Jin U. Kang *Editor*

Fiber Optic Sensing and Imaging

Springer

Fiber Optic Sensing and Imaging

Jin U. Kang
Editor

Fiber Optic Sensing and Imaging

Springer

Editor
Jin U. Kang
Department of Electrical and Computer Engineering
Johns Hopkins University
Baltimore
USA

ISBN 978-1-4614-7481-4 ISBN 978-1-4614-7482-1 (eBook)
DOI 10.1007/978-1-4614-7482-1
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013941502

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Optical fiber is one of the twentieth century's technology marvels. It made the modern day high-speed communications networks possible and this significant contribution was recognized by Charles K. Kao being awarded the Nobel Prize in Physics in 2009. However, unknown to most of the general public, optical fiber technology has made significant contributions to sensors and imaging technologies as well. It was recognized during the early part of the optical fiber development that optical fibers having small diameter, extremely long length, and flexibility have desired properties for developing endoscopic imaging devices and sensors. The long fiber length allowed the fiber optic sensors to have long interaction lengths which made them highly sensitive. The small diameter allowed the fiber optic sensors and imagers to be made compact and portable. The flexibility allowed the sensor and imagers to be placed in most tight spaces. Now, fiber optic imaging and sensing devices are being used for a wide range of applications such as in medical, environmental, manufacturing, and defense.

This book is designed to highlight the basic principles of fiber optic imaging and sensing devices. The book is by no means complete and comprehensive. But it is designed to provide the readers with a solid foundation in fiber optic imaging and sensing devices. It begins with an introductory chapter that starts from Maxwell's equations, mostly to show the readers where the governing equations in fiber devices come from and some idea as to how they are derived. The chapter ends with the derivation of the basic optical fiber characteristic equations and solutions (i.e., fiber modes). Chapter 2 reviews most common fiber optic interferometric devices which are bases for many fiber optic imaging and sensing systems. Chapter 3 discusses the basics of fiber optic imagers with emphasis on fiber optic confocal microscope. The fiber optic interferometric sensors are discussed in detail in Chaps. 4 and 5. Chapter 4 deals with fiber Bragg grating based sensor and various applications. Chapter 5 goes over in detail fiber Sagnac loop-based sensors. Chapter 6 covers optical coherence tomography in detail. Unlike other chapters, it goes into details the signal processing and systems level approach of the real-time OCT implementation that I hope the readers may find useful for building their own OCT systems. All the chapters start with theoretical derivation of devices characteristics. This is to help student reader with understanding the

underlying principle of fiber optic devices. Also useful forms of device characteristic equations are provided so that this book can be used as a reference for scientists and engineers in the optics and related fields.

This book could not have been written without the help of contributing authors, Do-Hyun Kim from FDA, Utkarsh Sharma and Xing Wei from Carl Zeiss Meditec, Li Qian from University of Toronto, Kang Zhang from GE Global Research, and Young-Geun Han from Hanyang University; my sincere thanks to them. Also help from one of my current students, Jaepyeong Cha, is greatly appreciated. I must say Brett Kurzman, Editor of Engineering/Applied Sciences, Springer must be one of the most patient people I know. Many thanks to his patience and accommodation.

Baltimore, March 2013 Jin U. Kang

Contents

Chapter 1
Optical Fibers

Jin U. Kang

1.1 Wave Equation

Light is an electromagnetic disturbance that propagates through space as a wave. Such electromagnetic (EM) wave is produced by a moving charge. An oscillating charge at a frequency f, produces EM wave with frequency f, schematically shown in Fig. 1.1:

1.1.1 Gauss's Law

Therefore, to understand the basic properties of light, you have to understand the basic laws that govern electric and magnetic fields. Gauss's Law for electric fields states that the total electric flux through any closed surface equals the net charge inside that surface divided by the electric permittivity of space, ε_0. This can be written in an integral form as:

$$\oint E \cdot dA = \frac{Q}{\varepsilon_0} \tag{1.1}$$

or differential form:

$$\nabla \cdot E = \frac{\rho}{\varepsilon_0} \tag{1.2}$$

Similarly, Gauss's Law for magnetic field states that the net magnetic flux through a closed surface is zero. This can be also written in an integral form:

J. U. Kang (✉)
Department of Electrical and Computer Engineering, Johns Hopkins University,
3400 N. Charles Street, Baltimore 21218, USA
e-mail: jkang@jhu.edu

