, 2005). The Cauchy parameters for dry gelatin are given by Martínez-Antón and Bernabeu (Martínez-Antón & Bernabeu, 1998).

Cauchy Parameters From Atchison & Smith, 2005				
	A	**B**	**C**	**D**
cornea	1.363471E+00	6.017972E+03	-6.770081E+08	5.916596E+13
aqueous	1.324294E+00	6.083028E+03	-7.076535E+08	6.160249E+13
Cauchy Parameters From Martinez-Anton & Bernabeu, 1998				
gelatin	1.530000E+00	1.788000E+03	5.730000E+08	

Fig. 11 - The Cauchy parameters for the cornea and aqueous humor have been determined based on dispersion measurements by several authors. The dispersion of dry gelatin has also been measured, and Cauchy parameters for this dispersion curve are also available. Gelatin is very similar in composition to collagen.

With these parameters, a set of dispersion curves can be generated over the wavelength range of 400nm to 900nm. From these curves, it is observed that the aqueous humor dispersion is very similar to the IAPWS water dispersion curve, but with a higher absolute value. It is also observed that the dispersion curve for dry gelatin intercepts the 1957 Maurice value (n_c=1.547) for dry collagen, at approximately 500nm (blue-green region of visible spectrum).

Using the Martinez-Anton Cauchy model for gelatin, the Atchison Cauchy model for aqueous humor, and the Leonard-Meek values (citing several sources) for collagen and matrix density, wavelength-corrected Gladstone-Dale models can be built. These Gladstone-Dale models closely match the Atchison Cauchy model for corneal dispersion (at the physiological hydration of H=3.2) in both magnitude and curvature. In the two-component Gladstone-Dale model (gelatin and aqueous humor), the density of the gelatin component is the 15:7 weighted-average (Maurice, 1957) for collagen (*?=1.41 g/mL)* and extrafibrillar matrix (*?=1.06 g/mL)*, giving an intermediate density value of (*?=1.3 g/mL)*. This two-component Gladstone-Dale model gives a very close match to the Atchison Cauchy model of corneal dispersion (at H=3.2). Using the literature values (Leonard & Meek, 1997) for collagen and matrix densities, a third Cauchy dispersion curve for the dry matrix component can be determined, which removes the residual error and gives putative refractive index values for the dry matrix component.

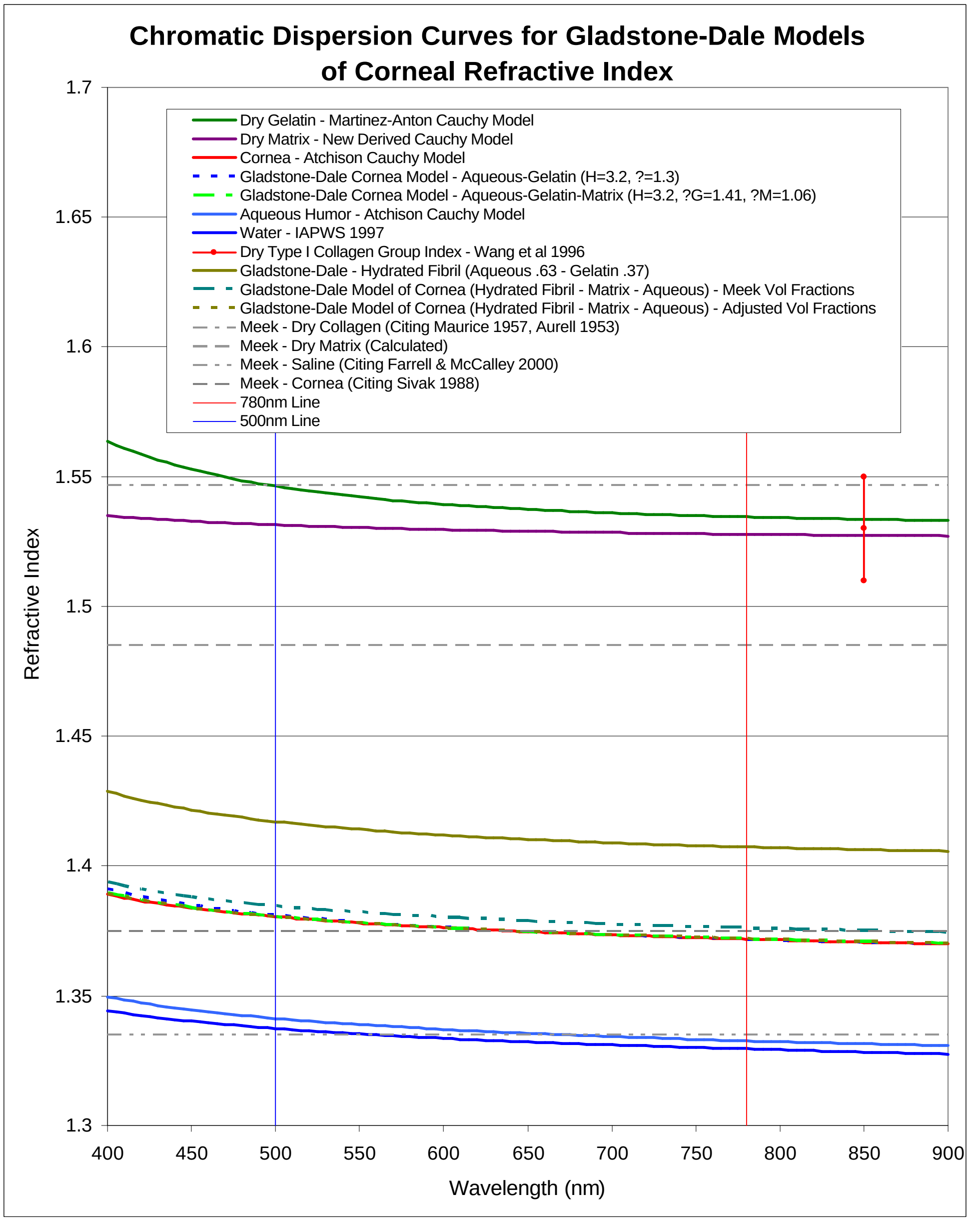

Fig. 12 - The dispersion curves for several substances are plotted for comparison, along with horizontal lines which show indices used in various Gladstone-Dale models of corneal refractive index. The dispersion curve for dry gelatin intersects the value (1.547) used in models of dry collagen near 500nm. The dry gelatin dispersion curve also approximates one measurement of dry collagen index, taken at 850nm. The literature value for corneal refractive index (1.375), used in the derivation of some published Gladstone-Dale models, intersects the corneal dispersion curve at ~640nm. The refractive index value used for the solvent component (1.335) intersects the water curve at ~550nm and the aqueous humor curve at ~675nm. Thus, it appears that published models may incorporate component refractive index values without regard for chromatic dispersion.

The two-component Gladstone-Dale model (aqueous humor and gelatin) has one assumption (density of the gelatin component is taken as the average of known corneal collagen and matrix densities) and no free parameters. The three-component Gladstone-Dale model (aqueous humor, gelatin, and matrix) has two free parameters. These are two Cauchy parameters to model the putative dispersion of the matrix component (A=1.5253 and B=1526).

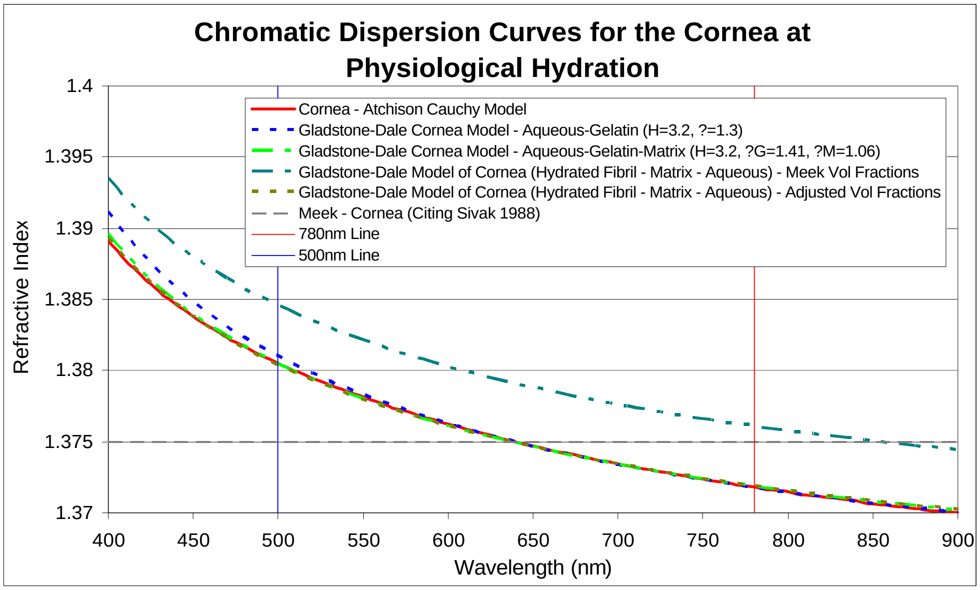

Fig. 13 - The dispersion of the cornea, given by the Atchison Cauchy model, is well-approximated by Gladstone-Dale models built from aqueous humor and dry gelatin. In the aqueous-gelatin model, the matrix is modeled as having dispersion properties identical to gelatin, and the known density and mass fractions are used to give an overall component density of 1.3 mg/mL. In the aqueous-gelatin-matrix model, the known densities of collagen and matrix, as well as the mass fraction, are used to calculate a putative dispersion curve for the dry matrix by solving for two Cauchy parameters.

Both of these Gladstone-Dale models (two-component and three-component) match the measured corneal dispersion curve within the limits of precision of the calculation, determined by the significant figures in the density data, mass fractions, and the hydration parameter. The three-component model, which includes a putative dispersion curve for the matrix component, is virtually identical to the corneal dispersion data. One recent measurement of dry Type I collagen group refractive index, taken at an 850nm central wavelength, gives the value as 1.53+/-.03 (Wang et al, 1996). This is very close to the refractive index value of gelatin at 850nm (1.533), as shown on the dispersion chart. It is concluded from this analysis that dry gelatin dispersion and dry corneal collagen dispersion are probably similar, and that dry gelatin refractive index values (at a particular wavelength) can be substituted for (unavailable) dry collagen refractive index values.

It is interesting to compare the individual dispersion curves of the cornea, aqueous humor, gelatin, and matrix with the values cited in previous models. The index values used for the solvent component (water at 1.333 or saline at 1.335) underestimate the index of aqueous humor (1.341, given by the Cauchy model) at 500nm. The value used for the corneal refractive index (1.375) also underestimates the value for cornea at 500nm (1.380, given by the Cauchy model).

In order to investigate the hydration-dependent refractive index of the cornea, new measurements were made using the synthetic aperture confocal pachymeter to measure the thickness and refractive index of rabbit corneas during slow dehydration. These measurements should serve to accurately determine the refractive index of the cornea at 780nm, and also place bounds on the error caused by hydration-dependent refractive index shifts. Accurate measurement of the hydration-refractive index relationship will enable the use of optical pachymetry for accurate CCT measurement at the NIR laser wavelength of 780nm. Finally, accurate measurements may shed light on the quality of existing biophysical models of this relationship, and indicate areas where the models may be improved upon.

b. Experimental methods

i. Samples

Fresh rabbit central corneal buttons from New Zealand White rabbits (sacrificed under another unrelated medical research protocol) were excised by trephine (Katena 7.5mm) from the animal within 10 minutes of death. After excision, the corneal epithelium was removed by gentle scraping with a scalpel blade. The corneal buttons were rinsed with a small amount of cold deionized water, and all of the buttons were placed into a 10 mL volume of cold deionized water. The corneas were stored at 4°C for a period exceeding 10 days. The corneas were kept in low-temperature storage, but no other special precautions (such as dialysis tubing) were taken to prevent leaching of proteins and glycosaminoglycans from the buttons into the bathing fluid.

ii. Hydration perturbation, drying protocol, and measurement

The corneas were allowed to reach an equilibrium hydration with the 10mL bathing fluid. Corneal hydration increased to a mean value of H=8.23 +/- .26 (SD) for the 10 corneal buttons used in further analysis. Corneas were removed from the bathing fluid, patted dry with lens paper to remove surface droplets, and placed onto the center of pre-massed Corning #1 22-mm square microscope coverslips. Each coverslip was placed onto the stage of a precision analytical balance (Sartorius) and an initial mass was recorded. Each cornea was allowed to dry for 780 minutes while remaining in place on the balance stage. The enclosure of the balance stage was open to the room air. The mass of the cornea and coverslip was recorded at each 60 minute drying interval for 780 minutes (13 hours). A final dry mass was recorded at 2880 minutes. Hydration values were then calculated for each measurement time point as the ratio of water mass to the final dry mass.

iii. Refractive index measurement

To measure the corneal refractive index, a synthetic-aperture confocal pachymeter system (described in a later chapter) is utilized to measure both the geometric and optical thickness simultaneously. The measurement principle for thickness and refractive index is similar to that described by Dirckx (Dirckx et al, 2005). In the Dirckx system, a thin slice of tissue on the order of 100 microns is placed on a microscope slide, and an immersion fluid of known refractive index is spotted onto an adjoining area of the slide. A coverslip is then placed onto the tissue slice, and the apparent optical thickness of the tissue layer and the fluid layer are each measured. The geometrical thickness of the tissue and adjacent fluid layer are assumed to be the same, and the refractive index of the tissue layer may then be calculated as the ratio of the geometrical thickness to the apparent optical thickness.

The synthetic-aperture confocal pachymeter system uses a similar confocal technique to measure the optical thickness, but a slightly-different method is used to determine the geometrical thickness of the sample as it dehydrates. The geometrical thickness T of the cornea is measured as the distance between the position of the top surface of the cornea and the position of the top surface of the microscope coverslip, measured in an area adjacent to the button. The assumption in this case is that the measurement instrument's Z-scanning reference position does not change, relative to the precision balance stage. This assumption can be verified by measuring the position of the coverslip surface before and after the thickness measurements. The reference fluid in this case is air. This reference fluid is nondispersive, an advantage compared to the Dirckx method.

This measurement arrangement is unique in several respects, compared to previous corneal refractive index measurements. Accurate measurement of corneal hydration is somewhat difficult, in that small amounts of material might be dislodged during various experimental manipulations. In this case, the cornea remains in place on the balance during the drying step, which minimizes errors due to transfer loss. The drying is slow, such that the hydration gradient across the stroma should be small. The optical and geometric thickness measurements are made across the entire thickness of the intact stroma, which helps to reduce the relative measurement uncertainty for an instrument with a given thickness measurement resolution, and gives a value that represents the average refractive index over the full thickness. Finally, the light source is a laser, with a well-defined central wavelength and relatively-narrow spectral bandwidth. This helps to define the proper input parameters for Gladstone-Dale models used to describe the data.

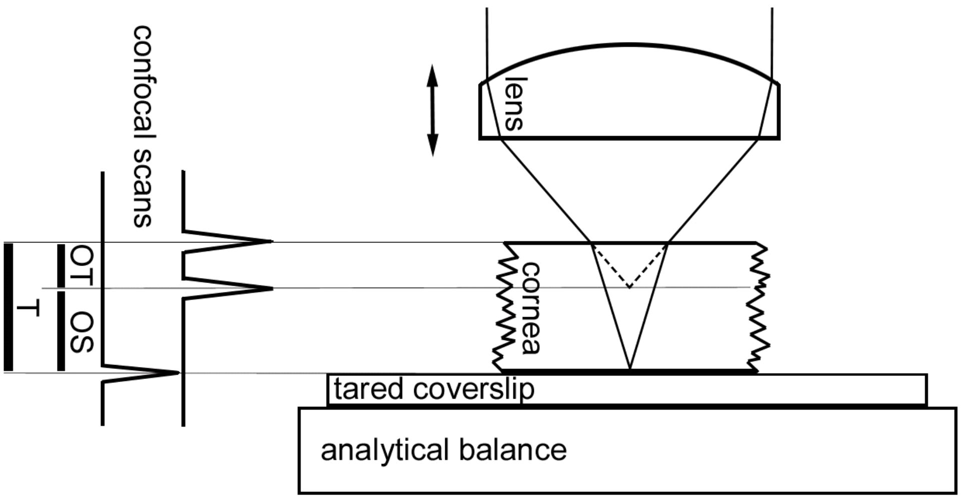

Fig. 14 - The optical and geometrical thickness of a cornea, as well as the mass, can be measured simultaneously using a Z-scanning confocal sensor. The geometrical thickness (T) is simply the difference between the top corneal surface position and the previously-measured position of the coverslip. The optical thickness (OT) is the distance between the top corneal surface position and the apparent bottom corneal surface position. The apparent position of the bottom corneal surface is shifted from the actual position, due to refraction of the measurement beam.

Because of this unique measurement system, the relationship between thickness, refractive index, and hydration can be determined experimentally with high accuracy and precision. A further extension, not undertaken in this work, would be to perform composition analysis on the final dry buttons in order to determine the final relative dry collagen and ground substance volume fractions.

c. Results and discussion

i. Corneal thickness as a function of hydration

As expected, measured corneal thickness was a linear function of hydration. However, for this group of rabbit corneas, the average dry thickness (91.7 microns) and physiological thickness (458 microns) differ from the implied values for dry corneal thickness given by the Hedbys-Mishima equation using their stated parameters (42 microns and 362 microns for small rabbits, 72 microns and 442 microns for large rabbits). The dry thickness of these corneas was also confirmed using calipers.

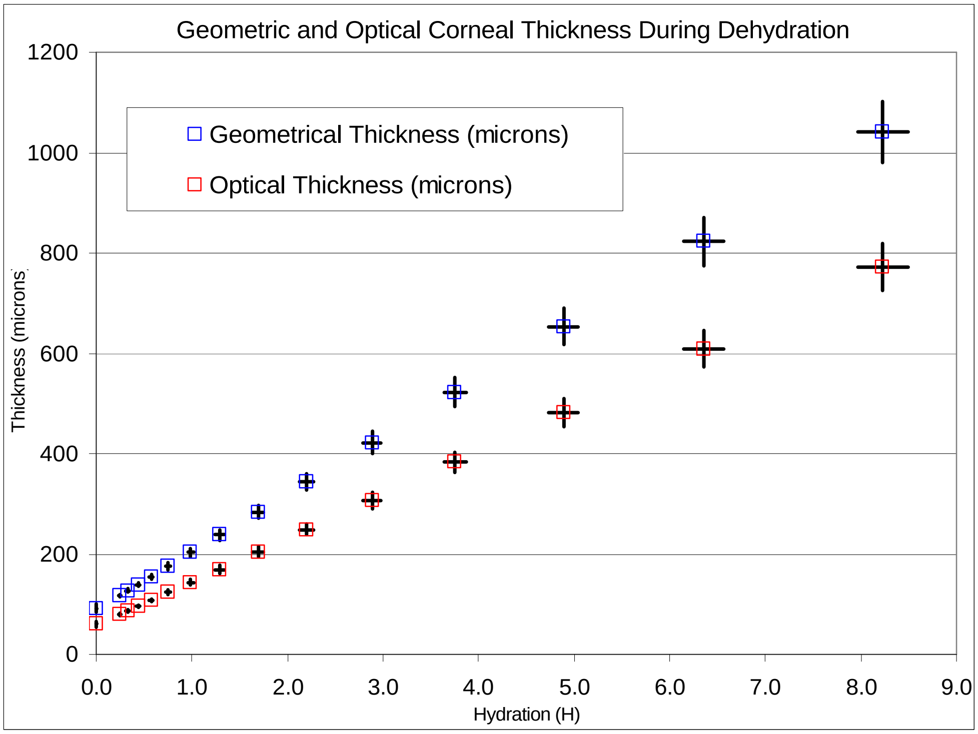

Fig. 15 – The geometrical thickness, optical thickness, and mass of corneal buttons was measured during slow dehydration. The hydration parameter H was calculated after a final dry mass was reached. Data from ten corneal buttons was pooled at each timepoint in the measurement series. A final dry thickness of ~90 microns was reached. While the data do show a very linear relationship between hydration and thickness, the slope and intercept values are different from the original parameters given by Hedbys and Mishima for large and small rabbits. Those parameters were based on measurements of corneas over a smaller range of H that did not extend to complete dehydration. Additionally, a different type of pachymeter instrument was used, and no correction for a shift in refractive index was made in the original Hedbys-Mishima data.

While the measured average value for the dry corneal thickness significantly exceeds the large-rabbit dry thickness value implied by Hedbys and Mishima's parameters, it is worth noting that the Hedbys-Mishima parameters were derived from a linear fit to a dataset that included no corneas below H=2, and the implied dry thickness value is therefore an extrapolation.

Also, because the original Hedbys-Mishima thickness measurements were made using a Maurice-Giardini optical pachymeter with no correction for hydration-dependent index shifts, the absolute thickness values they reported will be lower than the true geometrical thickness for hydrations lower than H=3.2, and slightly higher for hydrations above H=3.2. If this effect were compensated in the Hedbys-Mishima data, the implied dry geometrical thickness value would increase by approximately 10%.

Based on these observed differences, it is suggested that Hedbys-Mishima parameters obtained from literature sources should be used with caution in future quantitative modeling. The Hedbys-Mishima equation has been applied in many experiments over the years, usually to gauge one variable (hydration) by measuring another (thickness). Often, the method of thickness determination was not equivalent to the method used to originally obtain the cited parameters. Presumably, the animals in this study had different baseline thicknesses than the population used by Hedbys and Mishima. Based on the observed difference of baseline dry thickness among study populations and thickness measurement systems, it is argued that literature values for Hedbys-Mishima parameters should be applied to an uncharacterized experimental system with caution.

ii. Refined Gladstone-Dale model of refractive index and hydration

Corneal refractive index as a function of hydration was measured over a hydration range of H=8 to H=0. The shape of the drying curve is similar to previously-published data and models (Meek, Dennis, & Khan, 2003). A new wavelength-dependent Gladstone-Dale model of the refractive index-hydration relationship was developed.