J. U. Kang (ed.), *Fiber Optic Sensing and Imaging*,
DOI: 10.1007/978-1-4614-7482-1_1,

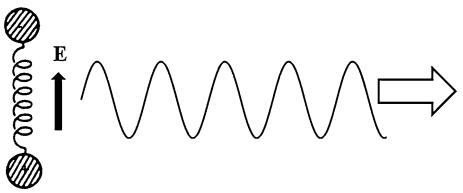

Fig. 1.1 Schematic of EM wave generation by a moving charge

$$\oint H \cdot dA = 0 \tag{1.3}$$

or differential form:

$$\nabla \cdot H = 0 \tag{1.4}$$

1.1.2 Faraday's Law

Faraday's Law of induction states that the line integral of the electric field around any closed path equals that rate of change of magnetic flux through any surface area bounded by that path. This can be written in integral form:

$$\oint E \cdot ds = \frac{d\Phi_m}{dt} \tag{1.5}$$

or differential form:

$$\nabla \times E = -\frac{\partial B}{\partial t} \tag{1.6}$$

Here, B is the magnetic displacement vector defined as: $B = \mu_0 H$. μ_0 is called magnetic permeability of space.

1.1.3 Ampere-Maxwell Law

Ampere-Maxwell Law states that the line integral of the magnetic field around any closed path is determined by the sum of the net conduction current through that path and the rate of change of electric flux through any surface bounded by that path. This can be written in integral form:

$$\oint B \cdot ds = \mu_0 I + \varepsilon_0 \mu_0 \frac{d\Phi_e}{dt} \tag{1.7}$$

or differential form:

$$\nabla \times H = J + \frac{\partial D}{\partial t} \tag{1.8}$$

D is the electric displacement vector defined as $D = \varepsilon_0 E + P$.

P is the electric polarization defined as $P = \varepsilon_0 \chi E$ where χ is called electric susceptibility and is an intrinsic property of the material system.

1.1.4 Maxwell's Equations

Combining the basic laws described above, one can assemble well-known Maxwell's Equations.

$$\nabla \bullet \boldsymbol{B} = 0 \tag{1.9}$$

$$\nabla \bullet \boldsymbol{D} = \rho \tag{1.10}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial \boldsymbol{t}} \tag{1.11}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial \boldsymbol{t}} \tag{1.12}$$

and

$$D = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \boldsymbol{B} = \mu_0 \boldsymbol{H} + \boldsymbol{M} \tag{1.13}$$

If we simply assume free space, i.e., J = M = P = 0, we can combine Eqs. 1.3 and 1.4 to obtain the following equation

$$\nabla \times \boldsymbol{E} = -\frac{\partial(\mu_0 \boldsymbol{H} + \boldsymbol{M})}{\partial \boldsymbol{t}} \tag{1.14}$$

And use it to derive the characteristic equation for electric field wave as:

$$\nabla \times \nabla \times \boldsymbol{E} = -\mu_0 \frac{\partial}{\partial \boldsymbol{t}} \nabla \times H \tag{1.15}$$

$$\nabla \times \nabla \times \boldsymbol{E} = -\mu_0 \frac{\partial^2 D}{\partial t^2} \tag{1.16}$$

$$\nabla \times \nabla \times \boldsymbol{E} = -\mu_0 \varepsilon_0 \frac{\partial^2 E}{\partial t^2} \tag{1.17}$$

Since

$$\nabla \times \nabla \times \boldsymbol{E} = \nabla(\nabla \bullet \boldsymbol{E}) - \nabla^2 E \tag{1.18}$$

Making paraxial beam approximation, i.e.,

$$\nabla(\nabla \bullet \boldsymbol{E}) = 0 \tag{1.19}$$

This simply implies that the curvature of the electric field is not large. From that we can derive the well-known Wave Equation that governs the propagation of light in free space.