In this new model, wavelength-dependent component refractive indices for water, dry collagen, and dry matrix are incorporated. The water index is taken from the IAPWS dispersion equation. The dry collagen index is taken from a published Cauchy equation for gelatin. The dry matrix index is taken from a derived Cauchy equation which completes a 3-component Gladstone-Dale model of the known corneal chromatic dispersion at physiological hydration.

Hydration-dependent volume fractions for each component are calculated by using literature values for the density of dry corneal collagen (*?=1.41 g/mL*) and dry extrafibrillar matrix

(*?=1.06 g/mL*), as well as the known density of water (*?=1.00 g/mL*) and the known ratio (15:7) of collagen mass to extrafibrillar matrix mass (Leonard & Meek, 1997).

This new model describes the refractive data measured at 780nm very well over the full range of hydration. The model is fairly insensitive to small adjustments of most defining empirical parameters. However, the refractive index parameter for the solvent component is crucial, because it determines the asymptotic behavior at large hydration values, and because it comprises the largest single volume fraction component over most of the hydration curve.

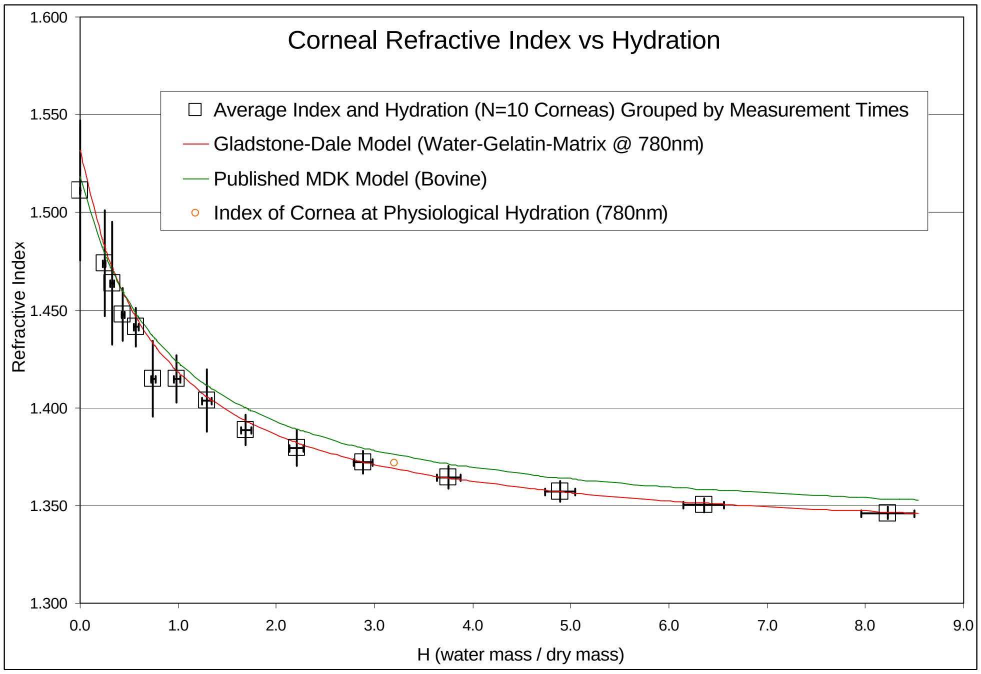

Fig. 16 – The refractive index of the cornea at 780nm as a function of hydration is calculated by dividing the measured geometrical thickness by the measured optical thickness. A Gladstone-Dale model, using the refractive indices of dry gelatin, water, and the putative dry matrix component and hydration-dependent volume fractions calculated from known component densities, provides a very good description of the measured data. The model published by Meek, Dennis, and Khan is plotted for comparison, as is the refractive index of a cornea at physiological hydration, taken from the Atchison Cauchy model. This value is higher, reflecting the difference between water and aqueous humor indices.

The refractive index of pure water at 20°C and 780nm (n_w=1.329) was used in the model of the measured rabbit data. This value, rather than the value of aqueous humor, was chosen based on the reasoning that the bathing fluid used for equilibration in these experiments more closely resembled water than aqueous humor. The refractive index of dry collagen, taken from the gelatin model, was n_c=1.534, lower than the value n_c=1.547 used by earlier researchers. The refractive index of dry matrix, taken from the derived Cauchy model, was n_p=1.528, significantly higher than the value of n_p=1.485 used by Meek and coworkers.

Compared to previously-published models, this model describes the measured rabbit cornea data more accurately. If input parameters (refractive index of solvent, collagen, and dry matrix) taken from published models (Meek, Dennis, & Khan, 2003) are used, this new model produces a nearly-identical curve for the hydration-index relationship as the earlier published model. It is possible that the new model is more accurate in some regard, or that previous published data (Meek, Dennis, & Khan, 2003) were measured with some systematic error, or both.

As an exercise, the new wavelength-dependent model was applied to this previously-published data. In order to apply the new wavelength-dependent model, a guess at the effective wavelength of the original refractive index measurements is necessary.

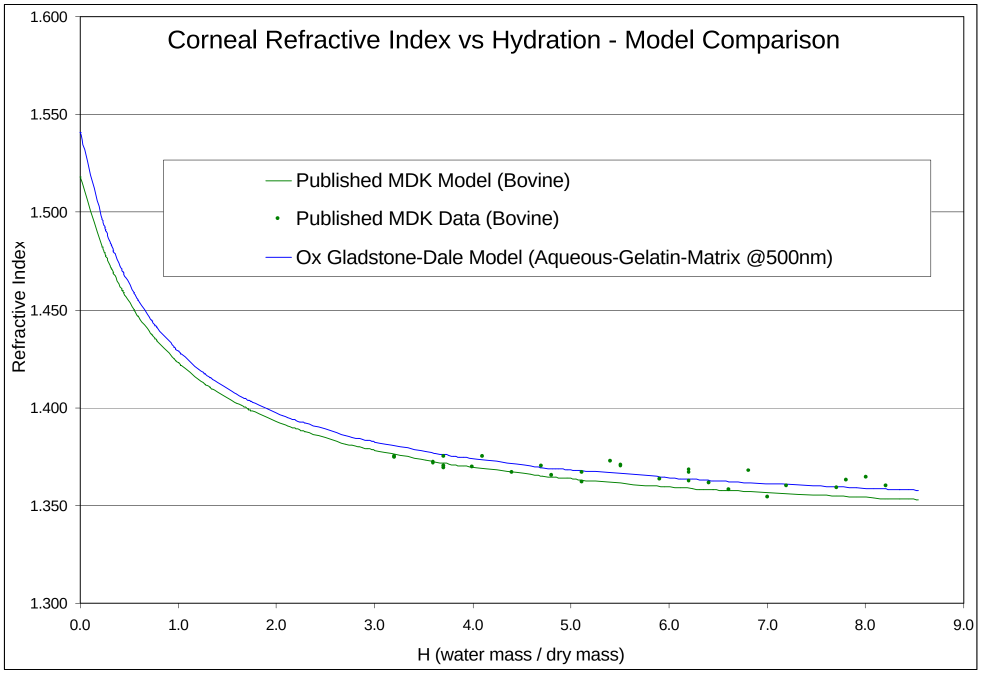

Fig. 17 – The model given by Meek, Dennis, and Khan predicted values for refractive index lower than the measurements indicated. A new Gladstone-Dale model, incorporating component indices (aqueous humor, dry gelatin, and dry matrix) at 500nm, gives a line that is closer to the middle of the measured data. The actual effective wavelength of the Abbe refractometer used to measure this data is unknown.

The instrument used to measure the bovine data was an Abbe 60 Series Refractometer (Bellingham & Stanley Ltd., Tunbridge Wells, England). The transmission light source for the original measurement was described in the paper as a bench lamp, while the reflection light source was an integral LED source, and the “detector” for this measurement was the observer’s eye. The light source used was an inexpensive table-top tungsten reading lamp purchased from a local store (Meek, personal communication, Sept. 26, 2007). The effective measurement wavelength was assumed to be 500nm, near the center of the visual system response. The previously-cited refractive index of dry collagen (n_c=1.547) also matches that of gelatin at this wavelength. The actual effective measurement wavelength for this instrument is unknown.

The Cauchy dispersion curve of gelatin was used to determine the index for dry collagen (n_c=1.547) at 500nm. The refractive index of dry matrix was taken from the derived Cauchy model of dry matrix (n_p=1.531) at 500nm. Finally, the refractive index of aqueous humor at 500nm was taken from the Cauchy model of aqueous humor (n_w=1.341).

The decision to use the wavelength-dependent index of aqueous humor (rather than water, as used to model the rabbit data) was made based on the fact that Meek and coworkers placed the corneal samples into dialysis tubing during hydration equilibration, and thus, the solutes responsible for the difference between aqueous humor and water may have remained in the corneas during swelling. Also, the solutions used to swell the corneas in the Meek experiment were sodium phosphate buffer solutions with a .03 micromolar NaCl ionic strength. The refractive index of salt water depends upon both salinity and wavelength. The value from the Cauchy model of aqueous humor at 500nm (n_w=1.341) was used in the new model describing the Meek bovine data, but the actual index of the solvent used in the Meek experiment remains unknown. The saline refractive index value used by Meek and coworkers (1.335) was cited from an earlier work. There was no refractive index measurement stated for this particular salt solution.

The choice of a proper refractive index for the solvent component of Gladstone-Dale models has been problematic. Leonard and Meek initially used n_w=1.333 (water at unspecified wavelength) in 1997. In the later publication (Meek et al, 2003), that value was replaced with n_w=1.335 (salt solution at unspecified wavelength and unspecified salinity, citing Farrell & McCally). For Gladstone-Dale models, the proper refractive index is the one that most closely approximates the component present for given measurement conditions. In the original Gladstone-Dale model for stromal refractive index (Maurice, 1957), the refractive index value of the solvent component was 1.335. This was meant to approximate aqueous humor at an unspecified wavelength. In the new model developed to describe the measured rabbit data, the solvent was modeled as water, since the equilibration bathing fluid was pure distilled water and no dialysis tubing was used to isolate the excised buttons. Not using dialysis tubing causes another uncertainty, as it is possible that soluble components in the matrix leached into the bathing solution. This leaching should represent a small error, since the refractive index of collagen and matrix (as modeled) are similar, and the absence of dissolved material (left in the bathing fluid at the time of button removal) will be compensated by the lower final dry mass which is used to calculate hydration. Additionally, a small volume of bathing fluid, 10mL, was used to equilibrate all of the corneal buttons as a group, and the temperature was kept low during immersion. The total amount of matrix dissolved into the bathing solution should be small under these conditions.

In summation, a new formulation of the Gladstone-Dale model for corneal refractive index as a function of hydration was developed. This model uses wavelength-dependent component refractive indices that reflect the dispersive properties of each model component. The model accurately describes the measured rabbit data, and provides an improved description of previously-published measurements from another research group, compared to their model.

The original purpose of the Meek-Dennis-Khan model was to examine the effect of hydration on scatter due to the mismatch between fibrillar and extrafibrillar indices. If an accurate model for the partitioning of water between fibrillar and extrafibrillar compartments is developed, the new model can be used to revisit this question. The new model presented may be further verified and extended in the future by directly measuring the chromatic dispersion of dried collagen and dried matrix material extracted from the cornea. Additionally, future Gladstone-Dale models of hydration-induced refractive index changes may benefit from chromatic dispersion measurements of the actual fluid used to swell the corneas, at pre-equilibrium and equilibrium time points.

A final conclusion deals with instrumentation. Optical pachymeters must multiply the measured optical thickness by an assumed refractive index to give the geometric thickness of the cornea. If the wrong refractive index is used, accuracy will be affected by a multiplicative error. If the wavelength of operation is poorly defined or unstable, this will represent an additional random source of error. Variations in corneal hydration will also affect the refractive index, but over the range of normal physiological fluctuations in hydration, this source of error is unimportant.

CHAPTER IV

RAMAN SPECTROSCOPY FOR HYDRATION MEASUREMENT

Several research teams have attempted to build instruments for measurement of corneal hydration by Raman spectroscopy. In each case, the systems required a legally-prohibited level of laser excitation power. The difficulties associated with Raman spectroscopy of the cornea were explored in order to determine key factors which limit system performance. An attempt was made to measure hydration using an instrument which differed from the typical approach (confocal sampling dispersive instrument with a cooled CCD), and used only low-cost components. The results were not completely satisfactory. Measurement of the composition of water-isopropanol dilutions using very poorly-optimized sampling optics was possible, and the low system noise floor implies a measurement SNR equal to the best published report is possible using a legal excitation power and low-cost components. Unfortunately, this measurement requires an impractically long measurement time, which precludes any real clinical utility. Nonetheless, with better detectors and a well-engineering optical sampling system, a filter-based system using two detector channels may yield system performance that is clinically useful and legally permissible.

a. Review of Raman scattering physics

Optical scattering is a light-matter interaction which may be elastic or inelastic in nature. In molecular scattering, the light interacts with the molecular electron cloud, resulting in a momentary transition from one energy state to another. The energy is quickly re-radiated as a new photon. In Rayleigh scattering, the molecule returns to the original state, and no energy is exchanged. The emitted photon has a wavelength very close to the original. Raman scattering is an inelastic scattering event with a discrete energy exchange which depends upon the specific inducible polarizability of a particular molecular bond. In Raman scattering, the molecule returns to a different vibrational state, and the difference in energy between the initial and final states is represented as a shift in the emitted photon energy relative to the incident photon energy. Correspondingly, the wavelength of the emitted photon is altered.

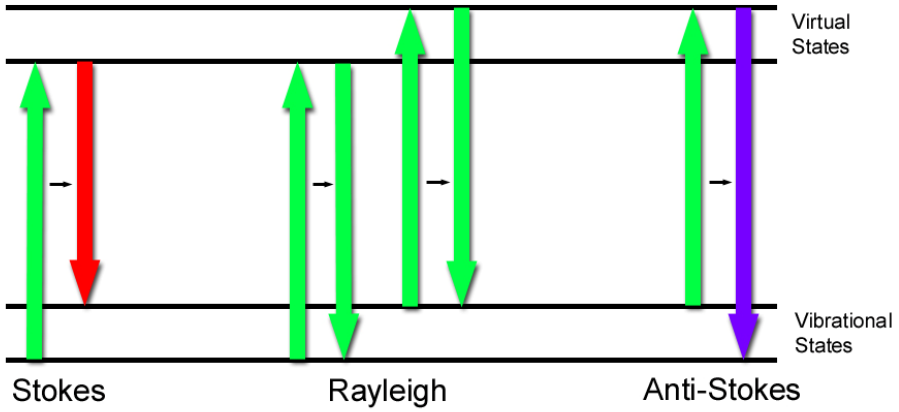

Fig. 18 - Raman and Rayleigh scattering can be represented by an energy diagram. Incident photons are depicted as upward arrows, while the resultant scattered photons are shown as downward arrows. In Rayleigh scattering, the system returns to its original vibrational energy state, and the scattered photon wavelength does not change. In Raman scattering, the system returns to a different vibrational state, and the energy difference is manifested as a shift in the scattered photon wavelength. The magnitude of the energy shift depends upon the characteristics of the vibrational system. However, the magnitude of the wavelength shift depends upon both the excitation wavelength and the energy shift magnitude. Excitation with longer wavelengths results in a larger shift in absolute wavelength, although the shift in energy remains identical.

Because the magnitude of the wavelength shift depends upon the type of bond, measurement of the spectral distribution of these scattered photons gives information about the molecular structure. This is the principle of Raman spectroscopy. According to the Boltzmann distribution, most molecules are initially in the lower vibrational state, and so the majority of Raman-scattered photons are shifted to a longer wavelength, producing a more intense Stokes-shifted spectral line. A smaller percentage of the molecules will initially occupy a higher vibrational state, and the resultant photon will have a shorter wavelength, producing a less intense anti-Stokes-shifted spectral line. The more-intense Stokes spectrum is usually the one that is measured, unless strong sample background fluorescence dictates measurement at wavelengths shorter than the excitation. The ratio of the Stokes and anti-Stokes spectra can also be used to measure the absolute temperature.

The scattered photon wavelength depends upon both the incident wavelength and the characteristics of the bond, but the magnitude of the shift depends only on the characteristics of the bond, and not the excitation energy. The intensity of the resultant spectral band is determined by both the incident excitation power and wavelength.

$$I \propto P_{in} \alpha^2 \omega^4$$

Equation 13

The intensity I is directly proportional to the incident optical excitation power P_{in} and is also proportional to the fourth power of the incident optical frequency ω and the squared polarizability α. Thus, Raman spectral intensity depends upon two parameters which may be freely chosen by the instrument designer, subject only to technical limits. In contrast, absorption spectroscopy requires that the incident photon energy must match the particular energy gap between the ground and excited state of the analyte molecule. Unfortunately, Raman scattering is an intrinsically-weak optical interaction. Raman scatter is approximately 1E-6 as intense as Rayleigh scatter for pure liquids, and only 1E-12 as intense for micron-sized particles. The choice of excitation wavelength is an extremely important design parameter for ophthalmic Raman instrumentation development, as there is interplay between this parameter and ocular safety, expected signal intensity, detector spectral responsivity, and background sample fluorescence.

The International Union of Pure and Applied Chemistry recommends standards for the presentation of chemical data. Among these standards, a frequently-ignored standard is given for the horizontal axis of Raman spectra. Because Raman spectra are shift spectra, with absolute wavelength or wavenumber values that depend upon the excitation wavelength, IUPAC states that the horizontal axis should be labeled Δcm^{-1}, although commonly, the delta symbol is omitted.

b. Raman spectrometer design variations

i. Excitation sources

In Raman's original observation of the Raman effect in the 1920's, filtered sunlight was the excitation source, producing a weak Raman spectrum even from large samples of very pure preparations of liquid benzene, with strong Raman scattering. The difficulty of producing a powerful, monochromatic excitation light source that could be focused onto a small sample volume was a considerable technical challenge which hindered the development of Raman spectroscopy for many years. High or low-pressure mercury lamps were the main light source for early instruments, often requiring multi-kilowatt power supplies and water cooling systems.

The first working laser was demonstrated by Maiman in May of 1960. In 1961, Townes proposed the use of lasers for Raman excitation sources. Weber and Porto first used a laser as a Raman excitation source in 1964. In 1966, Perkin Elmer developed the first commercial benchtop Raman spectrometer, which utilized a 4mW 632.8nm HeNe laser source and S20 PMT detector. Since those early efforts, laser sources have dominated practical Raman instruments. In the design of a Raman spectrometer, the choice of the excitation laser is crucial. The choice of excitation wavelength depends upon the expected sample autofluorescence, as well as detector spectral response for the wavelengths corresponding to the Raman shifts of interest. Additionally, certain types of lasers have superior beam quality, wavelength stability, and output stability. These factors are important, since excitation fluctuations can distort weak spectra.

ii. Sampling and collection optics

The details of Raman sampling optics can have a large effect on observed bandwidth, intensity, and even apparent spectral band position (Smith & Dent, 2005). Scattered light is typically collected at either 90° or 180° from the incident beam in most instruments. While Raman scattering is isotropic in direction, the two sampling angles are not always equivalent. This is because the 180° sampling strategy uses a confocal optical arrangement which greatly simplifies the task of matching the excitation and collection volumes. In the 90° collection geometry, care must be taken to ensure that these volumes overlap, which can be difficult if the sample is a solid, or liquid with an unknown or variable refractive index (Smith & Dent, 2005, McCreery, 2000). The numerical aperture of the collection optics is important. As Raman scatter is isotropic, a larger NA will result in a higher collected solid angle and a greater signal power.

iii. Dispersion mechanisms

Gratings and prisms are optical components used to split an incident beam into a spectrum. Gratings are comprised of a series of periodic structures, usually parallel ridges patterned on an optical substrate. These structures, with density d, will diffract an incident beam component λ (incident at angle Θ_i) into m orders at an angle Θ_m given by the grating equation.

$$d(\sin\theta_m(\lambda) + \sin\theta_i) = m\lambda$$

Equation 14

The multiple orders of the beam each receive a portion of the energy, and so energy is lost if only one detector is used. The efficiency of a grating can be improved by using a particular blaze angle, as well as a surface coating that is optimized for a particular wavelength range. Prisms, used in Newton's original demonstration of chromatic dispersion, are able to separate a beam using the material dispersion and law of refraction.

A Raman spectrometer may also use a Fourier-transform system to generate the spectrum. This type of Raman instrument is often used with high-power 1064nm Nd:YAG laser sources, with much lower sample fluorescence. This type of instrument is not relevant for ophthalmology for safety reasons.

One interesting possibility is to attempt a spectrometric measurement without dispersion. If the goal is quantitative spectrometric measurement, and the wavelengths of the spectral peaks of interest are known, fixed entities, it is possible to simply use bandpass optical filters, rather than dispersive elements, to measure the relative intensities of known spectral bands. A similar method has been used for fluorescence spectroscopy of the cornea. The success of this approach depends upon the existence of high-quality filters to reject Rayleigh scatter, while transmitting the Raman bands of interest. A second factor is the possibility that the detectors will sample a significantly-different portion of the excitation volume, since the measurement paths are not necessarily collinear, as in a dispersive spectrometer. This factor can be dealt with by suitable optical design, which may include a dichroic mirror element.

iv. Filters

Notch and edge filter development have enabled major improvements in Raman spectrometer performance. The need to reject intense Rayleigh light often necessitated the use of double or triple monochromator stages, with very low total light throughput. A single filter can now replace large and expensive monochromator stages, although the filter is suitable only for fixed excitation wavelengths. However, this restriction is minimal, since most instruments use only one laser excitation wavelength. While the major use of filters in Raman spectroscopy is Rayleigh rejection, filters can also be used in a design that performs a ratiometric measurement of previously-defined spectral bands, which achieves spectrometry without dispersion.

v. Detectors

Photodetectors are transducers which produce an electrical response to an input light signal. Photodetectors are crucial sub-systems of many scientific, medical, and industrial instruments, and many distinct types of detectors have been developed, each with particular strengths and weaknesses. These strengths and weaknesses make certain detectors unsuitable for use in clinical Raman spectroscopy, due to lack of spectral sensitivity, high noise, or fragility. In order to understand and compare these strengths and weaknesses, a review of photodetectors and performance metrics is necessary.

Noise-equivalent power is a figure of merit used to express detector sensitivity. NEP is defined as the optical signal power (in watts) that gives an electrical signal-to-noise ratio of 1 over a given electrical detection bandwidth. A detector with a smaller NEP is capable of detecting a weaker signal. NEP is also sometimes specified in watts per root Hz, which gives the power spectral density of the noise over a defined bandwidth. Related figures of merit include detectivity D (which is the reciprocal of NEP) and D*, which is the detectivity normalized over a specified detector unit area, typically in cm^2. However, NEP and D* are not the only relevant figures of merit for a detector. A measurement in the visible spectrum may be limited by the background and its noise, the shot noise of the signal, or the Johnson noise of the load (Hobbs, 2000).

The NEP of the silicon photodiode used in the current research is specified by the manufacturer at 14fW/rtHz at 20V reverse bias, while the dark current is 200pA at this bias. At 0V reverse bias, the dark current is 60pA, and so the NEP might be reduced by a factor of two.

Signal-to-noise ratios are a perpetual source of confusion in optical detection. This is because quantum detectors are square-law devices. Electrical power output from a quantum detector is proportional to the square of the incident optical power, because it is proportional to the squared photocurrent (Hobbs, 2000, Flanigan, 1995). An optical signal with a mean intensity of 10E6 photons per second has an SNR of 10E3 in a 1-second measurement due to photon shot noise statistics. However, if this light is detected by a quantum detector with a quantum efficiency of 1, the electrical SNR is now 10E6 in a 1-second measurement, proportional to the square of the photocurrent. Noise-equivalent power parameters take the square law into account, and so the parameter represents the optical power necessary to give a measurable electrical signal. Signal shot noise is rarely the dominant noise source, and so this square law behavior is rarely an issue.

Spectral responsivity is the ratio of the output current (A) to the optical input power (W) at a specified wavelength. Spectral responsivity depends upon the quantum efficiency, the photon energy, and any internal gain that the detector produces. For photodiodes, responsivities are typically about .32 A/W at 400nm, and 1.24A/W at 1240nm. A silicon photodiode of the type used in this research has a responsivity of .52A/W at 778nm and .54A/W at 806nm.

The quantum efficiency is a key parameter for all quantum detectors, including photodiodes as well as photomultipliers. The quantum efficiency is the ratio of the number of electron-hole pairs generated to the number of photons incident, before any gain takes place. The quantum efficiency of PMT's tends to be fairly low in the near-IR (S20 photocathode QE is <.03 at 800nm), while the silicon photodiode quantum efficiency may be greater than .6 in this region (and even greater than .9 if coated with an anti-reflective coating). Detectors with internal gain, such as PMT's, can reduce the importance of subsequent circuit noise, but low-noise gain cannot make up for poor shot noise performance due to low quantum efficiency (Hobbs, 2000).

The temporal bandwidth of a detector is related to both the maximum response speed to an applied signal impulse, and the dispersion of the impulse response. The response speed for a photodetector is often given as either the rise time from 10% to 90% of the maximum response, or the 3dB cutoff frequency. The response speed is dictated by the physics of the device. In a photodiode, the response speed is a function of the terminal capacitance, as well as the drift and diffusion times of the generated carriers. These factors depend upon the junction geometry and externally-applied biasing, as well as the load resistance that the generated current must flow through. The diffusion time for carriers that are generated outside of the depletion layer may be several microseconds. The response waveform of a photodiode may exhibit a response with two distinct time constants if the diffusion time is large and many carriers are generated outside the depletion layer. In a photomultiplier, the response speed and pulse dispersion are a function of the potentials between dynodes, as well as the number of dynode stages and details of the dynode geometry. The waveform of the detector's response, relative to the input optical waveform, is complex and difficult to calculate from first principles, especially of details of the device fabrication are unknown.

Photodetectors are classified by physical mechanism of operation. Two broad categories are quantum detectors and thermal detectors. Quantum detectors operate upon principles that directly convert photons into an electrical signal, while thermal detectors rely upon a multi-step process of absorbing photons, dissipating thermal energy, and electrical transduction of a corresponding increase in temperature.

Photoemissive detectors are based upon the external photoelectric effect. Photocathodes are thin layers of photosensitive semiconductor material. When a photon strikes the photocathode, an electron-hole pair is generated. The electron follows a zig-zag path as it scatters within the atomic lattice. The kinetic energy of the electron is proportional to the incident photon energy. Depending upon the kinetic energy level, this scattering may be either quasi-elastic or inelastic. In the quasi-elastic scattering regime, each scattering interaction removes only a small amount of kinetic energy, and therefore the scattered electron may undergo several scattering interactions and follow a long mean free path while maintaining a high kinetic energy. In the inelastic scattering regime, each scattering event removes a large amount of kinetic energy, generating secondary electron-hole pairs with low kinetic energy. If the photoelectron reaches the photocathode surface, it may then be emitted (externalized) into a vacuum, provided that its kinetic energy at this point exceeds the work function of the photocathode. This free electron is then drawn through the vacuum toward an anode (or dynode) held at a higher potential. The rate of electron emission is proportional to the intensity of the incident light (Donati, 2000).

The physics of the photoelectric effect determine several significant aspects of a photoemissive detector's performance. The detector will be insensitive to long-wavelength photons, since they possess insufficient energy to overcome the photocathode material's work function threshold. The detector will be insensitive to short-wavelength photons, since they generate electrons which quickly lose kinetic energy through inelastic scattering before they can be emitted at the surface. Between these extremes, the detector will produce an emitted electron for an incident photon with a certain statistical probability (Donati, 2000).

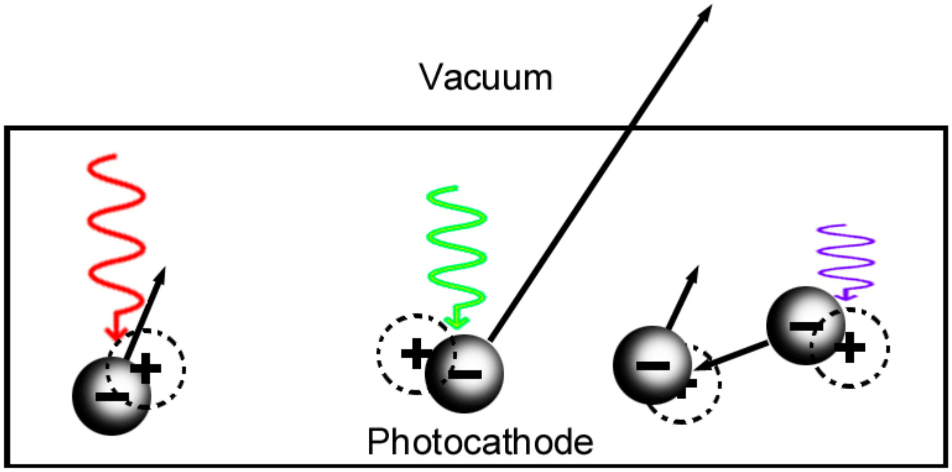

Fig. 19 - A photocathode's spectral response is governed by the work function threshold (which limits long-wavelength sensitivity) and the inelastic scattering regime (which results in secondary electron-hole pair generation). Photoelectron emission is only efficient over a range of intermediate incident wavelengths. NEA surface treatments can effectively reduce the work function, improving long-wavelength response.

In order to improve long-wavelength sensitivity, the photocathode surface may be treated in order to generate n-type defects on a p-type substrate, forming a negative-electron-affinity (NEA) photocathode. This is accomplished by exposing the photocathode to a Cesium vapor, and allowing an atomic layer to form. This n-type layer lifts the conduction band of the semiconductor substrate above the vacuum level. Electrons which come near the surface are able to exit by quantum mechanical tunneling.

In a vacuum photodiode (phototube) device, the photocathode is simply placed in a vacuum tube with an appropriate window material, and an anode to collect emitted photoelectrons. The anode may be a thin wire between the window and a hemi-cylinder photocathode (reflection photocathode) or a grid or plate placed on the far side of a transmission photocathode, relative to the window. The photocurrent is typically converted to a voltage by passage through a large load resistor. Phototubes can have a large active area, low dark current, and a large useful spectral range from 100nm to 1100nm with proper photocathode choice. The intrinsic response time is sub-picosecond, and phototubes are still used for short-pulse laser measurements. However, phototubes are inferior to solid-state devices with respect to size, ruggedness, microphonic sensitivity, magnetic field sensitivity, and bias voltage requirements (Donati, 2000).

The photomultiplier tube, or PMT, is an important type of photodetector, invented in 1919 by Joseph Slepian. Slepian was a prolific inventor at Westinghouse, retiring with 204 patents, second only to George Westinghouse (361 patents). The photomultiplier was the most sensitive detector type for visible photons for over 80 years. A photomultiplier is a modification of the phototube that contains a system of dynodes, which allow for a very high gain through charge multiplication, with minimal added noise resulting from the process.

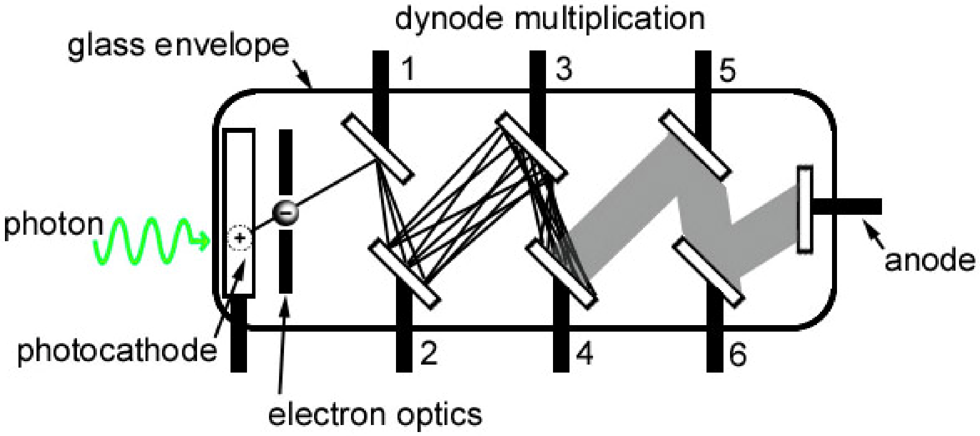

Fig. 20 - A photomultiplier tube provides a very high internal gain. An emitted photoelectron from the photocathode is accelerated towards a nearby electrode (dynode 1) that is held at a potential of ~100V, relative to the photocathode. A series of these dynodes, each at successively higher potentials, allows for a cascade multiplication and high gain. The resultant electrons are collected at the anode. External magnetic fields and vibrations can modulate the dynode multiplication efficiency, which makes the PMT susceptible to external interference and system crosstalk. A nearby chopper motor may thus introduce coherent interference.

Electrons emitted from a photocathode are accelerated toward the dynode, which has a potential difference of approximately 100V. The primary photoelectrons trigger the release of additional electrons from the first dynode through the process of secondary emission, and these electrons are, in turn, accelerated toward subsequent dynodes. The mean gain of each stage, which depends upon dynode material, incident electron energy, and temperature, can range from <10 to nearly 100. A photomultiplier typically employs 8 to 14 dynode stages, with an overall gain of 1E4 to 1E8 (Donati, 2000). The dynodes are often biased with a resistive voltage divider, which uses equal resistors of 100kO to 1MO for maximum linearity. Unequal resistors may be used to optimize time or amplitude resolution.

The photoemissive process itself does not damage the photocathode. Emitted electrons reach the anode, flow through a circuit, and eventually recombine with a hole in the photocathode, returning the device to its initial state. However, when illumination of the photocathode is high for an extended time period, the temperature of the photocathode rises due to resistive heating. At moderate temperatures, dopants undergo slow diffusion, resulting in a slow but progressive loss of performance. At higher temperatures, diffusion quickly destroys delicate structures such as NEA surface treatments, and a rapid and catastrophic performance loss is seen. The photocathode may also be damaged due to residual gas in the vacuum envelope. Emitted electrons may collide with gas atoms, producing positive and negative ions in the impact. The positive ions are then accelerated toward the negative photocathode, causing damage to the lattice upon collision.

Because of these effects, vacuum photodiodes and PMT's can be permanently damaged if exposed to room light. Vacuum photodiodes and PMT's also frequently have fragile glass envelopes. The typical MTTF for a PMT is 10,000 hours if it is kept out of bright light (Donati, 2000). PMT's are also subject to external magnetic interference, because an oscillating magnetic field will distort the path taken by electrons in the multiplication stages and modulate the signal. Sensitivity to magnetic fields is undesirable in systems which use an electric motor to modulate the signal, as this will create a coherent interference source. Finally, high-gain PMT's also require a stable, high-voltage power supply, which can be expensive. Despite all these limitations, PMT's are the most sensitive detectors for visible light, and cooled PMT's with photon counting electronics can detect average photocurrents on the order of 1E-20W, with several hours of integration time and suitable background correction. This represents an average photon flux of much less than 1 photon per second.

In addition to photoemissive detectors, a second class of quantum detectors uses the internal photoelectric effect to generate photocurrent. Internal photoelectric detectors, such as semiconductor photodiodes, are rugged and low cost devices with excellent linearity. Avalanche photodiodes and silicon photomultipliers are capable of low-level detection and photon counting because they have internal gain mechanisms, similar to PMT's.

Semiconductor photodiodes are the workhorses of photodetection, exhibiting excellent linearity over a large dynamic range, in a low-cost rugged device. Photodiodes are formed from a PN junction in a semiconductor, such as silicon. Photons which enter the junction are absorbed, generating an electron-hole pair, similar to the process in a photocathode. If this pair is generated in or near the depletion region of the junction, the large electric field will separate the carriers. The charge carriers then may form a current, flowing through a resistive path before finding each other and recombining. The kinetics of this process shows a two-part response, with a slow part comprised of the time necessary to diffuse into the depletion region, and a fast part comprised of the time to drift in the region of the field. The intrinsic response speed may be increased by increasing the size of the depletion volume. This is accomplished by adding a thick layer of low-doped "intrinsic" semiconductor, which forms the PIN photodiode.

The quantum nature of the photoconversion, and the high quantum efficiency, make photodiodes very linear transducers over a large dynamic range. The large bandgap of silicon makes thermal generation of carriers a rare event. Photodiodes become nonlinear when saturated with optical powers on the order of milliwatts, but a tightly focused beam at the microwatt level can also give nonlinear behavior (Hobbs, 2000)

Photodiodes are current sources, and most electronic instrumentation is designed to measure voltage. The standard solution is to use a transimpedance amplifier (TIA) configuration, which allows a linear current-to-voltage conversion at a much higher bandwidth, using higher resistance than a simple load resistor, which forms an RC time constant with the photodiode junction capacitance (typically 10's to 100's of picofarads). Many hundreds of pages have been written regarding transimpedance amplifier design and optimization. A thorough discussion of the topic is given by Graeme (Graeme, 1996). A discussion of photodiode amplifiers optimized for measurement of weak signals at low bandwidths is given by Eppeldauer (Eppeldauer, 2000). In most discussions, the importance of tertiary structures and construction techniques are neglected. For example, in extremely-high-gain transimpedance amplifiers, special precautions must be taken in order to minimize surface and volume leakage currents. Ions and trace oils from fingerprints on components can create conductive paths sufficient for femtoampere to picoampere current flow between a potential of a few volts. For this reason, it is very important to prevent contact with the feedback resistor and amplifier package, as well as the PCB. Sensitive current measurements often require special PCB layout techniques, using guard rings on the board to surround sensitive pins. If possible, it is sometimes greatly beneficial to dispense with the PCB altogether, and use air insulation around the sensitive nodes. A recent practical discussion of special construction techniques for low-level current measurement circuits is given by Rako (Rako, 2007).

Avalanche photodiodes are analogous to PMT's in some ways. The avalanche photodiode is a semiconductor device operated with a bias voltage that is sufficient to impart enough energy to free carriers that they can generate further carrier pairs in a process known as impact ionization. This process of carrier generation continues, resulting in an avalanche of free carriers. The APD must be carefully biased to a voltage between 90% and 99% of the breakdown voltage, and stabilized in temperature. The gain possible from this mechanism ranges up to a factor of 100, but smaller gains are less technically-challenging (Hobbs, 2000). The multiplication process adds noise, and the amount of excess noise is device-specific. Avalanche photodiodes can be used for photon counting, by reducing the bias after each event to stop the avalanche using a quenching resistor. Avalanche photodiodes have higher quantum efficiencies than PMT's in the near infra-red, and are more rugged. The active area is typically very small, measured in square microns.

A very new type of detector is the silicon photomultiplier (SiPM), which has recently become available commercially. The SiPM was developed in Russia in the mid-1980's. The SiPM has a high gain (up to 1e6), low bias voltage (30V), and combines features of APD's with a greater dynamic range and gain. The SiPM is an array of APD cells, each operated in Geiger mode with a series quenching resistor. Each unit in the array can be fired by a single photon, and the sum of the output current pulses approximate the input photon flux characteristic. The gain is inherently high, and the noise is minimal. These detectors are currently most sensitive in the visible, with maximum QE's of 40% at 600nm and QE's of 20% at 800nm. Like the solid-state photodiode and APD, the SiPM is rugged and compact, and immune to magnetic fields. As this detector becomes more common, it will be of great use in clinical medical instrumentation that requires low-light sensitivity in the NIR.

Charge-coupled devices are the last important class of solid-state detectors. CCD's were used in all prior systems for Raman measurement of corneal hydration. A CCD is an array of silicon photodetectors that integrate photogenerated charge carriers on tiny capacitors. After the measurement time is finished, the charge from each detector element (pixel) is read out by an amplifier in series. The limits on CCD performance are set by quantum efficiency, dark current, readout noise, and transfer losses. The minimum transfer efficiency is typically 99.99%, but when raised to the 1024th power, this small loss compounds to 90.2% transfer efficiency. With the best available AR coatings, the overall quantum efficiency can approach 85% (Hobbs, 2000).

Thermal detectors are covered in this review to complete the discussion. Thermal detectors detect incident radiation indirectly, by electronically sensing the change in temperature caused by absorption and dissipation of radiation. Thermal detectors are important in certain regions of the spectrum, such as the far IR, where quantum detectors no not exist. Thermal radiation detectors include bolometers, thermocouples, thermopiles, pyroelectric detectors, and Golay cells, among others.

c. Quantitative Raman spectrometry

i. Sources of variation and error in quantitative Raman spectroscopy

While qualitative Raman spectroscopy can be achieved with only moderate care, quantitative Raman spectroscopy is very difficult, due to the many factors which influence the measured strength of the scattering signal (Smith & Dent, 2005). These factors include reproducibility of excitation and collection optics during sample repositioning, fluctuations and drift in the optical power output of excitation sources, and the presence of strong interfering fluorescent components (McCreery, 2002).

Because each Raman instrument may have a complex spectral response curve, the relative intensity of peaks obtained on one instrument may differ from those obtained on another.

The differences in spectral response may be due to spectral characteristics of the detector, the spectral efficiency curve of the grating, or the spectral transmittance of elements such as Rayleigh rejection filters. NIST has developed transfer standards for quantitative Raman spectroscopy based on traceable standard lamps and well-characterized luminescent glass samples.

ii. Use of internal standards

Several of the difficulties of quantitative Raman spectroscopy are dealt with effectively by using internal standards to normalize the spectrum. An internal standard will effectively compensate for excitation power fluctuations as well as unrepeatable sampling optics, as long as the signal band and standard band are derived from the same sampling volume at the same time (McCreery, 2002). The use of internal standards is implicit in all previous attempts at Raman spectroscopic measurement of corneal tissue hydration. In this case, the intensity of the water (OH) band is measured relative to the intensity of the collagen (CH) band. This ratiometric approach greatly simplifies the measurement task.

iii. Raman spectra of mixtures as linear superpositions of component spectra

The use of internal standards implies that Raman spectra of mixtures are roughly linear superpositions of component spectra. This is not strictly true, but is approximately so (Manoharan et al, 1992). The shapes of Raman bands may be distorted due to interactions (such as hydrogen bonding) with other molecular species present in the system. This effect has been specifically observed for collagen and water. However, if the major Raman peaks for collagen and water are integrated over a wide spectral band, small shifts in the peak shape and central wavelength due to interactions will not greatly affect the ratiometric measurement.

iv. The OH-stretch Raman bands of water

Water is perhaps the most important and frequently-studied liquid. The Raman spectrum of water has complex subtleties, but may be understood to some degree using a simple Hooke's law model of a triatomic molecule. Disregarding changes in electronic energy, the energy of a molecule can be divided into certain vibration modes. The number of vibration modes is set by molecular degrees of freedom. For water, three degrees of freedom are taken for translational motion in any molecule, and an additional three are taken for molecular rotation. The total number of vibration modes is equal to three times the number of atoms (N=3) minus the translation and rotation degrees (6), giving a total of three modes.

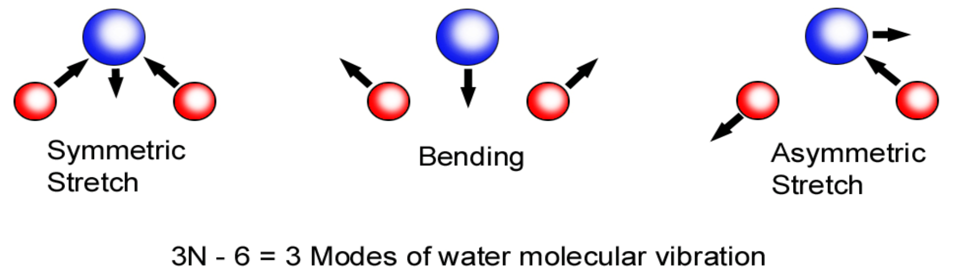

Fig. 21 - A Hooke's law model of water molecular vibrations shows that the symmetric stretch mode will cause a large polarization change and a small dipole change, indicating that Raman scattering from these modes will be strong, while infrared absorption will be minimal.