Wave Equation:

$$\nabla^2 E = \mu_0 \varepsilon_0 \frac{\partial^2 E}{\partial t^2} \tag{1.20}$$

Also note that the Wave Equation for the magnetic field can be derived in exactly the same manner and the following result:

$$\nabla^2 H = \mu_0 \varepsilon_0 \frac{\partial^2 H}{\partial t^2} \tag{1.21}$$

1.1.5 Solutions of Wave Equation

If we assume the direction of light propagation is along z axis, the wave equation can be simplified as:

$$\frac{\partial^2 E}{\partial z^2} = \mu_0 \varepsilon_0 \frac{\partial^2 E}{\partial t^2} \tag{1.22}$$

And since the speed of light, c is defined as:

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} \tag{1.23}$$

Finally, the simplified Wave Equation can be written as:

$$\frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} \tag{1.24}$$

Similarly for magnetic field it can be written as:

$$\frac{\partial^2 H}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 H}{\partial t^2} \tag{1.25}$$

The simplest plane wave solutions to both equations are sinusoidal waves, which can be written as:

$$E = E_0 \cos(kz - \omega t) \tag{1.26}$$

$$H = H_0 \cos(kz - \omega t) \tag{1.27}$$

Here k is called the propagation constant and ω is the angular frequency of light. They are defined as: $k = \frac{2\pi}{\lambda},\ \omega = 2\pi f$

Plug in the electric field solution to the electric field Wave Equation and one gets:

$$-k^2 E_0 \cos(kz - \omega t) = -\frac{1}{c^2} E_0 \omega^2 \cos(kz - \omega t) \tag{1.28}$$

Therefore, $c^2 = \frac{\omega^2}{k^2}$ has to be true. Using the definition of k and ω shown above, it can be reduced to the well-known definition of speed of light, $c = \lambda f$; therefore, the equation is satisfied.

It can also be shown from Maxwell's Equation, if we assume $E_y = E_z = 0$; we can relate electric field to magnetic field as:

$$\frac{dE_x}{dz} = -\mu_0 \frac{dH_y}{dt} \tag{1.29}$$

By using the sinusoidal solutions in Eq. (1.3) it can be shown that

$$\mathrm{kE}_0 = \omega \mu_0 H_0 \tag{1.30}$$

therefore:

$$\frac{E_0}{B_0} = \frac{E}{B} = \frac{\omega}{k} = c \tag{1.31}$$

Other plane wave solutions include:

$$E_{forward} = E_0 \sin(kz - \omega t) \tag{1.32}$$

$$E_{backward} = E_0 \sin(kz + \omega t) \tag{1.33}$$

$$E(z,t) = \frac{1}{2} E_0 e^{i(kz-\omega t)} + c.c. \tag{1.34}$$

$$E = E_0 e^{i(kz-\omega t)} \tag{1.35}$$

A more complicated spherical wave solution can also be found as:

$$E(r,\ t) = E_0 \frac{e^{i(kr-\omega t)}}{r} \tag{1.36}$$

Note that instantaneous intensity of the light is related to the electric field as:

$$I = \varepsilon_0 c |E|^2 \tag{1.37}$$

Therefore average intensity can also be written as:

$$\langle I \rangle = \varepsilon_0 c \left\langle |E_0 \cos(\omega t)|^2 \right\rangle = \frac{1}{2} \varepsilon_0 c E_0^2 \tag{1.38}$$

Typical unit for both intensities are in watts per centimeter squared or W/cm^2.

Light power in watts, therefore, is simply defined as intensity divided by area or P = I/A..

1.2 Refractive Index and Absorption

Refractive index and absorption are the two most fundamental optical properties of any optical material systems. Refractive index, in the most simplified sense, determines the speed of light in a material, as the speed of light in a material with refractive index, n, is defined as, $v = c/n$. Since most light waves are non-monochromatic and non-planar, the refractive index affects both temporal and spatial aspects of light waves. Absorption is usually the most dominant loss mechanism in optics, although one can always find a transparent band for any given material system.

To understand the origin of both refractive index and absorption, I will use a simple harmonic oscillator model to derive both refractive index and absorption. The harmonic oscillator is schematically shown in Fig. 1.2.

Here, E is the driving electric field that drives the harmonic oscillator; e is an electron with mass, m; x is the time-varying separation between the charges. Using the Abraham-Lorentz Equation of Motion, we can write the characteristic equation of this harmonic oscillator as:

$$\ddot{x}(t) + 2\gamma\dot{x}(t) + \omega_0^2 x(t) = \frac{e}{m}E(t) \tag{1.39}$$

Note $\ddot{x} = \frac{d^2x}{dt^2}, \dot{x} = \frac{dx}{dt}$

The first term of the equation is the acceleration term, the second term is the velocity term where γ is damping constant, and ω_0 is the resonant frequency. General solutions of the time-varying separation between the charges and the driving electric field can be written as:

$$x(t) = \frac{x_0 \mathrm{X}(t)}{2}\left(e^{-i\omega t} + e^{i\omega t}\right) \tag{1.40}$$

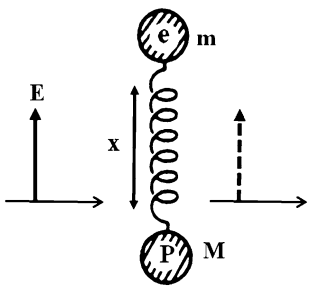

Fig. 1.2 Harmonic oscillator model

$$E(t) = \frac{E_0}{2}\left(e^{-i\omega t} + e^{i\omega t}\right) \tag{1.41}$$

where X(t) is a time-varying unit-less complex amplitude, which is an intrinsic property of the harmonic oscillator that determines how the oscillator responds to the driving electric field. This is the term that leads to the refractive index and absorption. By plugging in Eqs. (1.43) and (1.44) into (1.42) and neglecting small terms such as:

$$\ddot{X}(t) \text{ and } \gamma\dot{X}(t)$$

and ignoring negative frequencies, the equation can be written as:

$$-i\omega x_0\dot{X}(t) - \frac{x_0\omega^2}{2}X(t) - i\omega\gamma x_0 X(t) + \frac{\omega_0 x_0 X(t)}{2} = \frac{e}{m}\frac{E_0}{2} \tag{1.42}$$

This can be written in terms of $\dot{X}(t)$ as:

$$\dot{X}(t) = -\left[\gamma + \frac{i\left(\omega_0^2 - \omega^2\right)}{2\omega}\right]X(t) + \frac{ieE_0}{2\omega m x_0} \tag{1.43}$$

In steady-state, we can set $\dot{X}(t) = 0$ and write:

$$X(t) = \frac{ieE_0}{2\omega m x_0}\left[\frac{2\omega}{2\omega\gamma + (\omega_0^2 - \omega^2)}\right] \tag{1.44}$$

By using the rotating wave approximation, where $(\omega_0^2 - \omega^2) \approx 2\omega(\omega_0 - \omega)$ finally, the equation can be written as:

$$X(t) = \frac{ieE_0}{2\omega m x_0}\left[\frac{\gamma - i(\omega_0 - \omega)}{\gamma^2 + (\omega_0 - \omega)^2}\right] \tag{1.45}$$

As stated earlier, X is a complex value and can be defined as

$$X(t) \equiv U - iV \tag{1.46}$$

Therefore, U and V, the real and imaging part of X, respectively, can be written as:

$$U = \frac{eE_0}{2\omega m x_0}\frac{(\omega_0 - \omega)}{\gamma^2 + (\omega_0 - \omega)^2} \tag{1.47}$$

$$V = -\frac{eE_0}{2\omega m x_0}\frac{\gamma}{\gamma^2 + (\omega_0 - \omega)^2} \tag{1.48}$$

From the definition of electric polarization, P, it can be written as:

$$P = N(z)ex_0X \tag{1.49}$$

where N is the number of oscillators. Also from the definition of P in terms of electric susceptibility,

$$P = \varepsilon_0 \chi E_0 \tag{1.50}$$

$$P = N(z) e x_0 X \equiv \varepsilon_0 \chi E_0 \tag{1.51}$$

By setting two equations to equal, one can see that χ is related to X as:

$$\chi = \frac{N(z) e x_0 X}{\varepsilon_0 E_0} \tag{1.52}$$

Since χ is also a complex number defined as $\chi = \chi' + i\chi''$
The real and imaginary number of the electric susceptibility can be written as:

$$\chi' \equiv \frac{N(z) e x_0}{\varepsilon_0 E_0} U \tag{1.53}$$

$$\chi'' \equiv \frac{N(z) e x_0}{\varepsilon_0 E_0} V \tag{1.54}$$

From the Wave Equation in the previous section, if we don't ignore P, it can be shown that the real part of the electric susceptibility is related to the refractive index as:

$$n = 1 + \frac{\chi'}{2} \tag{1.55}$$

Therefore, n can finally be written as:

$$n = 1 + \frac{N(z) e^2}{4 \varepsilon_0 \omega m} \frac{(\omega_0 - \omega)}{\gamma^2 + (\omega_0 - \omega)^2} \tag{1.56}$$

Figure 1.3 shows the plot of n as a function of angular frequency. Values used are N = 10^{15} cm^{-3}, λ_0 = 1,500 nm, $\gamma = 10^{14}$ s^{-1}. One can deduce several facts from the equation and the plot.