The symmetric stretch mode causes a large change in polarization and a small change in the dipole, which gives strong Raman scattering and weak infrared absorption. The asymmetric OH stretching mode gives very weak Raman scattering, which is expected due to the small change in polarization. For the simple vibrations described, one might imagine that the Raman spectrum of water is dominated by sharp features characteristic of high-Q resonation. Indeed, the Raman spectrum of water vapor is dominated by a very sharp peak centered at 3657? cm^{-1} with a FWHM similar to the linewidth of the monochromatic excitation.

Unfortunately, the Raman spectrum of liquid water is considerably more complex, due to hydrogen bonding and other complex intermolecular interactions, such as those known to occur with collagen (Walrafen & Chu, 2000). The Raman OH symmetric stretch peak of liquid water is very broad, and is often decomposed into several Gaussian bands centered at 3028, 3242, 3386, 3484, and 3626 ? cm-1. The center is generally regarded as 3400 ? cm-1. Additionally, collagen has a small overlapping amide A Raman band at 3330 ? cm-1. The main collagen CH vibrational modes extend from 2850 to 3040 ? cm-1, and are centered at 2945 ? cm-1.

Although they interact, the OH symmetric stretch peak of water and the collagen CH vibrational modes can be used to measure hydration, as long as the measurement is taken over a sufficiently broad spectral band.

d. Review of previous systems for Raman tissue hydration measurement

i. Dermatologic systems

Raman spectroscopy has been used for experimental measurement of tissue hydration in both dermatological and ophthalmic applications. The dermatological system is much easier to engineer, because of the higher allowable excitation power. One such system used 850nm excitation at 100mW power (Caspers et al, 1998), along with a custom Raman spectrometer that used confocal collection optics, a .75NA microscope objective, and a back-illuminated, deep-depleted CCD cooled with liquid nitrogen. A measurement time of 60 seconds was used.

ii. Ophthalmic systems

An ophthalmic system for hydration determination by Raman spectroscopy has not yet been developed, despite multiple efforts. A primary reason for this is the extremely low limits on allowed ocular exposure to laser power, determined by CDRH, ANSI, and ICNIRP guidelines.

One system (Bauer et al, 1999) was able to measure hydration with an estimated SNR of 25 using 15mJ of 632.8nm HeNe excitation. An earlier report by this team used 25mW of 514nm laser light for a 3-second measurement to achieve an SNR of 30 (Bauer et al, 1998). These systems have the lowest reported optical excitation power levels, but they are still orders of magnitude above the legal limits.

The CDRH sets a limit of 20 microwatts for a 30,000 second (continuous) intraocular exposure to 632.8nm laser light. A higher limit, 385 microwatts, applies for times less than 10 seconds. The ANSI limit for HeNe exposure is 17.5 times higher than the CDRH limit, allowing 293 microwatts for 600 seconds. As most systems use 10’s of milliwatts excitation power, performance must be improved by orders of magnitude before these systems can be used clinically.

In addition to the problem of a small signal, Raman spectroscopy of tissue is often plagued by a large background from autofluorescence. Primary fluorophores in the cornea include flavoproteins, pyridine nucleotides, and glycosylated collagen (Rovati & Docchio, 2004). These endogenous fluorophores have peak excitations from 450nm to 366nm, and peak emissions from 440nm to 550nm. The fluorophores tend to be concentrated in the epithelium and endothelium, which is favorable if stromal tissue is the measurement target, and confocal optics can be used to restrict measurement depth. The lens has a much greater level of autofluorescence than the cornea, and so some depth discrimination in the sampling optics is essential. Because of fluorescence, the use of blue or UV excitation light is discouraged, even though the absolute Raman signal would be stronger.

e. System design

i. Excitation source selection

The experimental Raman hydration measurement system uses 632.8nm excitation from a helium neon laser. This laser is low-cost, and provides an output that is stable in power and wavelength using well-engineered power supply components. A semiconductor laser source at a similar wavelength could also have been chosen, but to achieve similar stability, the cost would be greater. A laser wavelength between 630 and 650 will minimally-excite tissue autofluorescence, while placing the peaks of interest into the peak range of spectral responsivity for Si photodiodes.

The choice of excitation wavelength determines the central wavelength for the Raman spectral peaks of interest, arising from OH and CH bonds in water and collagen. For 632.8nm, the collagen peak (2945 ? cm^{-1}) will be centered at 778nm, while the water peak (3400 ? cm^{-1}) will be centered at 806nm. The collagen peak extends from 771nm to 781nm, while the water peak extends from 787nm to 838nm. These peaks are both fairly broad. Since each pixel of a CCD detector will have associated readout noise, the binning of spectrally-broad peaks onto fewer detector elements may result in a useful improvement in signal to noise.

Given some limited data on system parameters, wavelength-dependent absolute Raman cross sections (~1.1E-30 cm^2/molecule-steradian), and collection efficiency and transmission efficiency at an undetermined number of surfaces, the initial photon budget was crudely estimated. A 20-microwatt input beam power at 632.8nm gives 3.13E-19 J/photon and a photon flux of 6.4E13 photons per second. Water contains 3.34E10 molecules per cubic micron. Within a 724 cubic micron cylindrical excitation volume, there are 2.4E13 molecules. Collecting over a .22NA (.086 steradians), with detector efficiency of .55A/W, the power is on the order of 2E-15 watts.

ii. Optical configuration

As previously described, the instrument used filters, rather than dispersion, to give the measurement channels selectivity for each measurand. After some difficulty, the laser emission was cleaned up using a standard laser line filter (CVI Laser F10-632.8-4-0.50), which removed a considerable amount of long-wavelength light. Bandpass filters centered at 780nm and 810nm (CVI Laser F10-810.0-4-0.50 & F10-780.0-4-0.50) were used on each photodiode amplifier, following a custom holographic notch filter (Chroma) to reject Rayleigh scatter and low-shift Raman scatter from glass.

The original concept for the instrument involved a fiber measurement probe, confocal sampling optics, and an NIR-grating to separate the channels. However, this was not pursued due to unforeseen technical difficulties and time constraints. The Raman scatter arising from the excitation power passing through a 2-meter fiber was considerable.

The use of special filtering techniques commonly employed in Raman fiber probes is probably necessary for this application. Using a fiber-coupled probe with a high-NA collection lens, a higher collection efficiency could be achieved, compared to the prototype in this research.

iii. Detection system

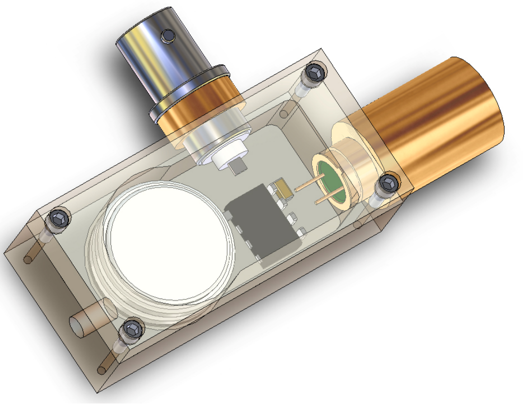

Fig. 22 - A photodiode transimpedance amplifier includes a photodiode, operational amplifier, feedback resistor, and internal Li-ion battery supply. The surface-mount 10G feedback resistor was soldered across the DIP pins to minimize stray capacitance. The critical signal node was wired in the air, to minimize leakage paths.

The detection system employed in this prototype required many iterations and significant experimentation. Several silicon photodiode detectors were explored, including the SFH-229 and the S-1223-01. These photodiodes have specified low NEP's (9E-15 W/rtHz and 1.4E-14 W/rtHz) and good spectral responsivities. The SFH-229 has a lower capacitance and a lower NEP, but the active area is lower. The final design used a collimated beam excitation and 90-degree collection with no external optics to collect the scattered light. A larger active area made the system less sensitive to alignment, at the expense of some noise. The transimpedance amplifier was constructed from a National Semiconductor LMC662 CMOS dual op-amp with rail-to-rail single supply operation and an input range including ground. The op amp is notable for ultra-low input bias current (2fA) and a gain of 126dB. The op amp package was an 8-pin DIP. The input-referred voltage noise has a 1/f characteristic with a knee at 200Hz and a magnitude of 20nV/rtHz RMS. The feedback resistor used was an Ohmcraft 10-gigaohm surface-mount chip in the 1206 package.

The construction of the preamplifier with a very high transimpedance gain must follow several rules. First, all leakage paths to the sensitive node must be eliminated. One strategy is to use only air-insulated wiring techniques, rather than PCB’s. Air is a very good insulator. The plastic used in the DIP package is also a good insulator, so long as it remains completely free of surface contamination such as fingerprints. Second, the stringent requirements for shielding dictate an internal battery supply inside the shielding. While this adds to the size of the detector module, it removes a major source of interference coupling from the external environment. Finally, BNC connectors are used to maintain the shielding integrity throughout the signal path.

iv. Signal recovery system

The signal recovery system used in this system is based on synthetic chopping and digital lock-in detection, discussed in a later chapter.

v. Performance

The measurement channel orthogonality is a measure of the independence of the channels, and the ability of the channels to reliably define a measured hydration stage. In an ideal system, measurements in the 780nm channel and 810nm channel would be fully independent.

$$P_{810nm} = [OH]a + [CH]b$$
$$P_{780nm} = [OH]c + [CH]d$$

Equation 15

At 20 microwatts excitation power and with no external collection optics, water gives a photocurrent of .57fA, a SNR of 29 in 2000 seconds of measurement time for this system. The delivered energy in this time is 20mJ. A previously-published system had an SNR of 30 using 25mW of 514nm light for 3 seconds, delivering 75mJ. Thus, this system, with very poor optics, was able to surpass the SNR/excitation power ratio and meet the power safety limit, but only by using a clinically-unrealistic measurement time. Using the ANSI limit for 632.8nm light (reduced by the chopping duty cycle), 293uw for 600 seconds, the SNR would be over 100.

The measurement channels in the current system are not fully independent. This may be due to incomplete spectral isolation, non-Raman background light, and distortions of the Raman peaks. The response of each channel to each component (in watts per molar concentration unit of the bond of interest) can be expressed as a linear system.

Ch. 1 (780nm)	**Ch. 2 (810nm)**	**2.0 mW 632.8nm excitation**
1.43E-16	5.21E-16	**watts per O-H bond concentration [M]**
1.03E-15	1.56E-16	**watts per C-H bond concentration [M]**

Fig. 23 - The channel response at a particular excitation power can be expressed in units of watts per molar bond concentration, with a partial response to each bond in each channel. The response ratios were 3.7 and 6.4.

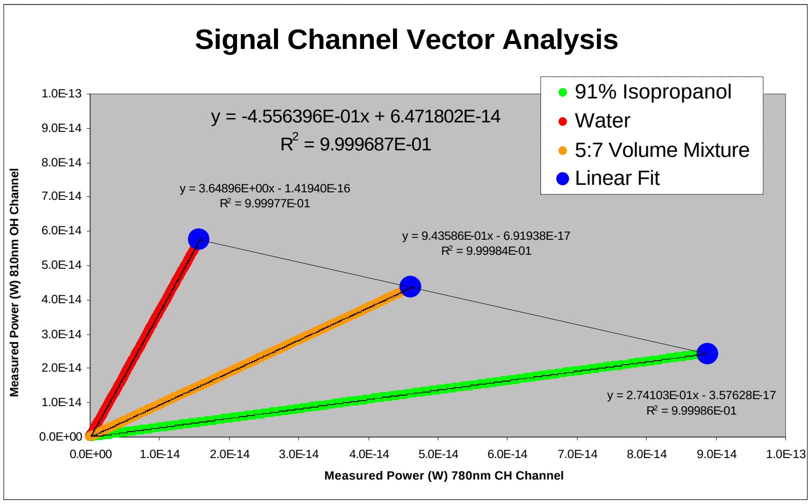

Fig. 24 - The response of the two channels to dilutions of water and 91% isopropanol can be plotted as vectors. With 2mW excitation (chopped at 50% duty), the signals were each on the order of 5E-14W in a 2000-second measurement time. The system was extremely linear in all respects.

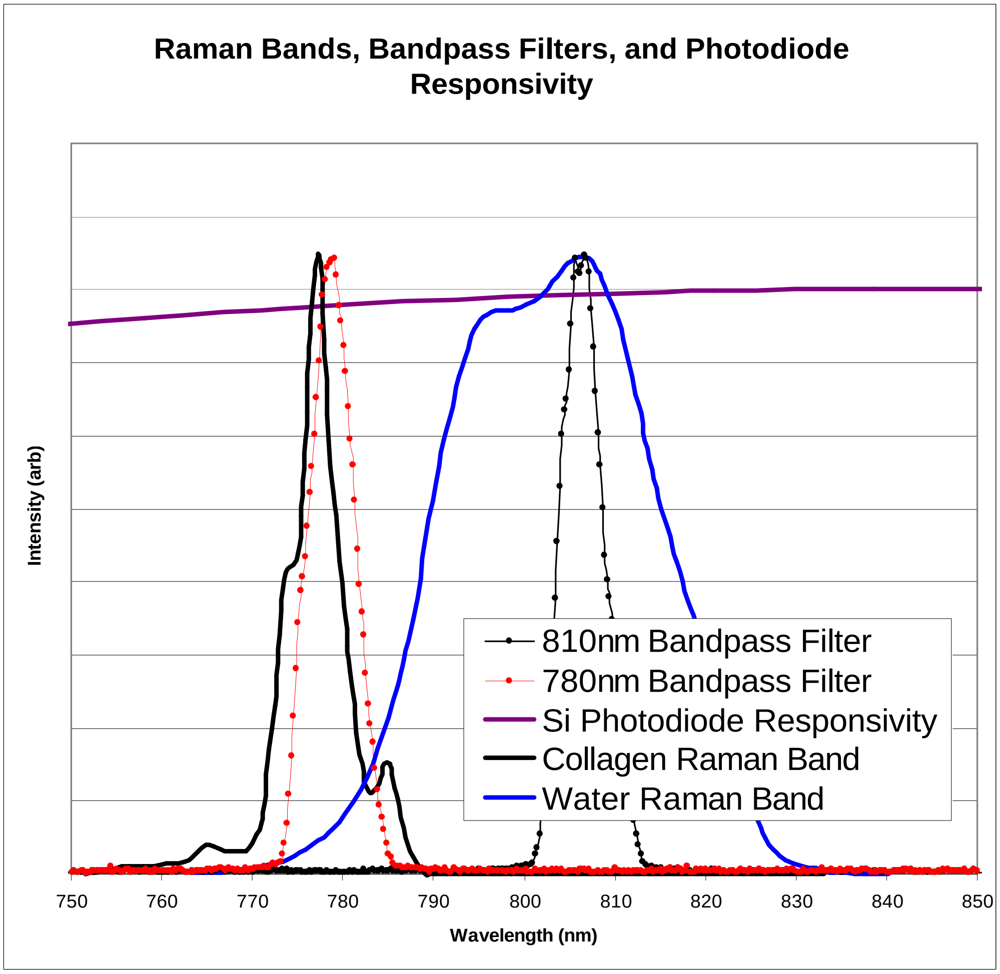

Fig. 25 - The band-pass characteristics of the optical filters used on each photodiode amplifier channel are shown, along with the Raman bands for water and collagen using 632.8nm excitation. The bandwidth of the 780nm filter is appropriate, but a similar filter centered at 778nm would collect a higher signal power. A much wider matching bandwidth could be used for the water detection channel, as the Raman peak for water is quite broad. The responsivity of silicon photodiode detectors is high across this region of the spectrum.

The system limit of detection may be evaluated from the response characteristic and the system noise floor. For a 2000-second measurement time, the limit of detection is roughly 1 molarity unit of the OH bond, corresponding to a water concentration of less than 1%, with 2mW CW excitation. A more-appropriate performance metric is given by the channel orthogonality coefficients, as they will reflect the signal power change per molar unit, which may then be compared to the noise floor for a given measurement time.

The system noise floor was very low, compared to previously-described chopped photodiode radiometers systems of a similar construction (Eppeldauer, 2000). The observed noise at 2000 second measurement times was 2.17E-17 watts. This low noise floor allows for very low-level measurements, but the long measurement time (30 minutes) is unsuitable for clinical applications. However, the current system has many sub-optimal components, and this study provides some encouragement that Raman spectroscopy can be used to measure corneal hydration using a legal excitation power, if improvements to the optics and detector subsystems are made.

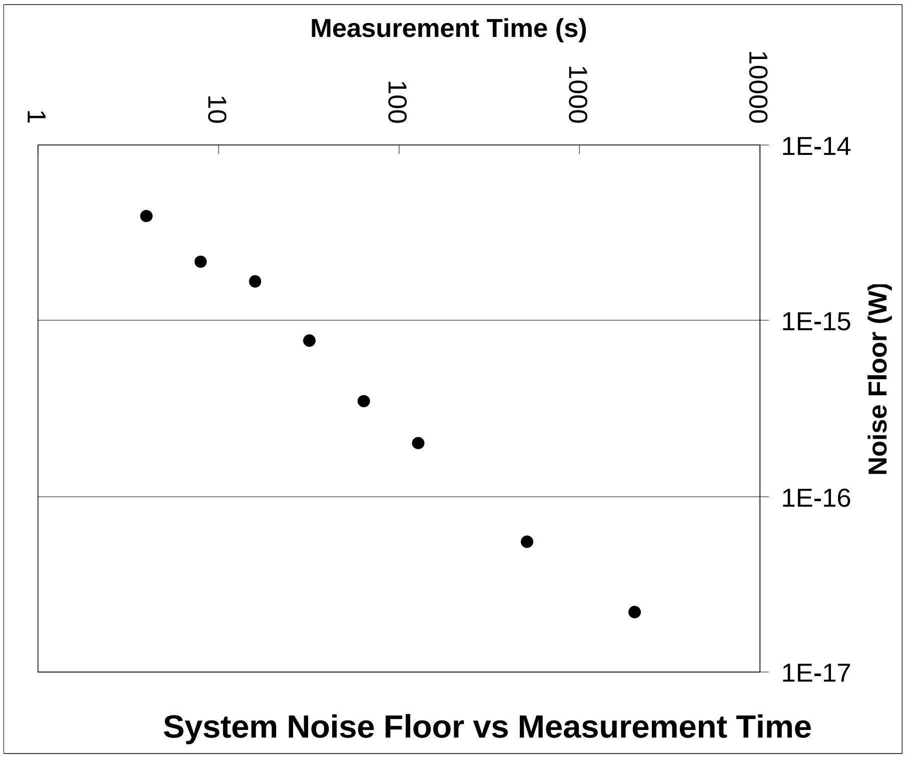

Fig. 26 - The system is capable of detecting very weak signals, but only by using very long measurement times. A measurement time of 2000 seconds has a noise floor of 2E-17 watts, using the synthetic chopper and lock-in amplifier described in a later chapter. This system has a dynamic range of 8 orders of magnitude, determined by the 7.2V battery power rails of the transimpedance amplifier. The digitizer has a 10V maximum range.

vi. Areas for future improvement

The optical system used in the current prototype system was very far from optimal. The system used no focusing or collection optics, resulting in a very low etendue, or collected solid angle. The first basic improvement that should be made to this system is in the collection optics. A commercial fiber-coupled Raman probe, employing an excitation line filter, Rayleigh notch filter, and perhaps the channel bandpass filters could be designed that would improve the sampling characteristics and collected solid angle by a large amount.

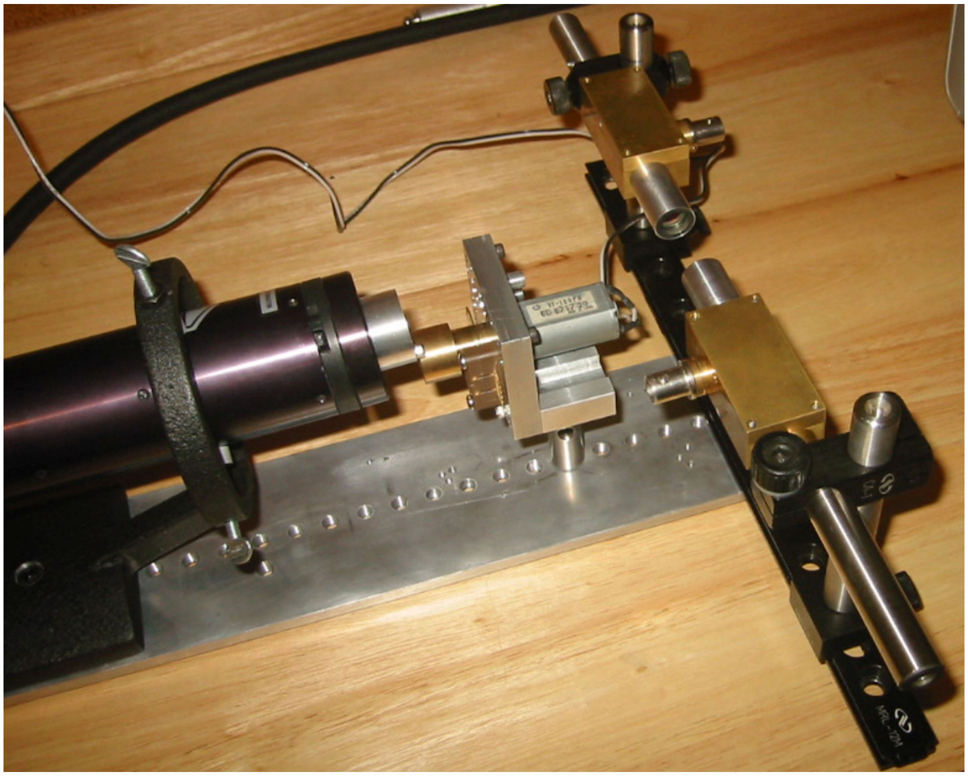

Fig. 27 - The prototype Raman spectral hydration measurement system is shown. Two photodiode transimpedance amplifiers with filters for Rayleigh rejection and Raman isolation are oriented opposite of each other, at 90 degrees relative to the incident beam. A synthetic chopper is used for modulation and lock-in detection. A HeNe laser excitation source with a 632.8nm bandpass filter is shown at left.