1. Index is highest near the resonant frequency and decreases as the frequency is decreased.
2. Even in off-resonance, the index is not constant.
3. In the absence of other resonances, the index at the center of the resonance is 1.

Similarly from the Wave Equation in the previous section, it can be shown that the imaginary part of the electric susceptibility is related to the loss coefficient, α, as:

$$\alpha = k_0 \frac{\chi''}{2} \tag{1.57}$$

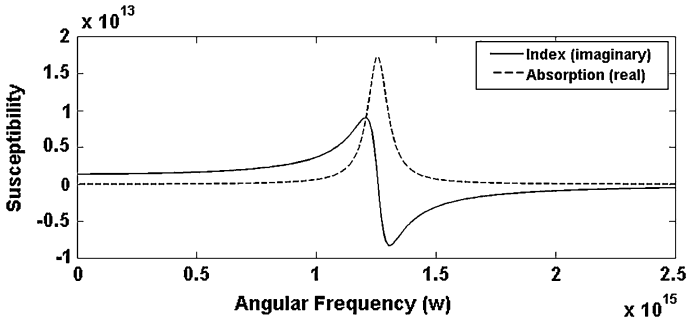

Fig. 1.3 Susceptibility as a function of angular frequency

Therefore the loss coefficient can be written as:

$$\alpha = k_0 \frac{N(z)e^2}{4\varepsilon_0 \omega m} \frac{\gamma}{\gamma^2 + (\omega_0 - \omega)^2} \tag{1.58}$$

For SiO_2 glasses the main resonance lies below $\sim$350 nm and between visible and near IR regime, for all practical purposes, there are no absorptions.

1.3 Slab Waveguides

Light travels in free space with velocity equal to $c = 3 \times 10^8$ m/s and a propagation constant of $k_0 = \omega_0/c$. The velocity and direction of the light depend on the basic physical properties of the media the light propagates in. Since the light is an ultrahigh frequency electromagnetic wave, how the electric and the magnetic waves of the light interact with the media will determine how the light propagates. In most cases—at least within the scope of this book—we will only deal with cases where the light travels in dielectric media; thus, only the interaction of the electric field of the light with the media will be considered.

The light with a finite beam size will diffract and diverge while propagating. The beam can be focused or collimated by using lenses and mirrors. Alternatively, light can be transmitted through dielectric conduits, i.e., optical waveguides.

Examples of widely used optical waveguides include slabs, strips, and fibers. They are described in more detail in this chapter.

Slab waveguides are one of the basic and simplest waveguide structures. Basic properties of slab waveguides will be reviewed in this section. This will be a great segue to a more complicated cylindrical waveguide (optical fiber) described later in this chapter. Figure 1.4 shows a slab waveguide structure with a guiding layer of index n_g and substrate and cover materials of index n_s and n_c, respectively. Without losing generality, we assume that they have the relation $n_g > n_s > n_c$. Snell's law states that the incident and refracted light satisfy the condition,

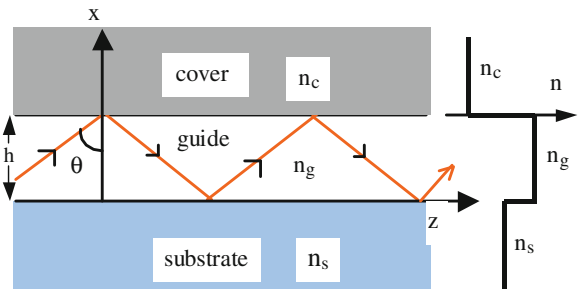

Fig. 1.4 Side view of a slab waveguide showing the zigzag waves corresponding to a guided mode

$n_1 \sin\theta_1 = n_2 \sin\theta_2$, where θ_1 and θ_2 are the angles between the incident or refracted light and the boundary normal, respectively. When the incident angle is larger than the critical angle θ_c, defined as $\theta_c = \sin^{-1}(n_2/n_1)$, light gets reflected 100 %. When this is the case for both top and bottom interfaces between the guiding layer, the cladding, and the substrate, light travels along z-axis in a zigzag pattern, as shown below:

1.3.1 Total Internal Reflection and Phase Shift at the Boundary

Consider light incident at an angle θ_1 on the interface of two different media which are isotropic, homogeneous, lossless, and having refractive index n_1 and n_2, respectively. As shown in Fig. 1.5. Two different cases can be considered, depending on the orientation of the electric and magnetic field relative to the incident surface normal.

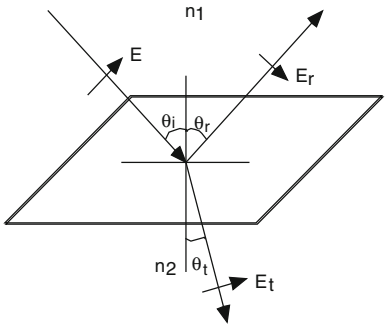

Fig. 1.5 Light reflection and transmission at the boundary

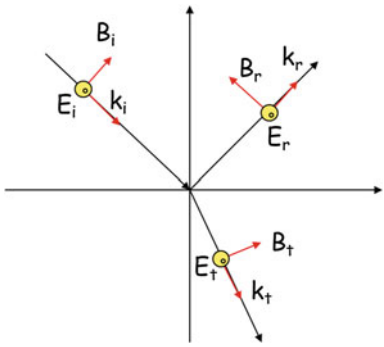

Fig. 1.6 TE wave

1.3.1.1 Transverse Electric Reflection

If the electric field is parallel to the plane of incidence, this is referred to as transverse electric (TE) field. This is schematically shown in Fig. 1.6.

Since fields parallel to the plane of incidence have to be continuous, we can write two boundary conditions:

$$E_i + E_r = E_t \tag{1.59}$$

$$B_{ix} + B_{rx} = B_{tx} \tag{1.60}$$

The boundary condition for the magnetic field can be written in terms of electric field as:

$$n_i E_i \cos\theta - n_i E_r \cos\theta = n_t E_t \cos\theta_t \tag{1.61}$$

By combining two equations:

$$n_i E_i \cos\theta - n_i E_r \cos\theta = n_t (E_i + E_r) \cos\theta_t \tag{1.62}$$

We can obtain reflection efficient, r, for TE as:

$$r_{TE} = \frac{E_r}{E_i} = \frac{n_i \cos\theta - n_t \cos\theta_t}{n_i \cos\theta + n_t \cos\theta_t} \tag{1.63}$$

When the incidence angle is zero, the equation reduces to a well-known reflection equation as:

$$r = \frac{n_1 - n_2}{n_1 + n_2} \tag{1.64}$$

Note that the power reflection coefficient, R, is defined as $R = |r|^2$

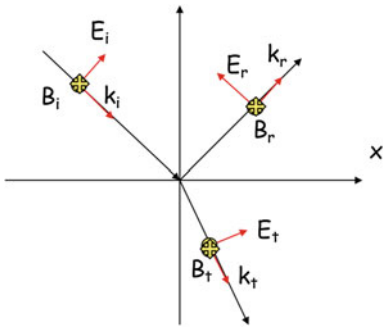

Fig. 1.7 TM wave

1.3.1.2 Transverse Magnetic Reflection

Similarly when the magnetic field is parallel to the plane of incidence, this is referred to as transverse magnetic (TM) field. This is schematically shown in Fig. 1.7.

The same boundary conditions exist as the TE case and this can be written as:

$$B_i + B_r = B_t \tag{1.65}$$

$$E_{ix} + E_{rx} = E_{tx} \tag{1.66}$$

These terms can be written in terms of electric fields as:

$$n_i E_i + n_i E_r = n_t E_t \tag{1.67}$$

$$E_i \cos\theta - E_r \cos\theta = E_t \cos\theta_t \tag{1.68}$$

Combining these two equations and solving for the ratio of reflected field to the incident field results in the reflection coefficient for the TM wave:

$$r_{TM} = \frac{E_r}{E_i} = \frac{n_t \cos\theta - n_i \cos\theta_t}{n_t \cos\theta + n_i \cos\theta_t} \tag{1.69}$$

Using the Snell's equation,

$$n_1 \sin\theta_i = n_2 \sin\theta_t \tag{1.70}$$

The Fresnel equations in terms of incident angle only can be written for TE and TM waves:

$$r_{TE} = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t} = \frac{n_1 \cos\theta_i - \sqrt{n_2^2 - n_1^2 \sin^2\theta_i}}{n_1 \cos\theta_i + \sqrt{n_2^2 - n_1^2 \sin^2\theta_i}} \tag{1.71}$$

$$r_{TM} = \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_2 \cos\theta_i + n_1 \cos\theta_t} = \frac{n_2^2 \cos\theta_i - n_1\sqrt{n_2^2 - n_1^2 \sin^2\theta_i}}{n_2^2 \cos\theta_i + n_1\sqrt{n_2^2 - n_1^2 \sin^2\theta_i}} \tag{1.72}$$