The second major improvement would be to use filters that are better matched to the Raman bandwidths. The 780nm collagen filter has a good spectral width, but is centered at a longer wavelength than the optimal location. The filter for the OH Raman spectral band could be widened by 3X, which would result in a much larger effective signal. A wider bandpass might allow more background and worsen the channel orthogonality, but this change is worth investigating.

In addition to these optical changes, the use of a well-designed dispersive spectrometer might also be warranted. This option was abandoned due to time constraints, but a dispersive system might increase overall throughput while allowing greater overlap in the sampled volume seen by each detector. Finally, the use of single-photon counting detectors, such as APD's or the new SiPM's, may reduce the noise floor enough to make clinically-useful measurements at a legally-permissible power level.

CHAPTER V

SYNTHETIC CHOPPING AND OPTIMAL PHASE-SENSITIVE DETECTION

This dissertation concludes with a chapter that describes a new method and apparatus for amplitude modulation and signal recovery, synthetic chopping. The synthetic chopping technique should have broad experimental applicability, because it enables high-performance lock-in signal recovery using only low-cost equipment. The synthetic chopper is a very simple modification of the traditional light chopper, but it is significant because it allows for straightforward implementation of the rules for optimal phase-sensitive detection, while at the same time reducing the importance of absolute mechanical precision and frequency stability in the chopper. This is because the effective bandwidth of the phase-lock loop is increased tremendously in the synthetic chopper method, reducing the aperture uncertainty to the unavoidable aperture jitter of the ADC. The result is a very robust and effective system that costs much less than a traditional chopper and lock-in amplifier. This system was used to enable the measurement of Raman signals with low-cost S-1223-01 photodiodes and un-optimized TIA pre-amplifiers, giving a noise floor of 2E-17W with a 2000-second measurement time. This compares favorably with results obtained by NIST researchers using hand-selected photodiodes, thermally-stabilized preamplifiers, and commercial digital lock-in amplifiers (Eppeldauer, 2000).

a. Noise, interference, cross-talk, offset, and drift

Noise and interference are fluctuations in a signal that obscure the signal information during measurement. Cross-talk, offset, and drift also complicate measurement. However, noise is unique among these sources of measurement fluctuation, since it arises from fundamental physics, rather than avoidable flaws in the system.

Noise is a fundamental process, intrinsic to the physics of the system. Noise is also usually defined as a mean-zero process, meaning that the average value of an ensemble of noise measurements will be zero. Finally, noise is also typically defined as a random process, with an intrinsically-unpredictable amplitude and phase at any instant. This last element of the definition is somewhat problematic, as unpredictability cannot be proven by experiment. While non-randomness in a pattern, signal, or process may be demonstrated by providing a concise description of the pattern, order, or information content, randomness cannot be demonstrated in a similar manner, because absence of evidence (of information content) is not evidence of absence. A process is not unpredictable merely because no one has published a method to predict it.

While it is generally impossible to prove that a process is random using experimental data, pure mathematical descriptions of probability distributions, such as Poissonian and Gaussian distributions, which stipulate lack of statistical correlation between events, have been shown to statistically describe certain measured processes very well. Thus, the idiomatic assumption of randomness in a process can have the useful consequence of producing an accurate description of its statistical behavior. Additionally, the physical theories at the foundation of noise theory, such as thermodynamics and quantum mechanics, have a very good experimental track record. No contradictory evidence has been found to show that noise (thermal or shot) is not a random process. The physical equations describing thermal and shot noise set a standard limit for system noise. Using this standard, the performance of a real measurement system can be evaluated.

Types of noise are often described by the power spectral density in each unit of bandwidth. White noise is one very important distribution. White noise from thermal agitation of electrons in a resistor was first observed by John B. Johnson in 1927 and explained theoretically by Harry Nyquist (Johnson, 1928, Nyquist, 1928). True white noise is a theoretical construct, because a noise source with finite power over an infinite bandwidth would require infinite energy. Real noise sources are regarded as white if the power spectral density is uniform over a defined range of bandwidths. White noise arises in systems with many degrees of freedom that are uncorrelated in time. Important sources of white noise include thermal (Johnson) noise and shot noise, which are discussed in detail later.

Another common noise power spectral distribution is has a power that is approximately inversely proportional to frequency over a defined bandwidth. This noise, called 1/f noise, is more problematic in measurement, because simply increasing the averaging time does not improve the signal to noise ratio, as it does for white noise. Theoretical explanations for 1/f noise are not as well-developed as those for white noise. Electronic 1/f noise was observed by Johnson in 1925 in an experiment designed to detect Schottky's theoretical shot noise. In the time since, a general physical theory has still not been formulated. Specialized physical models exist for some important electronic systems, but no general model explains the ubiquitous 1/f character seen in many systems in biology, astrophysics, and geophysics. It should be noted that 1/f noise is also a theoretical construct, since it would require infinite energy at DC (f=0). However, no real measurement can be taken at DC, because this measurement would also require infinite time, and so the theoretical problem is mathematically different than for white noise. Many long experiments (taking data over years and decades) show no limit for 1/f noise behavior.

Interference, in contrast to true noise, is the result of a non-random (typically man-made) energy source that couples into the measurement channel. Interference is, in principle, predictable. However, removing or compensating interference is not always straightforward, since the information which describes the interference is not always available. For example, 24-hour periodic variations in ambient light or ambient temperature which occur naturally due to daylight can be compensated in a measurement, but only if the measurement time is logged.

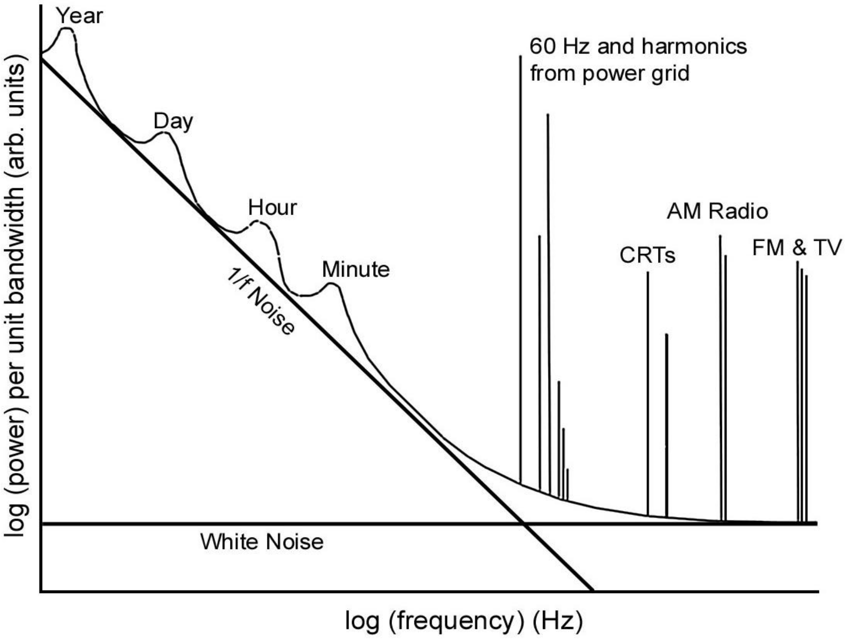

Fig. 28 - A signal may be obscured by noise and interference. Depending upon the physical origin, the spectral distribution of the measured power may show a white or 1/f distribution, which are added to give a fundamental noise floor for a system. Peaks may arise from natural and man-made sources which have a defined frequency. For example, temperature-dependent effects may show a daily periodicity, while the electrical power grid will give a sharp interference at either 60Hz or 50Hz and their harmonics.

Additionally, interference invariably adds true noise to the system, since any additional energy coupled into a system will have its own associated quantum and thermal fluctuations. These fluctuations will add in quadrature with noise already present in the system. For this reason, it is best to minimize interference as much as possible, by shielding the system from external energy coupling, using enclosures, baffles, apertures, thermal insulation, and electrical and magnetic screening and isolation techniques.

Cross-talk is a subset of interference, referring to interference that arises within part of the measurement system itself and couples back into another part. Because it is internal to the system, shielding and isolation is more difficult. However, compensation is easier, since the cross-talk properties of a system can be measured, and additional information necessary to remove cross-talk effects can be measured along with the signal of interest.

Offset is a static deviation from true zero in the system. A truly-static offset can be removed by subtracting a constant value that only needs to be measured once. Unfortunately, offset is seldom truly static. The more common situation is drift. Drift is a slow variation in the measurement baseline, but drift can also describe a variation in system gain. Drift effects can be minimized using special measurement techniques, particularly those which provide for alternating cycles of signal and baseline measurement. Ultimately, drift is reduced by using inherently stable systems, and then providing for active stabilization to remove the residual effects.

i. Photon and current shot noise

Shot noise is a type of white noise that arises in signals carried by uncorrelated discrete (quantized) events. Electronic shot noise was discovered by Schottky in 1918. The Schottky formula states that a constant current of I amps will show fluctuations i_{noise} equal to the square root of two times the product of the current, elemental charge q, and the observation bandwidth B.

$$i_{noise} = \sqrt{2qIB}$$

Equation 16

In a signal with full shot noise and a large number of quanta, the RMS magnitude of the fluctuation is equal to the square root of the average number of quanta observed in the measurement time. In a beam of natural light produced by a large number of independent elemental emitters, the photons also have shot noise. The average intensity is defined as the average number of photons which arrive at a given point in a given time interval, measured over an ensemble of infinite intervals. The actual number of photons that arrive in any specific time interval may be higher or lower than the average value, and the distribution will be a Poisson distribution.

$$f(x,y,t) = \frac{(yt)^x e^{-yt}}{x!}$$

Equation 17

The Poisson distribution, published in 1838 as part of a mathematical analysis of the criminal justice system, describes the probability of occurrence of a particular number of discrete events with a known average rate of occurrence in a particular time interval. The equation gives the probability that x discrete events with a known average frequency of occurrence y that will occur in a fixed time t, with the requirement that the individual events are absolutely uncorrelated and independent of one another. In the limit of a large average event frequency (>1000 events per unit time), the Poisson distribution is well-approximated by a Gaussian distribution with a mean y and variance y.

Shot noise is important in low-level photodetection. In most experiments, shot noise will determine the absolute theoretical limit for the signal to noise ratio. Most light sources have full photon shot noise, and detected photocurrents invariably have full electron shot noise, due to the stochastic nature of electron-hole pair generation. Other important shot noise sources include the shot noise of the dark current in the photodiode, and the shot noise of the bias current in an amplifier. Low-level photodetection systems using photodiodes are usually far from the photocurrent shot-noise limit, since other noise sources dominate. The measured SNR, compared to the SNR due to the photocurrent shot noise limit, is therefore a measure of the optimality of a photodetection system.

It is important to appreciate the requirement of event independence for Poissonian behavior. In the case of current in a macroscopic highly-conductive metallic wire, a train of charge carriers flow past a specified point in a highly-correlated and evenly-spaced manner, due to the Pauli exclusion principle. Inelastic electron-phonon scattering effects smooth out the flow of the discrete electrons. This current will not have full shot noise, despite the use of this example in many textbooks which teach shot noise statistics. However, current through a barrier, such as a p-n junction, Schottky diode, or tunnel diode, will show full shot noise. The quanta must be isolated, so that correlations do not emerge from interactions.

Photocurrent Shot Noise Limited SNR's

	Photocurrent (A)	**Noise (A/rtHz)**	**SNR**	**SNR**
	I	i_{noise}	B=1Hz	B=.01Hz
mA	1.00E-03	1.79007E-11	5.59E+07	5.59E+08
μA	1.00E-06	5.66069E-13	1.77E+06	1.77E+07
nA	1.00E-09	1.79007E-14	5.59E+04	5.59E+05
pA	1.00E-12	5.66069E-16	1.77E+03	1.77E+04
fA	1.00E-15	1.79007E-17	5.59E+01	5.59E+02
aA	1.00E-18	5.66069E-19	1.77E+00	1.77E+01

Fig. 29 - A limit on achievable signal-to-noise ratios, determined by the Schottky equation, can be calculated for given photocurrent signal levels. A 1-attoampere signal can just be detected in a 1Hz bandwidth (.5 seconds), but can be detected at an SNR of nearly 18 if a .01Hz bandwidth (50 seconds) is used. At low photocurrent values, it is very difficult to reach the shot noise limit using photodiode detectors, as other noise sources, such as the shot noise of the dark current or input bias current, begin to dominate.

ii. Thermal (Johnson) noise

In addition to shot noise, all photodetection systems have thermal noise, also known as Johnson noise (Johnson, 1928, Nyquist, 1928). Johnson noise is a voltage noise that is a consequence of thermal agitation of charge carriers in a resistor. The formula for Johnson noise voltage states that the voltage is equal to the square root of four times the product of Boltzmann's constant, the resistance (in ohms), the absolute temperature (in degrees kelvin), and the measurement bandwidth (in Hz). The theoretical explanation for Johnson noise, given by Nyquist, is a direct consequence of the equipartition theory of thermodynamics, which assigns a total energy per degree of freedom equal to Boltzmann's constant multiplied by the absolute temperature. Johnson noise can be reduced by reducing the temperature, the resistance, or the measurement bandwidth. However, it is difficult to substantially reduce the absolute temperature. Reducing the temperature from ambient (293°K) to liquid nitrogen temperatures (77°K) will only reduce the noise by a factor of two. Reducing the resistance across the terminals reduces noise for a transducer that produces a voltage signal.

$$V_{Johnson} = \sqrt{4k_b RTB}$$

Equation 18

Johnson noise may also be equivalently modeled as a current noise source in parallel with the resistor. In this case, a slightly-modified expression is used to calculate the RMS value of the current noise. This expression is used to analyze noise in current-producing transducers.

$$i_{Johnson} = \sqrt{4k_b TB / R}$$

Equation 19

For a photodiode with a given shunt resistance, one may observe that reducing the shunt resistance will effectively increase the Johnson noise, because the transducer produces a current signal (the photocurrent), rather than a voltage signal. The same is true in transimpedance amplifier design. A larger transimpedance feedback resistor will result in a higher signal to noise ratio, because the increased voltage noise is proportional to the square root of R, while the signal increases directly with R. Thus, a transimpedance amplifier should be used with the highest resistance that is compatible with the bandwidth and dynamic range requirements. However, if the feedback resistor is made larger than the photodiode shunt resistance (typically 10 gigaohms or more), then photocurrent will return to ground through the diode, rather than the feedback resistor. For this reason, photodiodes may be specially selected for high shunt resistance in the most demanding low-light applications (Eppeldauer, 2000).

In both forms of the Johnson noise equations, it is evident that noise is proportional to the square root of the bandwidth. Thus, in order to reduce the noise by a factor of two, the bandwidth must be reduced by a factor of four, and this can be accomplished by integrating four times as long. This simple technique is effective for reducing white noise, such as thermal or shot noise. Unfortunately, a more sophisticated approach is needed for dealing with 1/f noise.

iii. Noise with a 1/f power spectral density in photodetection

A unified discussion of the many physical models of 1/f mechanisms in various electronic devices has been developed by van Der Ziel (van Der Ziel, 1988). Collision-free devices such as vacuum photodiodes and the secondary emission multiplier stages always have a fundamental 1/f noise characteristic, while collision-dominated devices (including Schottky barrier diodes, PIN diodes, BJT's, Si JFETS) can show fundamental normal collision 1/f, Umklapp 1/f, intervalley scattering 1/f, and Hooge-type 1/f. The amount of 1/f noise, and the frequency above which white noise dominates, will depend intimately on the device geometry and fabrication, including impurities which vary from batch to batch. In addition to the analog circuitry, 1/f noise also enters the measurement process in the time base. Even the most stable ovenized quartz oscillators show frequency flicker noise. In order to achieve 1E-14 Allan deviation frequency stability over a measurement time of a few seconds, the crystals must be operated in the lower range of HF (5-10MHz) and must be hand-selected (Rubiola & Giordano, 2007).

Because significant 1/f noise is present in many detectors, electronic oscillators, and also in amplifiers and even mechanical systems, the noise spectrum of individual systems must be measured to determine the total noise spectrum of the system, which will indicate a break-frequency above which 1/f noise is minimized and white noise dominates. Experiment, rather than a simple equation, is the only way to determine whether 1/f noise is dominant in a particular system. Once the 1/f break frequency for the system has been determined, the signal information should be placed into this region of the frequency spectrum, which can be achieved with modulation.

iv. Quantization noise and ADC errors

Another source of noise is quantization noise. While digital signal processing has many advantages over analog signal processing, the input data from a sensor must first be digitized, and this process is the source of both noise and error. The use of digital post-processing is extremely dependent upon the quality of the digitized data (Hobbs, 2000). Quantization noise is additional noise added during digitization, a consequence of the limited number of discrete values that must approximate the measured value. The quantization process itself (in an ideal digitizer) adds a white noise of $12^{-1/2}$ ADU's (Hobbs, 2000). For example, in a 16-bit converter (65536 ADU's) with a 10V input range, the quantization noise is equal to 44 microvolts RMS. Real ADC's are far from ideal, and suffer from both integral nonlinearity and differential nonlinearity. In order to assess the quantization noise performance of a real system, the ADC datasheet lists a parameter for an effective number of bits (ENOB). This parameter should be used in calculations of expected system performance.

v. System noise due to phase jitter and frequency instability

Phase jitter and frequency instability are major limits of lock-in amplifier performance. Phase jitter and frequency instability may arise in the sampling process (ADC aperture jitter) as well as the external reference signal circuitry and the PLL system. Ultimately, the important characteristic for a lock-in system isn't the absolute stability, but rather, the ability of the phase-lock loop to follow the phase and frequency of the signal. For example, a perfectly stable lock-in amplifier with a perfect ADC clocking system, but a phase-locking mechanism which has low bandwidth, will fail to follow true changes in the phase and frequency of the signal, and measurement accuracy will suffer. However, if the signal and the external reference are extremely stable in absolute terms, then even a slow PLL with low bandwidth will not compromise the measurement. The effect of phase jitter and frequency instability must be evaluated in the context of the entire system, and subunits with intrinsic instability must be matched with high-bandwidth compensation and tracking mechanisms. The system performance is limited most by the weakest link.

vi. Interference and cross-talk

Fluctuations of a measurement due to interference and crosstalk are not fundamental, and can be minimized with improvements in the system design. Interference will manifest as narrow peaks in a measured noise spectral curve that extend above the noise floor set by white noise and 1/f noise. Interference is effectively removed by shielding sensitive portions of the system from the environment (Ott, 1988). For systems employing modulation and phase-sensitive detection, the effect of residual interference that is able to penetrate shielding may be minimized by careful selection of the signal modulation and sampling frequencies and sampling time. These parameters are set such that the signal and interfering sources fall onto different (relative prime) integer multiple of the system frequency resolution, and an integer number of both the signal and interference periods are present in the measurement (Zhang & Soukas, 1991).

Crosstalk will also manifest as a narrow peak rising above the broadband noise spectrum. However, crosstalk may be distinguished from interference because it will generally have a constant phase relationship, relative to the signal. Crosstalk is in many ways more problematic, as it may arise from a component “inside” the shield, and will “move with” the signal, which in many ways precludes techniques used to minimize interference.

The worst type of crosstalk for lock-in amplifiers is crosstalk that is synchronous with the signal itself, causing coherent interference. The crosstalk from the reference signal of a chopper into the signal path of a lock-in amplifier has been identified as the source of the noise floor limit in sensitive experiments to characterize avalanche photodiode noise (Lau et al, 2006). Similarly, the crosstalk from electronics used to drive resonant MEMS choppers has limited the performance of these devices (Wolffenbuttel & de Graaf, 1995).

Crosstalk is also problematic for signal recovery systems which use electro-optic modulators, because these systems require high-frequency high-voltage drive systems that broadcast energy at exactly the wrong frequency. Crosstalk is not removed by coherent averaging, because it is a coherent interference source. The front-end electronics are, by design, necessarily sensitive to the fundamental frequency. A stable crosstalk will present as a stable offset that may be removed by measurement and subtraction.

Fluctuations in the crosstalk energy (which are necessarily present) may have a 1/f nature, which is not reduced by further averaging. It is therefore desirable to minimize all electronic switching at the fundamental frequency, and design all analog reference signal circuitry to operate at a very high harmonic of the fundamental.

The perfect modulator places the moduland, and only the moduland, into the modulation frequency band. Because there is no need for electronic switching of a reference circuit at the fundamental frequency, the synthetic chopping technique offers the possibility of dramatically reducing this source of cross-talk.

b. Overview of signal recovery techniques

i. Introduction

Signal recovery techniques are special measurement strategies used to recover a small signal that is buried in a large amount of noise. Signal recovery techniques make use of additional *a priori* information about the signal in order to narrow the measurement bandwidth until the signal-to-noise ratio is suitable. Ultimately, the effectiveness of the bandwidth-narrowing process depends upon the amount of *a priori* information used, and the accuracy of that information.

ii. Coherent averaging

Coherent averaging is a signal recovery technique that can be used to recover an unknown waveform that is buried in white or 1/f noise. In fact, the coherent averaging process, with baseline correction, will remove all noise. However, crosstalk from another part of the system may couple coherently into the signal path, and will remain. In coherent averaging, also called multiple-time averaging or boxcar averaging, a repetitive signal is repeatedly measured (over n measurement scans), and the results from each measurement scan are added into the same bins. The resulting sums are then divided by the number of scans n, and the result is a signal waveform with an improvement in SNR proportional to the square root of the number of scans.

In order to use coherent averaging, the signal needs to be repetitive and a suitable reference trigger signal must be present. Ideally, the signal will be periodic (repetitive with equal periods) and the reference trigger signal will be tightly-coupled to the experimental system. If there is a coherent crosstalk source (which may arise from the reference signal, and is also tightly-coupled to the system), then this noise will pass through the coherent averaging operation, and distort the measured signal waveform.

The basic solution for this problem is to ensure that coherent electronic effects occur at a sufficiently high frequency, where the relative sensitivity of the signal path is lower. A more detailed analysis shows that phase-locked clocks with integral, relative-prime frequencies over the physical modulation cycle give the most noise immunity.

iii. Lock-in amplification

The lock-in amplifier is an important instrument that may be regarded as a special type of filter with a very narrow bandwidth that is able to track a signal of interest in the presence of noise. The lock-in amplifier was invented by Robert H. Dicke in the 1940's (Dicke, 1946).

Modulation and lock-in detection are generally used in extremely adverse signal-to-noise conditions, and so the technique has become synonymous with signal recovery. However, the technique also improves the precision and accuracy of measurements with a high input SNR, and so these concepts are worth exploring, even for less troublesome experimental systems (Meade, 1982). The lock-in amplifier is a well-developed concept, and numerous variations have been implemented depending upon performance and cost constraints.

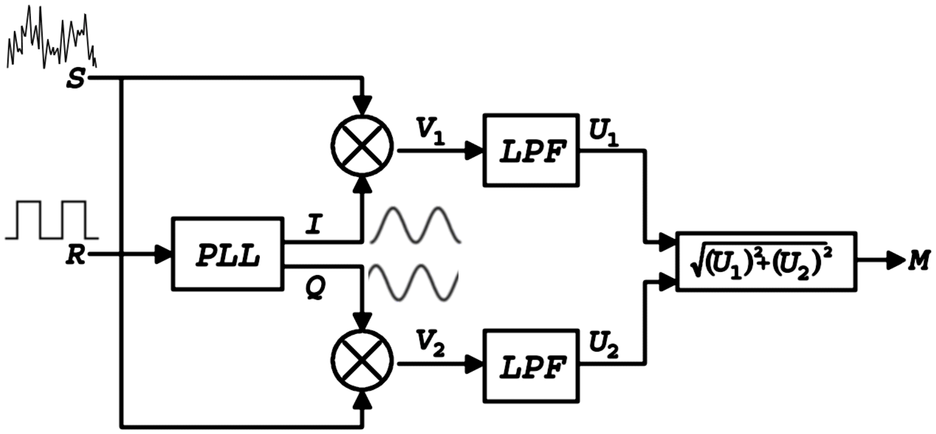

Fig. 30 - A typical dual-channel lock-in amplifier uses a signal S and external reference R to extract the magnitude of a single Fourier component of the signal. The external reference is input to a phase-lock-loop, which outputs sine and cosine waveforms (in-phase and quadrature internal reference waveforms). These waveforms are mixed with the signal waveform, and the output is low-pass filtered. These outputs give the amplitude and phase of the signal, relative to the external reference. The root of the sum of the squares gives the magnitude. The external reference signal may couple into the sensitive signal pathway, creating a source of coherent interference that raises the system noise floor.

The lock-in amplifier is able to track and measure small signals because of information supplied by the external reference. The external reference is a waveform that is locked in phase and frequency to the signal waveform. Because of this locking, the lock-in amplifier is able to properly rectify a signal, even if it drifts in frequency during the measurement. The external reference waveform is used to derive internal reference waveforms, which are mixed with the noisy signal to extract the true signal component. After mixing, the signal component has been rectified, and may be isolated with a simple low-pass filter operation. Lock-in amplifiers are generally insensitive to offset and baseline drift in the front-end amplifier, because the signal information is shifted away from DC and AC coupling into the lock-in is typically used. However, all stages, from the preamplifier to the mixer and end-stage filters, must have sufficient dynamic reserve such that noise does not overdrive a portion of the system into a nonlinear or saturated regime.

Lock-in amplifiers are usually classified by the implementation of the mixing and low-pass filter stages. The analog and digital versions of the lock-in amplifier are briefly reviewed, and a more in-depth discussion of the reference waveform will then be undertaken.

In an analog lock-in amplifier, the mixer is implemented as a switched amplifier. Information from the external reference waveform is used to switch the signal path between inverting and non-inverting amplifier paths, each with an identical gain. In its simplest implementation, a single amplifier is used, and the signal passes alternately through a gain of +1 and -1, switched with 50% duty cycles. This square-wave mixer has an input acceptance which includes the fundamental frequency and its odd harmonics. More complex analog mixers use a network of switched amplifiers to approximate a sine wave, and reduce the input acceptance of the odd harmonics. This is known as a Walsh mode demodulator. In either case, there is little flexibility in the input acceptance characteristic, because it is determined by the hardware configuration. Additionally, an analog mixer may suffer from drift, offset, and gain stability issues. Finally, care must be taken to ensure that both the signal and reference waveforms have exactly 50% duty cycles. Otherwise, an offset will result. For these reasons, analog lock-in amplifiers are now mainly used in applications (such as analog control loops) where digital lock-in amplifiers are not fast enough, due to limited digitizer speed.

The digital lock-in amplifier uses digital signal processing to implement both the mixer and the low-pass filter. The digital mixer is simply a digital multiplication of the digitized input signal with an internal reference waveform stored in memory. In principle, this internal reference waveform could have any shape, but in practice, it is typically either a square or sine wave. This restriction has the consequence of unnecessarily-limiting the input acceptance to the fundamental frequency or the fundamental plus the odd harmonics.

A more appropriate internal reference waveform is one which corresponds to the actual signal waveform at hand (Wilmshurst, 1985), with appropriate weighting according to the measured noise statistics (Doering & Harvey, 1989).

c. Optimal methods for lock-in signal recovery

Optimal methods for lock-in signal recovery have been presented by several authors. An emphasis is placed on gains that can be achieved by tailoring the internal reference waveform to the signal and noise characteristics that are actually present in the system, and on proper sampling to minimize aliasing and leakage effects which can corrupt sampled data. Proper sampling is particularly important for minimization of coherent interference effects.

i. Wilmshurst analysis on optimal signal recovery

Wilmshurst's text on signal recovery devotes several chapters to the design of phase-sensitive detectors (Wilmshurst, 1985). In particular, it is highlighted that for systems with white noise, the noise error is minimized when the internal reference waveform is identical to the average signal waveform, and matched in phase. The proof of this statement is given in section 9.3, and it is shown that the standard deviation of the noise integral is minimized when the reference and signal waveform are matched. If the waveform is periodic with an integral of zero over the period, the system will have no sensitivity to DC offset. Wilmshurst also presents a system for matched filtering in a PSD in section 9.6. In this section, it is shown that a system with non-white noise must be whitened by using a using a filter with an inverse transfer function relative to the noise spectrum, and that the optimal weighting function for this system results from passing the noiseless signal waveform through the same inverse filter.

ii. Doering-Harvey optimal linear phase-sensitive algorithm

In the analysis of Doering and Harvey (Doering & Harvey, 1987), a numerical method for realizing the optimal digital phase-sensitive detector is presented. In this paper, the optimal weighting waveform is shown to result from multiplication of the inverse of the measured noise correlation matrix with a noiseless signal waveform. This is equivalent to the Wilmshurst analysis, and it may be observed that the inverse of the white-noise correlation matrix (which is also the identity matrix), when multiplied by a noiseless signal waveform, results in the same waveform. It is shown both theoretically and experimentally that such systems are more efficient (30%-40%) at measuring data of a fixed quality, and is particularly beneficial for fast chopping with slow detectors. The authors conclude with four rules for optimal phase-sensitive detection.

The first rule is to chop the signal with as much of a square-wave shape as possible, to maximize the geometric factor. The second rule is to use the optimal window, as determined by the method described. The third rule is to refrain from over-filtering the signal in analog steps before the A-D conversion step. This is because the finite bandwidth of the detector has, in most cases, eliminated high frequency noise, and further filtering will only reduce the signal power. The fourth rule given is to sample at greater than twice the effective bandwidth of the signal and noise.

A final observation of the authors concerns maximizing overall system sensitivity. It is noted that since the optimal technique dramatically reduces the importance of the detector bandwidth in determining signal to noise (because the optimal method reduces the importance of the detected waveform geometric factor), it is better to use detectors with a lower NEP, even at the cost of reduced bandwidth.

iii. Zhang-Soukas rules for optimal phase-sensitive detection

Zhang and Soukas concentrate on rules for improving the PSD by improving the quality of digital data itself. This is accomplished by proper selection of sampling and modulation frequencies, relative to discrete interference peaks (Zhang & Soukas, 1991). In this work, the authors show the connection between phase-sensitive detection and Fourier analysis. Four rules are given which determine the proper frequencies. First, the fundamental frequency of the signal of interest should be placed on an integer multiple of the frequency resolution. This will serve to concentrate the signal energy into a given frequency component. The second rule is to place any strong interference signals on a different integer multiple of the frequency resolution. Again, this serves to concentrate energy from the interference sources into a different portion of the frequency spectrum. For example, if the noise is a 60Hz sine wave and the frequency resolution of the system is 20Hz, then the noise will be concentrated at 60Hz, and the signal should be placed at 20Hz, 40Hz, 80Hz, of another integer multiple of the frequency resolution. The third rule states that if the noise cannot be placed on a multiple of the frequency resolution, it should be placed as close to a multiple as possible. The fourth rule states that when there is sampling frequency uncertainty, a shorter measurement time will give rise to a smaller phase and amplitude error. For a given sampling frequency, a longer measurement time will give rise to finer frequency resolution. However, this finer resolution brings greater sensitivity to slight misalignments, drift, or jitter in the phase and frequency of the signal, relative to the sampling clock. This is a very important rule, because it stands in contrast to the general rule which states that a longer measurement time will always reduce noise.

A longer measurement time will, indeed, provide a better rejection of random noise in a system with perfect clocks and a perfectly-stable signal frequency. Since the system frequency stability is never perfect, a shorter sampling period may actually improve the signal to noise ratio because it reduces the sensitivity to phase and frequency drifts. Again, like 1/f noise, this is a property of the system which cannot be calculated from a simple physical model, and must be characterized by measurement.

d. Amplitude-modulated signal waveforms

In order to utilize phase-sensitive detection, the signal of interest must first be modulated. Several types of amplitude modulators are commonly used to modulate optical beams in instrumentation. Amplitude modulators vary considerable in the achievable frequency and depth of modulation.

i. Review of amplitude modulators and their performance limits

Optical amplitude modulators are a diverse group of devices employing simple or complex physical methods to predictably reduce the amplitude of a beam of light. In addition to mechanical choppers, there are rotating and oscillating mirrors, liquid crystal modulators, electro-optic modulators, and acousto-optic modulators.

Compared to mechanical choppers, electro-optic modulators and acousto-optic modulators have a much greater maximum modulation frequency. Electro-optic modulators are based on the Pockels effect or the Kerr effect, which use a high-voltage signal (400V-10kV) to modulate the polarization and phase of a beam. In conjuction with a polarization analyzer, they can be used to modulate the amplitude by modulating the polarization. The depth of modulation achieved by these devices is usually not more than 1000:1, due to polarization nonuniformity. A properly designed narrowband Pockels cell, which is optimized to operate in a resonant mode, may be driven in a resonant circuit which reduces the necessary voltage by a factor of Q. In this application, modulation frequencies of 100kHz to above 1GHz can be achieved from a 30V supply. However, these devices are expensive, have a low throughput (etendue), and require significant electronic circuitry operating at the fundamental modulation frequency, which predisposes these systems to crosstalk problems.

The major strengths of rotating mechanical choppers are low cost, high etendue, compatibility with a broad region of the EM spectrum (and even particle beams), low crosstalk, and extremely large depth of modulation. Major weaknesses of rotating mechanical choppers include moving (wearing) parts, vibration, size, frequency instability, and phase jitter. With proper design, these weaknesses can be brought to within acceptable limits.

ii. Random errors and deterministic errors

The modulation from a chopper has random and deterministic errors. Deterministic errors are deviations from the desired modulation characteristic that are mainly due to the limited precision of the chopper fabrication technology. For example, the desired mark-space ratio for a chopper is generally 1:1. This type of modulation has important properties for the measurement of weak signals. If the ratio of the “on” to “off” periods is significantly different from 1:1, then two effects will result. First, a more sophisticated system of phase-sensitive detection is required, in order to adjust the offset of the internal reference waveform such that the integral of the modulation waveform over a period is zero. Second, a slight statistical defect will be present, since the data will not equally sample signal and background. It has been shown that for very weak signals, any reduction in the uncertainty of the signal gained by increasing the “on” time is more than offset by an increase in the uncertainty of the noise due to reduced “off” time (Pearce, 1972). Deterministic errors may also be caused by diffraction at the edge of the chopper blade, or by effects due to a nonuniform beam intensity profile. These deterministic effects can be measured and compensated (Pelizzari & Postol, 1977, Naszodi, 1979, Vincent 1986).

Random errors in chopper modulation are the result of random errors in chopper frequency and phase, which will be present in any system due to effects of friction. The effect of these random errors can be reduced by increasing measurement time, as well as the rotational mass of the chopper. The effects of random errors can also be reduced by increasing the effective bandwidth of the system used to track the chopper motion, such that deviations from the mean rotational frequency are compensated in the output.

iii. Coherent interference and system crosstalk

As previously mentioned, coherent interference sources, which often arise as a result of system crosstalk, are problematic because they exist at the exact frequency, and often the exact phase, where the lock-in system is most sensitive. While it is often possible to measure the coherent interference waveform by blocking the signal, it is not always possible to measure the signal without the interference. A general strategy for minimizing the effects of coherent interference is to ensure that electronic switching occurs at a much higher frequency than the fundamental signal frequency, so that the receiver is less-sensitive. After this strategy has been pursued, it may still be necessary to correct the measured signal waveform, by subtracting the coherent interference waveform collected over an identical measurement time.

iv. Waveform measurement techniques

The waveform of the signal should be measured directly, in order to account for deterministic non-uniformities in the chopper, reference disk, or beam intensity profile. Also, the bandwidth and phase delay characteristics of the detector and amplifier must be measured. The waveform may easily be measured by coherently averaging a measurement of many individual revolutions in a high-signal state. After the measurement, this waveform should be corrected for offset, such that the integral over the full disk revolution is zero. This offset correction will allow the waveform to be used in phase-sensitive detection without causing a DC sensitivity.

e. Noise spectrum measurement

The noise spectrum may be measured in a manner similar to the signal waveform, with the system running but no sample in place. The sampled values from each revolution of the chopper may be regarded as vectors, and a correlation matrix may be computed from many of these vectors (tens of thousands). The correlation matrix will contain very complete information on the noise. Coherent interference will manifest as very high positive and negative correlations with a well defined phase relationship. Other areas of the matrix reflect the broadband noise characteristics, such as the presence of white noise in the measurement bandwidth.

f. Synthetic chopping

Synthetic chopping and lock-in detection is a modification of regular chopping that primarily involves the system of external reference signal generation, internal reference signal generation, and phase-locking between these signals. A transducer with high angular measurement resolution is used to directly trigger analog to digital sampling, replacing the role of a traditional hardware or software phase-locked-loop. The result is a PLL that follows deterministic and random errors in the chopper frequency and phase with a greater bandwidth than is possible using a typical reference generator. Because the synthetic reference signal is indexed, it may be used to give a 1:1 nondegenerate representation of the modulation characteristics of the chopper, which also incorporates the response characteristics of the detector as well. This indexed modulation waveform can then serve as the internal reference waveform for lock-in detection.

i. Description of the method

The synthetic chopping method relies upon a modified chopper with a high-resolution indexed rotary encoder. This encoder provides a reference signal that is mechanically phase-locked to the modulation waveform. Additionally, the synthetic chopper can be designed to take the rules of optimal phase-sensitive detection into consideration, but using an encoder with a large relative-prime number of cycles per revolution.

The synthetic reference signal is a signal derived from a physical feature, such as an optical or magnetic encoder track, that is distinct from the structure that modulates the incident beam. The reference signal is, ideally, derived from a feature with a relative-prime spatial frequency that is greater than twice the spatial frequency of the modulation feature, so that both they Nyquist sampling and optimal spectral sampling rules are satisfied. A relative-prime sampling strategy ensures that aliasing between harmonics of the modulation and reference does not take place at harmonics below the product of the relative primes. This is analogous to the relative-prime spatial frequencies of a Vernier gauge, which allows for unambiguous interpretation of the spatial beat frequency over the full unit cycle.

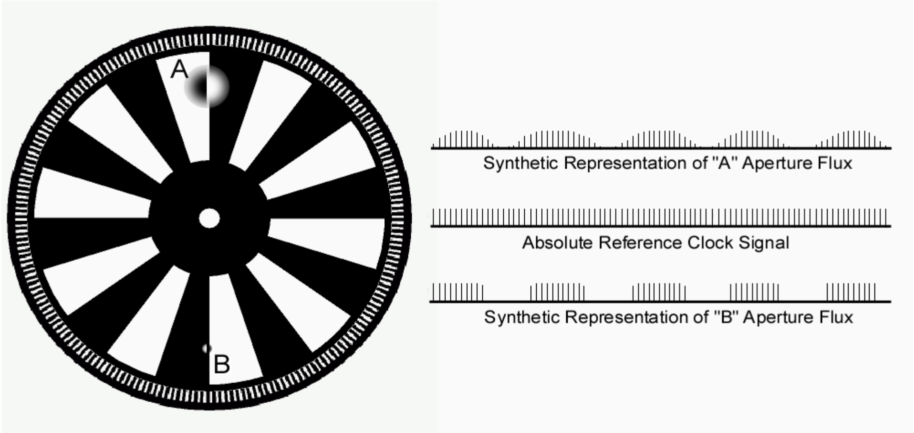

Fig. 31 - The synthetic reference signal is derived from an encoding structure with a high angular spatial frequency that is physically-locked to the modulating structure. The encoding structure has a known number of points, as well as an index feature. This encoding structure can be used to trigger sampling, which may be repeated to perform a very accurate coherent-averaging measurement of the modulation, that incorporates non-uniformities of the modulator and the encoder, as well as the beam and detector characteristics.

In order to perform optimal phase-sensitive detection, the waveform of the measured signal must be measured. This measured waveform incorporates complex optical effects relating to the beam profile, apertures, and chopper blades, as well as the phase and frequency response of the detector and amplifier circuitry.

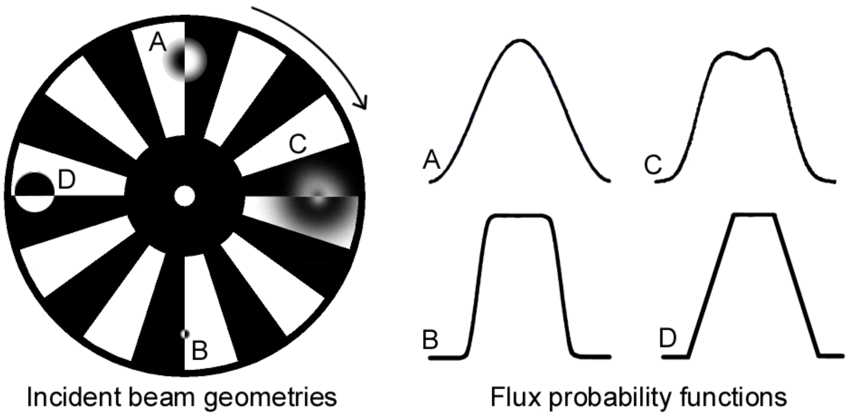

Fig. 32 - The modulation waveform achieved by a chopper depends upon characteristics of the chopper blade, as well as the incident beam and other apertures which may be present in the system. The flux probability functions are then input to a detector, which will further distort the waveforms due to its own unique bandwidth and phase delay characteristics.

The signal waveform can be measured, even for very weak signals, by coherent averaging techniques. Briefly, the signal and noise are sampled over many full revolution periods of the chopper, and the sampled values are added coherently into a one-dimensional array with a dimensionality equal to the number of reference signal transitions per period. Since broadband noise will have a random frequency and phase relationship, the coherent averaging operation will produce a waveform that is composed of the signal, as well as any coherent interference generated by the modulation process.

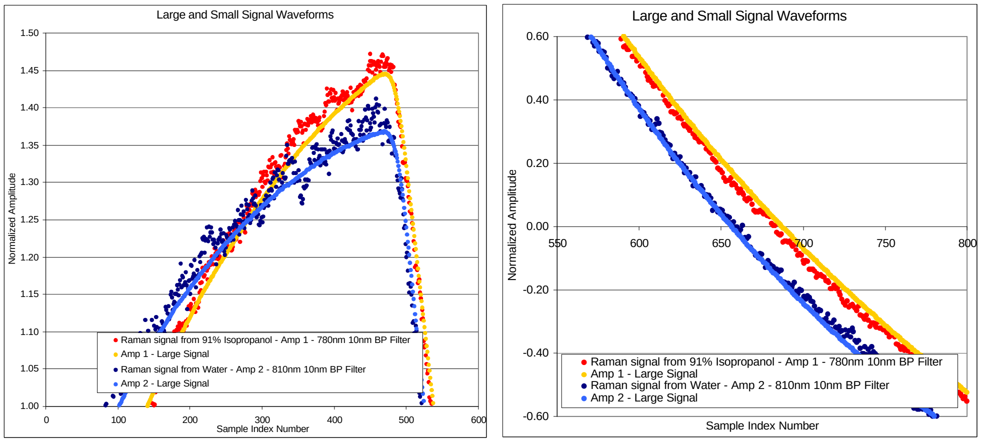

Fig. 33 - Modulation waveforms may be measured by direct coherent averaging of a weak signal, or by coherently averaging a stronger signal derived from an identical beam and aperture profile. The results for waveforms generated from nanowatt and sub-picowatt signals show the extremely linear behavior of the photodiode and transimpedance amplifier. The waveform from the small signal, although much noisier, is nearly identical to the waveform from the large signal. Departures from the large signal waveform are due to proportionately-stronger sources of additive noise (from quantization, shot noise, and thermal noise) as well as coherent interference from the modulator. Despite these effects, the difference in bandwidth between two photodiode and transimpedance amplifiers of nominally-identical construction is evident in both the strong and weak signal measured waveforms.