If $n_1 > n_2$ the reflection coefficients for both TE and TM become complex when $n_2^2 < n_1^2 \sin^2\theta_i$ or the incident angle is greater than $\sin^{-1}\frac{n_2}{n_1}$. This results in total internal reflection or 100 % of the light gets reflected. The angle at which this occurs is called the **critical angle** and is defined as: $\theta_c = \sin^{-1}\frac{n_2}{n_1}$

Since they are complex values, phase shift occurs when both TE and TM waves reflect at the boundaries, which are given by:

$$\tan\phi_{TE} = \frac{\sqrt{n_1^2 \sin^2\theta_i - n_2^2}}{n_1 \cos\theta_i} \tag{1.73}$$

$$\tan\phi_{TM} = \frac{n_1^2}{n_2^2}\frac{\sqrt{n_1^2 \sin^2\theta_i - n_2^2}}{n_1 \cos\theta_i} \tag{1.74}$$

By substituting subscript *1* to *g* (the guiding material) and subscript 2 to either c (the cover) or s (the substrate material), we obtain the phase shifts acquired at those boundaries. From the above Eqs. (1.73 and 1.74), we can see that the phase shift, ϕ_{TE} or ϕ_{TM}, increases from 0 at the critical angle to $\pi/2$ at the grazing incidence.

1.3.2 Dispersion Curves

To study the time-independent waves guided by a step-index thin-film waveguide, we begin with time-independent Maxwell's equations for isotropic, nonmagnetic media with $\mu = \mu_0$ and $\varepsilon = \varepsilon_0 n^2$

$$\nabla \times \mathbf{E} = -j\omega\mu_0\mathbf{H} \tag{1.75}$$

$$\nabla \times \mathbf{H} = j\omega\varepsilon_0 n^2\mathbf{E} \tag{1.76}$$

$$\nabla \cdot \mu_0\mathbf{H} = 0 \tag{1.77}$$

$$\nabla \cdot \varepsilon_0 n^2\mathbf{E} = 0 \tag{1.78}$$

All the guided field components must vary as $e^{-j\beta z}$ and the propagation constant β needs to be determined. The self-consistency condition- also known as the transverse resonance condition- states

$$2k_0 n_g h\cos\theta - 2\phi_s - 2\phi_c = 2m\pi \tag{1.79}$$

where m is an integer. This equation describes the constructive interference condition for the transverse components of the field. We can see that only certain discrete values of θ can satisfy the above equation. These are called modes and the integer m identifies the mode number.

For the mth guided mode of a slab waveguide, the propagation constant $\beta_m = \omega/\upsilon_p = k_0 n_g \sin\theta_m$, is the z-component of the wave vector $k_0 n_g$. The propagation constants are bounded by the plane wave propagation constants of substrate and the guide layer: $k_0 n_s < \beta < k_0 n_g$. At the cutoff frequency, the propagation constant β assumes the value of the lower bound $k_0 n_s$. As ω (or the thickness h) increases, β approaches its upper bound $k_0 n_g$ and the number of guided modes increases. While the guided modes are discrete, the radiation modes are continuous.

Based on the descriptions above, the effective index, N, of the guided mode can be defined as:

$$N = \beta/k = n_g \sin\theta \tag{1.80}$$

which is bounded by $n_s < N < n_g$. Define normalized frequency V as:

$$V = kh\sqrt{n_g^2 - n_s^2} \tag{1.81}$$

and a normalized guide index b related to the effective index N as:

$$b = \frac{(N^2 - n_s^2)}{(n_g^2 - n_s^2)} \tag{1.82}$$

The index b is zero at cutoff and approaches unity far away from it. For TE mode, by using (1.80 and 1.81), we can rewrite equation (1.79) to be:

$$V\sqrt{1-b} = m\pi + \tan^{-1}\sqrt{b/(1-b)} + \tan^{-1}\sqrt{(b+a)/(1-b)} \tag{1.83}$$

where a is the indicator of the waveguide asymmetry defined as

$$a = \frac{(n_s^2 - n_c^2)}{(n_g^2 - n_s^2)} \tag{1.84}$$

For a sputtered glass slab waveguide, with $n_g = 1.62$, $n_s = 1.515$, $n_c = 1$, and $a = 3.9$, Fig. 1.8 shows the normalized dispersion curves of the guide index b as a function of the normalized frequency, V, of the first four guided modes.