Using the synthetic chopper, coherent interference at the fundamental frequency is reduced, since no circuit is switching at this frequency. Instead, the switching occurs at a very high integer harmonic, which will couple into the signal path less strongly. Other sources of coherent interference, such as electrical commutation and magnetic noise or vibration noise from the motor, can be identified, measured, and eliminated from the signal waveform through subtraction, FFT-based weighting adjustments, or the noise correlation matrix method.

The simplest technique for coherent interference removal is to measure the modulation waveform by coherently averaging the modulated signal, and subtracting from this the waveform obtained by coherently averaging the noise for the same length of time. This will effectively remove stable additive coherent interference.

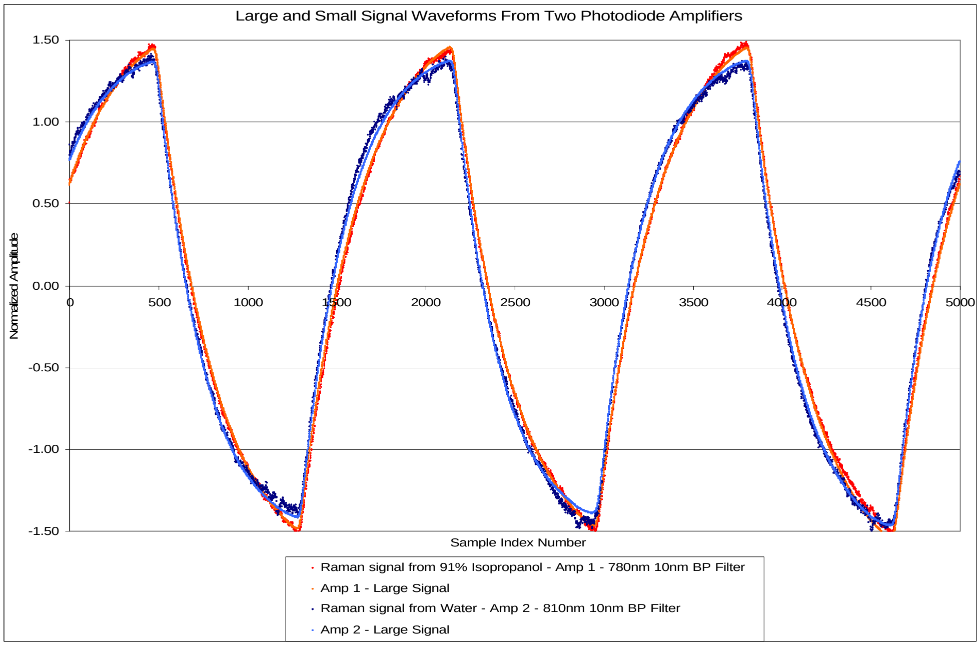

Fig. 34 - Modulated signal waveforms recovered by coherent averaging incorporate characteristics of the modulator, as well as the phase and frequency response of the detector and amplifier system. Shown are signal waveforms from coherent averaging of weak (Raman) and strong (nanowatt) signals measured over 32,768 revolutions. The waveforms, normalized for an RMS power of 1 over the period, show distinct differences in the bandwidth of two different photodiode transimpedance amplifiers constructed from the same components. The synthetic chopping system allows sources of coherent interference, such as electrical and magnetic interference from a motor, to be identified by their phase and frequency characteristics. A suitable strategy to remove or compensate the interference may then be adopted, and its effectiveness evaluated by repeated measurement. In the measured waveform of the weak signal from the second amplifier (dark blue points), small spurs are seen near the maximum and minimum areas of the modulation waveform. These spurs are due to coherent interference from the commutation switching points in the DC motor, used to rotate the chopper. They are much less evident in the waveform measured using a stronger optical signal power. This coherent interference can be removed from a weak measured signal waveform by subtracting the coherent waveform measured when the beam is off from the waveform measured when the beam is on. After normalizing the amplitude according to RMS power, offset correction is achieved by adding or subtracting a constant value to all samples in order to make the integral over 1 full period equal to zero. This corrected waveform can then be used directly as the internal reference signal for lock-in detection, and will be optimal for signals buried in a white-noise background. More advanced spectral weighting and coherent interference removal techniques may be used for measurements where the noise distribution is not white, or where the coherent interference amplitude is not stable. Coherent interference arising in the modulator or reference signal circuitry is often the limiting factor for lock-in detectors.

Optimal spectral weighting is computationally more involved. After obtaining the signal waveform and removing coherent interference by subtraction, a noise correlation matrix must be computed from a measurement of many cycles of the noise. This noise correlation matrix is a square matrix with a dimensionality equal to the number of reference signal samples. The matrix elements are the cross-correlation between the elements of the noise vectors, with values between

-1 and 1, with a value of 0 for no correlation. The noise correlation for white noise is thus the identity matrix, as the diagonal values (1's) will reflect perfect correlation of each value with itself, while the off-diagonal values (0's) will reflect absence of correlation between all other values. After the noise correlation matrix has been computed, it must be inverted. The inverse noise correlation matrix is then multiplied by the measured modulation waveform vector, and the resulting vector will be an optimally-weighted version of the original with respect to the noise power spectral distribution (Doering & Harvey, 1987).

The Vernier principle of measurement resolution improvement is a consequence of number theory, and may be used to draw an intuitive analogy for optimal sampling rules. The Vernier principle was first employed by the Portuguese navigator Pedro Nunes (Petrus Nonius, 1502-1578), and later developed by the French mathematician Pierre Vernier (c.1584-c.1638). If a unit quantity (unit length or angle) is subdivided into two different integral numbers of parts, and these numbers are relative primes that share no common factors, then a measurement can be taken with a subunit resolution equal to their product. The absolute precision of the division process and reading process need not be equal to the effective subunit resolution obtained, due to the scale magnification effect, as well as additional information provided by the symmetrical mismatch pattern from neighboring pairs on the scales.

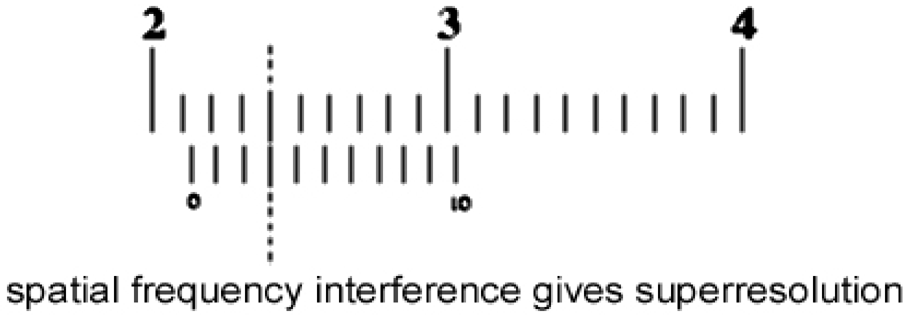

Fig. 35 - Vernier measurement increases measurement resolution by using the interference between sliding scales. The scales divide a unit quantity into two relative prime numbers (9 & 10), which share no common factors. The spatial frequencies of the divisions interfere uniquely. One pair on the scales always matches best. The symmetrical pattern formed by slight leading and lagging phases of neighboring pairs provide additional information, which reduces the effects of imprecision in the scales themselves. Thus, Vernier measurement allows for measurement resolution better than the scale resolution, or even the visual resolution of the observer, because of the relative prime relationship between the scales, which is known *a priori*. The Vernier measurement principle is closely related to aliasing and leakage, which affect the accuracy of discrete Fourier transforms. Aliasing errors would result if the scales were not divided into relative primes (5 & 10). In this case, there is an ambiguity in the matching point, and several points match equally well. Leakage errors would result if the divided unit quantity on one scale is not equal to the other.

The Vernier principle can be also be used to understand the effects of aliasing and leakage errors in discrete Fourier transforms. The discrete Fourier transform provides spectrum information on multiples of the frequency resolution. Aliasing errors are typically explained as arising when the Nyquist criterion is violated, and the sampling rate is lower than twice the highest signal component frequency present. However, undersampling systems are capable of recovering high-frequency components of periodic repetitive waveforms, provided that the sampling is performed in an interleaved fashion which effectively samples the wave at a higher frequency over the course of several signal periods. Aliasing, then, is the general consequence of discrete sampling at a number of points which do not visit unique portions of the repeating signal waveform over several periods. Leakage errors are typically regarded as resulting from limited (finite) sampling time. However, a system which samples a periodic waveform over a finite time can also have zero leakage (of the signal), provided that an exact integer number of signal waveforms are sampled and the data are sampled long enough to give a mean-zero value for the noise. In the context of optimal signal recovery, these concepts have been explored to give optimal rules for choice of sampling frequencies and measurement times, relative to discrete signal and interference sources (Zhang & Soukas, 1991).

The Vernier measurement principle analogy, though inexact, is a straightforward and intuitive example of how the spectral properties of a large structure can be used to increase effective measurement resolution.

Measurements based on the interference between two periodic structures, such as transmission gratings or machine threads, utilize an important principle. The measurement accuracy is greater than the accuracy of the periodic structures upon which the measurement is based (Sayce, 1972). This is because of averaging effects which minimize the importance of local defects in the periodic structures. In contrast, measurements using ordinary scales are always less precise than the scale itself. This principle has been re-discovered several times in history.

One important use of this principle is to improve the quality of periodic structure fabrication. The first accurate modern machine tools were built by Henry Maudslay (1771-1831). Maudslay realized that accurate threads were necessary for precision machining and manufacturing. Hand-cut nuts and bolts had existed as fasteners for several hundred years, with meshing that was necessarily coarse and crude, due to imprecision. Without large gaps, the non-uniform threads would not mate. The key innovation of Maudslay was to use averaging to improve the quality and uniformity of the threading itself in a self-improving process. In the Maudslay approach, a crude thread was cut into a metal cylinder stock by hand, using files and chisels. A special elastic nut was then formed by clamping on a tubular shell that was lined with a blend of elastic material such as pitch and sawdust. The nut extended over several periods of the thread, and its position was set by the elastic averaging properties of the female threading. A steel blade affixed to the end of this nut was then used to cut an extension of the thread onto the bar stock, with greater accuracy and uniformity that the initial thread. This process can be repeated several times, with several different nuts of random length, until the result is a thread that is nearly perfect in uniformity over a large length. This basic process of improvement of uniformity by elastic averaging was re-discovered by Sir Thomas Merton (Merton, 1950). Machine tools based on the Merton nut were used by the British National Physics Laboratory to create improved spectroscopic diffraction gratings and transmission gratings for optical position measurement. These transmission gratings were, in turn, used for precision electronic control of machine tools, a key development in modern manufacturing history (Sayce, 1972). The underlying concept from the early history of machine tools returned in a full technological circle.

Grating-based measurement principles are one source of improved precision in synthetic chopping, because they increase the absolute angular measurement precision. The synthetic reference signal may be generated using a rotary encoder system, such as the US Digital EM1-1-1250, that has a segmented photodiode detector which matches the encoder disk. The clock pulses from this transducer are derived from a measurement of several periods of the encoder pattern, and the precision of the measurement is fundamentally improved. By contrast, the reference signal from an ordinary chopper is invariably based on a single edge transition, whether based on a photo-interrupt sensor or a pick-off of the modulated beam.

This will cause both deterministic phase errors (due to non-uniform chopper disk blades, which may be wider or narrower than one another) and also random phase errors. The deterministic errors are eliminated by using an indexed synthetic reference waveform, while the random errors are reduced by using a transducer which averages over several periods of the encoder disk pattern. Because a synthetic chopper may use a grating-based measurement principle for generation of the reference signal, it is fundamentally superior to any single-point reference transducer system.

ii. Implementation

An implementation of the synthetic chopper was designed and built for recovery of the weak Raman signals from OH and CH bonds, using a low excitation power and simple detector. The chopper is an enclosed design, with a housing made from machined aluminum plate. The housing provides threaded mounting holes for the motor, optical rotary encoder, and external mounting. Additionally, threaded mounts for SMA fiber optic couplings are provided in some locations. The chopper wheel is machined from brass sheet stock using a 4-axis CNC milling machine, with the 4th axis used to rotate the chopper and a single axis used to cut. This is more precise than cutting using X and Y axis milling movements. A low-frequency chopper wheel (3 cycles per revolution) and a high-frequency chopper wheel (127 cycles per revolution) were constructed so that low and high modulation frequencies could be achieved within the motor speed limits (4-100 Hz).

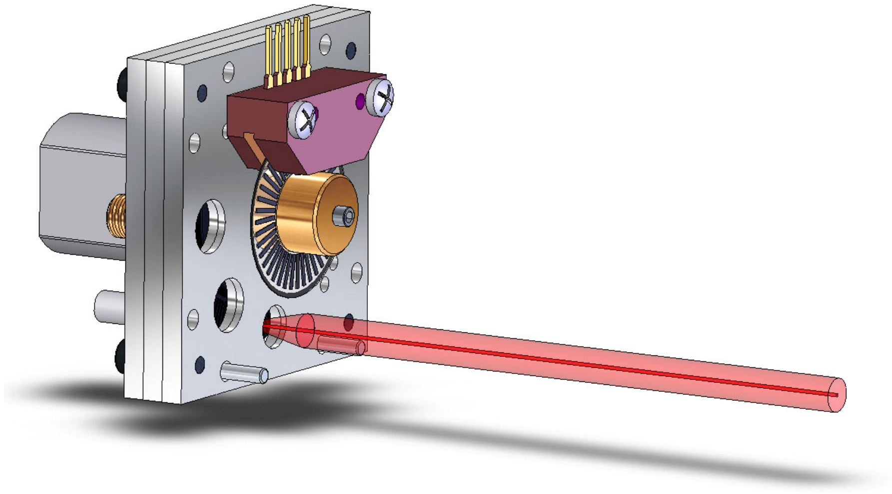

Fig. 36 - The synthetic chopper is constructed by mounting a modulator disk and optical rotary encoder disk on the same hub. The rotary encoder pulses are used to directly trigger sampling, resulting in a PLL system with a very high effective loop bandwidth that is able to remove the effect of deterministic errors in the disks.

The synthetic reference signal encoder system was built using a US Digital EM1-1-1250 indexed rotary optical encoder, run in 4X mode to give 5000 cycles per revolution. The TTL signal from the LS7184 IC was used to directly trigger analog-to-digital conversion. An adjustable resistor was included in order to ensure that the TTL pulse width is sufficient to trigger acquisition for different ADC chips. The usual requirement is for a pulse width of at least a few 10's of nanoseconds duration.

The data acquisition system is a National Instruments 4 channel 16-bit data acquisition board (National Instruments PCI-6122). This board is unique in that it accepts an external TTL signal to trigger the acquisitions. In other respects it is very standard. Newer generations of this product line are now available that offer more than double the sampling rate, and more channels as well.

Several virtual instruments were programmed in the LabVIEW programming language. These virtual instruments are used to measure the signal waveform and noise spectrum, and also correct the signal waveform according to coherent interference, incoherent interference and broadband noise.

A digital lock-in amplifier is also implemented in software. This digital lock in amplifier has shown the ability to realize a 2.17E-17 W noise floor in 2000-second measurements of the output of a photodiode transimpedance amplifier. This exceeds published results from well-designed systems of a similar architecture, using commercial digital lock-in amplifiers and selected high-shunt resistance photodiodes (Eppeldauer, 2000).

iii. Cost and performance of a complete system

The cost of the synthetic chopper and lock in amplifier is less than $2000. The majority of the cost is the ADC board (National Instruments PCI-6122), $1800. This gives a four channel instrument, at one fourth the cost of a single-channel commercial digital LIA. Additionally, this card is capable of sampling at up to 500ks/s on each channel, twice the rate of the SRS850. The complete system has a very low noise floor, due to the minimal level of coherent interference, and the quality of the phase-locking system achieved using the synthetic chopper.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Synthetic optical methods have been explored for clinically-relevant measurements of corneal biophysical parameters. A refined model of the hydration-refractive index relationship has been developed, and a new series of measurements have been taken with good correspondence to the model. A major conclusion is that prior Gladstone-Dale models of corneal refractive index under hydration perturbation fail to account for dispersive properties of the components. Wavelength-dependent component refractive indices must be used to accurately model measured data. Additionally, the refined model may be used to reinterpret existing measured data, and guide the direction of future research in this area. Uncertainty in the corneal refractive index due to hydration shifts is minimal over the normal physiological hydration range.

While a full system for corneal hydration measurement by Raman spectroscopy at a safe power level was not demonstrated, the nontraditional system architecture explored will provide a new path with some waypoints for future investigators.

This research was intended to improve the state of clinical eye care by focusing on well-defined shortcomings of current ophthalmic instrumentation with regard to evolving evidence-based medicine. A dedicated focus on translational research has led to new and clinically-applicable techniques to improve one area of clinical instrumentation, and has improved the state of the art in another area. In addition, a moderate improvement in the quantitative accuracy of an important biophysical model has also been achieved.

- Jesse Hart Shaver, September, 2007

BIBLIOGRAPHY AND REFERENCES

A

Arner RS, Rengstorff RH. Error analysis of corneal thickness measurements. *American Journal of Optometry and Archives of American Academy of Optometry* p.862-865 (October 1972)

Asensio I, Saleh R, Alonso L, Palanca-Sanfrancisco JM, Sanchis-Gimeno J. Corneal thickness values before and after oxybuprocaine .4% eye drops. Cornea, Volume 22, Number 6, p.527-532 (August 2003)

Atchison DA, Smith G. Chromatic dispersions of the ocular media of human eyes. J. Opt. Soc. Am. A, Vol. 22, No. 1, p.29-37 (January 2005)

Atta HR. New applications in ultrasound technology. British Journal of Ophthalmology, Volume 83, p.1246-1249 (1999)

B

Bashkatov AN, Genina EA, Kochubey V, Tuchin VV. Estimation of wavelength dependence of refractive index of collagen fibers of scleral tissue. Proc. SPIE Vol. 4162, p.265-268 (November 2000)

Bartoli A, Poggi P, Quercioli F, Tiribilli B. Fast one-dimensional profilometer with a compact disc pickup. Applied Optics, Volume 40, No. 7, p.1044-1048 (March 2001)

Bauer N, Hendrikse F, March WF. In vivo confocal Raman spectroscopy of the human cornea. *Cornea* 18(4):483 (1999)

Bauer NJC, Wicksted JP, Jongsma FHM, March WF, Hendrikse F, Motamedi M. Noninvasive assessment of the hydration gradient across the cornea using confocal Raman spectroscopy. IOVS, Vol. 39, No.5, p.831-835 (1998)

Bechmann M, Thiel MJ, Neubauer AS, Ullrich S, Ludwig K, Kenyon KR, Ulbig MW. Central corneal thickness measurement with a retinal optical coherence tomography device versus standard ultrasound pachymetry. Cornea, Volume 20, Number 1, p.50-54 (2001)

Benjamin KD, Armitage AF, South RB. Harmonic errors associated with the use of choppers in optical experiments. *Measurement*, 39 (8) 764-770 (Oct. 2006)

Benschop J, van Rosmalen G. Confocal compact scanning optical microscope based on compact disc technology. Applied Optics, Vol. 30, No. 10, p.1179-1184 (April 1991)

Blair DP, Sydenham PH. Phase sensitive detection as a means to recover signals buried in noise. *Journal of Physics E: Scientific Instruments*, Vol. 8, p.621-672 (1975)

Bodenheimer T. High and rising health care costs. Part 2: Technologic Innovation. *Annals of Internal Medicine*, 142:932-937 (2005)

Bohnke M, Chavanne P, Gianotti R, Salathe RP. High-precision, high-speed measurement of excimer laser keratectomies with a new optical pachymeter. Ger J Ophthalmol. 5(6):338-42 (1996)

Born M, Wolf E. Principles of Optics, 7th Ed. Cambridge University Press (1999) ISBN 0 521 642221

Brandt JD, Beiser, JA, Kass MA, Gordon MO, and the OHTS group. Central Corneal Thickness in the Ocular Hypertension Treatment Study. Ophthalmology, Vol. 108, No. 10, p.1779-1788 (2001)

C

Caspers PJ, Lucassen GW, Wolthuis R, Bruining HA, Puppels GJ. In vitro and in vivo Raman spectroscopy of human skin. Biospectroscopy, Vol.4, pS31-S39 (1998)

Chakrabarti HS, Craig JP, Brahma A, Malik TY, McGhee CNJ. Comparison of corneal thickness measurements using ultrasound and Orbscan slit-scanning topography in normal and post-LASIK eyes. *Journal of Cataract and Refractive Surgery,* Vol. 27, p.1823-1828 (2001)

Chandley PJ. A simple technique for reducing the phase jitter of a light chopper. *Infrared Physics*, Vol. 22, p.311-312 (1982)

Chase B. A new generation of Raman instrumentation. Applied Spectroscopy, Volume 48, Number 7, p.14A-19A (1994)

D

Daimon M, Masamura A. Measurement of the refractive index of distilled water from the near-infrared region to the ultraviolet region. Applied Optics, Vol. 46, No. 18, p.3811-3820 (June 2007)

Dayanir V, Sakarya R, Özcura F, Kir E, Aktunç T, Özkan BH, Okyay P. Effect of corneal drying on central corneal thickness. Journal of Glaucoma, Volume 13, Number 1, p.6-8 (February 2004)

Dicke RH. The measurement of thermal radiation at microwave frequencies. *Review of Scientific Instruments*, Vol. 17, No. 7, p.268-275 (July 1946)

Dirckx JJJ, Kuypers LC, Decraemer WF. Refractive index of tissue measured with confocal microscopy. Journal of Biomedical Optics, Vol. 10(4), p.044014-1-8 (July-August, 2005)

Dixon PK, Wu L. Broadband digital lock-in amplifier techniques. Review of Scientific Instruments, Volume 60, Number 10, p.3329-3336 (October 1989)

Doering CR, Harvey PM. Optimal signal-to-noise in digital phase lock amplifiers. *Applied Optics*, Vol. 26, No. 4, p.633-642 (1987)

Donati S. Photodetectors: Devices, Circuits, and Applications. Prentice Hall PTR (2000) ISBN 0-13-020337-8

Dougherty PJ, Wellish KL, Maloney RK. Excimer laser ablation rate and corneal hydration. *American Journal of Ophthalmology*, Vol. 118, No. 2, 169-177 (1994)

Doughty MJ, Zaman ML. Human corneal thickness and its impact on intraocular pressure measures: a review and meta-analysis approach. Major Review. *Survey of Ophthalmology*, 44(5):367-408 (2000)

Dueker DK, Singh K, Lin SC, Fechtner RD, Minckler DS, Samples JR, Schuman JS. Corneal thickness measurement in the management of primary open-angle glaucoma: A report by the American Academy of Ophthalmology. Ophthalmic Technology Assessment. *Ophthalmology*, Vol. 114, p.1779-1787 (September 2007)

E

Ehlers N, Hjortdal J. Corneal thickness: measurement and implications. Review. *Experimental Eye Research*, 78;543-548 (2004)

Ehlers N, Bramsen T, Sperling S. Applanation tonometry and central corneal thickness. Acta Ophthalmologica, Vol. 53, p.34-43 (1975)

Ehrmann K, Ho A, Schindleim K. A 3D optical profilometer using a compact disc reading head. Measurement Science and Technology, Volume 9, p.1259-1265 (1998)

Eppeldauer GP. Noise-optimized silicon radiometers. Journal of Research of the National Institute of Standards and Technology, Volume 105, Number 2, p.209-219 (2000)

Erckens RJ, Jongsma FHM, Wicksted JP, Hendrikse F, March WF, Motamedi M. Drug-induced corneal hydration changes in vivo by non-invasive confocal Raman spectroscopy. *Journal of Raman Spectroscopy*. 32:733-737 (2001)

F

Fan K-C, Chu C-L, Mou J-I. Development of a low-cost autofocusing probe for profile measurement. Measurement Science and Technology, Volume 12, p.2137-2146 (2001)

Fatt I, Harris MG. Refractive index of the cornea as a function of its thickness. *American Journal of Optometry and Archives of American Academy of Optometry* p.383-386 (May 1973)

Faust RC. Refractive Index Determination by the Central Illumination (Becke Line) Method. *Proc. Phys. Soc.* LXVIII, 12-B, 4D, p.1081-1094 (1955)

Fisher BT, Masiello KA, Goldstein MH, Hahn DW. Assessment of transient changes in corneal hydration using confocal Raman spectroscopy. *Cornea* 22(4):363-370 (2003)

Fisher WG, Piston DW, Wachter EA. Phase sensitive demodulation in multiphoton microscopy. *Microscopy and Microanalysis*, 8(3):191-202 (2002)

Fishman GR, Pons ME, Seedor JA, Liebmann JM, Ritch R. Assessment of central corneal thickness using optical coherence tomography. J Cataract Refract Surg, Volume 31, p.707-711(2005)

Flanigan DF. Discrepancies between two formulations of signal-to-noise ratio for background-limited detection. Applied Optics, Volume 34, Number 15, p.2721-2723 (1995)

Fogt N, King-Smith PE, Tuell G. Interferometric measurement of tear film thickness by use of spectral oscillations. J. Opt. Soc. Am. A, Vol. 15, No. 1, p.268-275 (January 1998)

Freegard TJ. The physical basis of transparency of the normal cornea. *Eye*, Volume 11, p465-471 (1997)

Freund DE, McCally RL, Farrell RA. Effects of fibril orientations on light scattering in the cornea. J. Opt. Soc. Am. A, Vol. 3, No. 11, p.1970-1982 (November 1986)

Friedman MH. A quantitative description of equilibrium and homeostatic thickness regulation in the in vivo cornea. *Biophysical Journal*, Volume 12, p.648-665 (1972)

G

Gao L, Fan H, Cheng ACK, Wang Z, Lam DSC. The effect of eye drops on corneal thickness in adult myopia. Cornea, Volume 25, Number 4, p.404-407 (May 2006)

Gaujoux T, Borderie VM, Yousfi H, Bourcier T, Touzeau O, Laroche L. Comparison of optical low-coherence reflectometry and ultrasound pachymetry in measuring corneal graft thickness. Cornea, Volume 26, Number 2, p.185-189 (February 2007)

Geist J. Waveform-independent lock-in detection. *Review of Scientific Instruments*, Vol. 43, No. 11, p.1704-1705 (November 1972)

Gillis A, Zeyen T. Comparison of optical coherence reflectometry and ultrasound corneal pachymetry. Bull Soc Belge Ophtalmol, Vol. 292, p.71-75 (2004)

Goldmann H, Schmidt T. Über applanationstonometrie. Ophthalmologica, Volume 134, p.221–242 (1957)

Gordon MO, Beiser JA, Brandt JD, Heuer DK, Higginbotham EJ, Johnson CA, Keltner JL, Miller JP, Parrish, RK II, Wilson MR, Kass MA, and the OHTS Group. The Ocular Hypertension Treatment Study: Baseline Factors That Predict the Onset of Primary Open-Angle Glaucoma. Archives of Ophthalmology, Vol. 120, p.714-720 (2002)

Graeme J. Photodiode Amplifiers: Op Amp Solutions. McGraw-Hill (1996) ISBN0-07-024247-X

H

Hanlon EB, Manoharan R, Koo TW, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, Feld MS. Prospects for in vivo Raman spectroscopy. Review. *Phys. Med. Biol.*, 45, R1-R59 (2000)

Harper CL, Boulton ME, Bennett D, Marcyniuk B, Jarvis-Evans JH, Tullo AB, Ridgeway AE. Diurnal variations in human corneal thickness. *British Journal of Ophthalmology*, Volume 80, No. 12, p.1068-1072 (December 1996)

Harvey AH, Gallagher JS, Levelt Sengers JMH. Revised formulation for the refractive index of water and steam as a function of wavelength, temperature, and density. Journal of Physical and Chemical Reference Data, Vol. 27, No.4, p.761-774 (1998)

Hedbys BO, Mishima S. Flow of water in the corneal stroma. *Experimental Eye Research*, 1:262-275 (1962)

Hedbys BO, Mishima S. The thickness-hydration relationship of the cornea. *Experimental Eye Research*, 5(3):221-228 (1966)

Hobbs PCD. Building Electro-Optical Systems. John Wiley and Sons, Inc. (2000) ISBN 0-471-24681-6

Holzman JF, Elezzabi AY. A dispersion-free high-speed beam chopper for ultrafast-pulsed-laser applications. *Measurement Science and Technology*. **14** p.N41-N44 (2003)

Horowitz P, Hill W. The Art of Electronics, 2nd Edition. Cambridge University Press (1989) ISBN 13-978-0-521-37095-0

I

Ihanamäki T, Pelliniemi LJ, Vuorio E. Collagens and collagen-related matrix components in the human and mouse eye. *Progress in Retinal and Eye Research* **23** p.403-434 (2004)

Ivarsen A, Stultiens BAT, Møller-Pedersen T. Validation of confocal microscopy through focusing for corneal sublayer pachymetry. Cornea, Volume 21, Number 7, p.700-704 (2002)

J

Johnson J. The Schottky effect in low frequency circuits. Physical Review, Volume 26, Issue 1, p.71-85 (1925)

Johnson J. Thermal agitation of electricity in conductors. Physical Review, Volume 32, Issue 1, p.97-107 (1928)

Johnson M. Photodetection and Measurement. McGraw-Hill (2003) ISBN 0-07-140944-0

K

Kaye GI. Stereologic measurement of cell volume fraction of rabbit corneal stroma. Arch. Ophthalmol., Vol. 82, p.792-794 (1969)

Kim WS, Jo JM. Corneal hydration affects ablation during laser in situ keratomileusis surgery. *Cornea*, 20(4):394-7 (2001)

Kim YL, Walsh JT, Goldstick TK, Glucksberg MR. Variation of corneal refractive index with hydration. *Physics in Medicine and Biology*, 49:859-868 (2004)

Kniestedt C, Lin S, Choe J, Nee M, Bostrom A, Sturmer J, Stamper RL. Correlation between intraocular pressure, central corneal thickness, stage of glaucoma, and demographic patient data: prospective analysis of biophysical parameters in tertiary glaucoma practice populations. *Journal of Glaucoma*, 15(2):91-97 (2006)

Kniestedt C, Nee M, Stamper RL. Accuracy of dynamic contour tonometry compared with applanation tonometry in human cadaver eyes of different hydration states. *Graefe's Arch Clin Exp Ophthalmol*. 243:359-366 (2005)

Kuypers LC, Decraemer WF, Dirckx JJJ, Timmermans J-P. A procedure to determine the correct thickness of an object with confocal microscopy in case of refractive index mismatch. Journal of Microscopy, Vol. 218, 68-78 (2005)

L

Lau KS, Tan CH, Ng BK, Li KF, Tozer RC, David JPR, Rees GJ. Excess noise measurement in avalanche photodiodes using a transimpedance amplifier front end. Measurement Science and Technology, Volume 17, p.1941-1946 (2006)

Leonard DW, Meek KM. Refractive indices of the collagen fibrils and extrafibrillar material of the corneal stroma. *Biophysical Journal*. Volume 72, p.1382-1387 (March 1997)

Li HF, Petroll WM, Møller-Pedersen T, Maurer JK, Cavanagh HD, Jester JV. Epithelial and corneal thickness measurements by in vivo confocal microscopy through focusing (CMTF). Current Eye Research, 16:214-221 (1997)

Li L, Uttamchandani D. Design and evaluation of a MEMS optical chopper for fibre optic applications. *IEE Proceedings - Science, Measurement and Technology*, Volume **151**, Issue 2, p. 77-84 (March 2004)

M

Manchester TP. Hydration of the cornea. *Transactions of the American Ophthalmological Society*. 68:427-461 (1970)

Manoharan R, Baraga JJ, Feld MS, Rava RP. Quantitative histochemical analysis of human artery using Raman spectroscopy. *J. Photochem. Photobiol. B: Biol*. Volume 16, p.211-233 (1992)

Marsich MM, Bullimore MA. The repeatability of corneal thickness measures. *Cornea*, Volume 19, Number 6, p.792-795 (2000)

Martínez-Antón JC, Bernabeu E. Spectrogoniometry and the WANTED method for thickness and refractive index determination. *Thin Solid Films*, 313-314, p.85-89 (1998)

Maurice DM. The structure and transparency of the cornea. *Journal of Physiology*. Volume 136, p.263-286 (1957)

Maurice DM, Giardini AA. A simple optical apparatus for measuring the corneal thickness and the average thickness of the human cornea. *British Journal of Ophthalmology*, Volume 35, p.169-177 (1951)

McCreery RL. Photometric standards for Raman spectroscopy. Handbook of Vibrational Spectroscopy, John Wiley & Sons. (2002)

McLaren JW, Nau CB, Erie JC, Bourne WM. Corneal thickness measurement by confocal microscopy, ultrasound, and scanning slit methods. *American Journal of Ophthalmology*. Vol. 137, No. 6, p.1011-1020 (2004)

McQuistan RB. On radiation modulation. *Journal of the Optical Society of America*. Vol. 49, No. 1. p.70-74 (January 1959)

McQuistan RB. On an approximation to sinusoidal modulation. *Journal of the Optical Society of America*, Vol. 48, No. 1. p.63-66 (January 1958)

Meade ML. Advances in lock-in amplifiers. *Journal of Physics E: Scientific Instruments*, Vol. 15, p.395-403 (1982)

Medeiros FA, Sample PA, Weinreb RN. Corneal Thickness Measurements and Visual Function Abnormalities in Ocular Hypertensive Patients. American Journal of Ophthalmology, Vol. 135, No.2, p.131-137 (2003)

Meek KM, Leonard DW, Connon CJ, Dennis S, Khan S. Transparency, swelling, and scarring in the corneal stroma. *Eye*, Volume 17, p.927-936 (2003)

Meek KM, Dennis S, Khan S. Changes in the refractive index of the stroma and its extrafibrillar matrix when the cornea swells. *Biophysical Journal*, Volume 85, p.2205-2212 (October 2003)

Merton T. On the reproduction and ruling of diffraction gratings. *Proceedings of the Royal Society of London, Series A. Mathematical and Physical Sciences*, Vol. 201, No. 1065, p.187-191 (March 1950)

Miglior S, Albe E, Guareschi M, Mandelli G, Gomarasca S, Orzalesi N. Intraobserver and interobserver reproducibility in the evaluation of ultrasonic pachymetry measurements of central corneal thickness. British Journal of Ophthalmology, Volume 88, p.174-177 (2004)

Modis L, Langenbucher A, Seitz B. Scanning-slit and specular microscopic pachymetry in comparison with ultrasonic determination of corneal thickness. *Cornea*, Vol. 20, No. 7, p.711-714 (2001)

Møller-Pedersen T. On the structural origin of refractive instability and corneal haze after excimer laser keratectomy for myopia. Acta Ophthalmologica Scandinavica, Volume 81, Supplement 237, p.6-20 (February 2003)

Myrowitz EH, Ren S, Chuck RS. Comparison of central corneal thickness measured by four different pachymeters. Eye and Contact Lens, Volume 33, Number 3, p.156-160 (2007)

N

Nam SM, Lee HK, Kim EK, Seo KY. Comparison of corneal thickness after the instillation of topical anesthetics: proparacaine versus oxybuprocaine. Cornea, Volume 25, Number 1, p.51-54 (January, 2006)

Naszodi LJ. Consideration of the real modulation in correlation TOF technique. *Nuclear Instruments and Methods*. Vol. 161, p.163-165 (1979)

Neelkantan K, Dattagupta S, Rajappan KP. Signal to noise enhancement of lock-in amplifiers. *Review of Scientific Instruments*, Vol. 51, No. 2, p.251-252 (February 1980)

Nemanich RJ, Solin SA, Doehler J. Inexpensive high-speed dentist drill light chopper and its use in rejecting luminescence background from Raman spectra. *Review of Scientific Instruments*, Vol. 47, No. 6, p.741-744 (June 1976)

Nemeth G, Tsorbatzoglou A, Kertesz K, Vajas A, Berta A, Modis L. Comparison of central corneal thickness measurements with a new optical device and a standard ultrasonic pachymeter. J. Cataract Refract Surg, Volume 32, p.460-463 (2006)

Nolte DD, Regnier FE. Spinning-disk interferometry: The BioCD. Optics and Photonics News, p.48-53 (October 2004)

Nyquist H. Thermal agitation of electric charge in conductors. Physical Review, Volume 32, Issue 1, p.110-113 (1928)

O

O'Donnell C, Maldonado-Codina C. Agreement and repeatability of central thickness measurement in normal corneas using ultrasound pachymetry and the OCULUS Pentacam. *Cornea,* Vol. 24, No. 8, p.920-924 (2005)

Oshika T, Yoshitomi F, Oki K. The pachymeter guide: A new device to facilitate accurate corneal thickness measurement. Jpn J Ophthalmol, Vol. 41, p.426-427 (1997)

Ott HW. Noise Reduction Techniques in Electronic Systems, 2nd Edition. John Wiley and Sons (1988) ISBN 0-471-85068

P

Pakravan M, Parsa A, Sanagou M, Parsa CF. Central corneal thickness and correlation to optic disc size: a potential link for susceptibility to glaucoma. British Journal of Ophthalmology, Volume 91, Number 1, p.26-28 (January 2007)

Pearce, JB. On optimum chopping. *Nuclear Instruments and Methods*, Vol. 103, p.435-436 (1972)

Pelizzari CA, Postol TA. Effect of imperfect beam modulation in neutron time of flight spectroscopy using mechanical correlation choppers. *Nuclear Instruments and Methods*. Vol. 143, p.139-144 (1977)

Q

Quan X, Fry ES. Empirical equation for the index of refraction of seawater. Applied Optics, Vol. 34, No. 18, p.3477 to 3480 (June 1995)

R

Rainer G, Petternel V, Findl O, Schmetterer L, Skorpik C, Luksch A, Drexler W. Comparison of ultrasound pachymetry and partial coherence interferometry in the measurement of central corneal thickness. Journal of Cataract and Refractive Surgery, Volume 28, p.2142-2145 (2002)

Rainer G, Findl O, Petternel V, Kiss B, Drexler W, Skorpik C, Georgopoulos M, Schmetterer L. Central corneal thickness measurements with partial coherence interferometry, ultrasound, and the Orbscan system. *Ophthalmology*, Vol. 111, No.5, p.875-879 (2004)

Rako P. Measuring nanoamperes. Electronic Design News, p.42-52 (April 26, 2007)

Rovati L, Docchio F. Autofluorescence methods in ophthalmology. Journal of Biomedical Optics, Volume 9, Number 1, p.9-21 (2004)

Rubiola E, Giordano V. On the 1/f frequency noise in ultra-stable quartz oscillators. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 54, No. 1, p.15-22 (January, 2007)

Ryu YK, Oh C, Lim J-S. Development of a noncontact optical sensor for measuring the shape of a surface and thickness of transparent objects. Optical Engineering 40(4) 500-502 (2001)

S

Sabetti L, Renzetti A, D'Alessandri L, Balestrazzi E. Eventual error caused by dehydration with pachometry. Ophthalmologica, Volume 215, p.97-101 (2001)

Sayce LA. Gratings in metrology. *Journal of Physics E: Scientific Instruments*, Vol. 5, p.193-198 (1972)

Schmid GF. Axial and peripheral eye length measured with optical low coherence reflectometry. Journal of Biomedical Optics, Volume 8, Number 4, p.655-662 (October 2003)

Seizer PM, Yen WM. High-speed light chopper wheel for modulation spectroscopy. *Review of Scientific Instruments*, Vol. 47, No. 6, p.749-750 (June 1976)

Shaver JH, Shen JH, Walker AL. A low-cost optical pachymeter for the diagnosis of glaucoma and ocular hypertension. Proc. of the 26th IEEE EMBS, 2330-2333 (2004)

Shildkrot Y, Liebmann JM, Fabijanczyk B, Tello CA, Ritch R. Central corneal thickness measurement in clinical practice. Journal of Glaucoma, Volume 15, Number 5, p. 331-336 (October 2005)

Sivak JG, Mandelman T. Chromatic dispersion of the ocular media. Vision Research, Volume 22, Issue 8, p.997-1003 (1982)

Smith E, Dent G. Modern Raman Spectroscopy: A Practical Approach. John Wiley and Sons (2005) ISBN 0-471-49668-5

Smith JW. The transparency of the corneal stroma. Vision Research, Volume 9, Issue 3, p.393-396 (1969)

Smith WJ. Modern Optical Engineering, 3rd Ed. McGraw-Hill Co. (2000) ISBN 0 07 136360 2

Sydenham PH, Thorn T. Handbook of Measuring System Design. John Wiley and Co. (2005) ISBN 0-470-02143-8

T

Toshiyoshi H, Fujita H, Ueda T. A self-excited chopper made by quartz micromachining and its application to an optical sensor. *Proc. IEEE Workshop on Micro Electro Mechanical Systems*, Oiso, Japan, Jan. 25-28, 1994 p.325-330

U

V

van Der Ziel, A. Noise: Sources, Characterization, Measurement. Prentice-Hall, Inc. (1970) ISBN-10 : 0136231659

van Der Ziel, A. Unified presentation of 1/f noise in electronic devices: fundamental 1/f noise sources. Proceedings of the IEEE, Vol. 76, No.3, p.233-258 (March 1988)

Vincent D. Amplitude modulation with a mechanical chopper. *Applied Optics*, Vol. 25, No. 7 p.1035-1036 (April 1986)

W

Wallin W. Wave synthesizing light chopper. *Journal of the Optical Society of America*. Vol. 45, No. 4. p.287-292 (April 1955)

Walrafen GE, Chu Y-C. Nature of collagen-water hydration forces. A problem in water structure. Chemical Physics, Volume 258, p.427-446 (2000)

Walti R, Bohnke M, Gianotti R, Bonvin P, Ballif J, Salathe RP. Rapid and precise in vivo measurement of human corneal thickness with optical low-coherence reflectometry in normal human eyes. Journal of Biomedical Optics, Volume 3, Number 3, p.253-258 (July 1998)

Wang XJ, Milner TE, Chang MC, Nelson JS. Group refractive index measurement of dry and hydrated type I collagen films using optical low-coherence reflectometry. Journal of Biomedical Optics, Volume 1, Number 2, p.212-216 (April 1996)

Wang J, Fonn D, Simpson T, Jones L. Precornea and pre- and postlens tear film thickness measured indirectly with optical coherence tomography. Investigative Ophthalmology and Visual Science, Volume 44, Number 6, p.2524-2528 (2003)

Wenthen FT, Snowman LR. High speed light choppers. *Applied Optics*, Vol. **12**, No. 4, p.822-824 (April 1973)

Wickham L, Edmunds B, Murdoch IE. Central corneal thickness: Will one measurement suffice? Ophthalmology. Vol.112, No.2, p.225-228 (2005)

Wilmshurst TH. Signal Recovery from Noise in Electronic Instrumentation. Adam Hilger Ltd. (1985) ISBN 0-85274-783-7

Windecker R, Fleischer M, Tiziani HJ. Low-coherence fiber-optic sensor with a large numerical aperture for topographic measurements. Applied Optics, Volume 37, Number 19, p.4080-4083 (July 1998)

Wolffenbuttel RF, de Graaf G. Noise performance and chopper frequency in integrated micromachined chopper-detectors in silicon. *IEEE Transactions on Instrumentation and Measurement*, Vol. 44, No. 2, p. 451-453 (April 1995)

X

Y

Yoshida S, Minami K, Okada K, Yamamoto H, Ueyama T, Sakai K, Kurata Y. Optical pickup employing a hologram-laser-photodiode unit. Japanese Journal of Applied Physics, Volume 39, Part 1, Number 2B, p.877-882 (February 2000)

Ytteborg J, Dohlman CH. Corneal edema and intraocular pressure. II. Clinical results. *Archives of Ophthalmology,* 74(4):477-484 (1965)

Z

Zhang SY, Soukas AV. Optimization of phase sensitive detectors using fast Fourier transform. Measurement Science and Technology, Volume 2, p.13-18 (1991)