1.3.3 Field Distributions

TE mode has three field components: E_y, H_x and H_z.

$$H_y = E_x = E_z = 0 \tag{1.85}$$

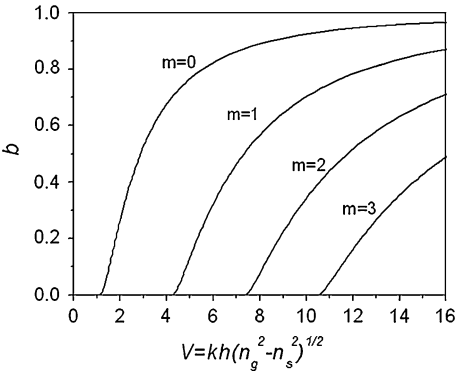

Fig. 1.8 Normalized dispersion diagram of a planar slab waveguide with $n_g = 1.62$, $n_s = 1.515$, $n_c = 1$

$$H_x = -(\beta/\omega\mu_0)E_y \tag{1.86}$$

$$H_z = (j/\omega\mu_0)\partial E_y/\partial x \tag{1.87}$$

$$\partial^2 E_y/\partial x^2 = (\beta^2 - n^2 k_0{}^2)E_y \tag{1.88}$$

where $k_0{}^2 = \omega^2\mu_0\varepsilon_0$. Solving the differential Eq. (1.88) subject to the boundary conditions at $x = 0$ and $x = h$ provides the mode field distributions. From the above analysis, we know $k_0{}^2 n_g{}^2 - \beta^2 > 0$ and $k_0{}^2 n_s{}^2 - \beta^2 < 0$ $k_0{}^2 n_c{}^2 - \beta^2 < 0$, such that possible solutions of *Ey* is a linear combination of sine and cosine functions inside the guiding layer, while they are exponentially decaying functions in the cover and substrate layers. Therefore, for the guiding layer region:

$$E_y = C_1 \cos(k_g x + \phi), \quad k_g = \sqrt{k_0{}^2 n_g{}^2 - \beta^2}, \quad 0 \le x \le h \tag{1.89}$$

$$E_y = C_2 e^{\gamma_s x}, \quad \gamma_s = \sqrt{\beta^2 - k_0{}^2 n_s{}^2}, \quad x \le 0 \tag{1.90}$$

$$E_y = C_3 e^{-\gamma_c (x-h)}, \quad \gamma_c = \sqrt{\beta^2 - k_0{}^2 n_c{}^2}, \quad x \ge h \tag{1.91}$$

The unknown constants, C_1, C_2, C_3 and ϕ, can be determined by the boundary conditions; *Ey* and *dEy/dx* are continuous. From these we can finally write down the field distribution for TE modes:

$$E_y = C_1 \left[\cos(V\sqrt{1-b}\frac{x}{h}) - \sqrt{\frac{b}{1-b}} \sin(V\sqrt{1-b}\frac{x}{h}) \right], \quad 0 \le x \le h \tag{1.92}$$

$$E_y = C_1 e^{-V\sqrt{b}(x/h)}, \quad x \le 0 \tag{1.93}$$

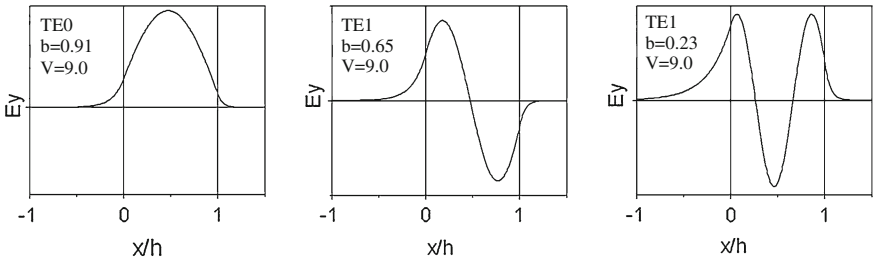

Fig. 1.9 Field distributions of first 3 guided TE modes in a symmetric dielectric waveguide (V = 9.0, with $n_s = 1.515$, $n_g = 1.62$, $n_c = 1$, and a = 3.9)

$$E_y = C_1 \left[\cos(V\sqrt{1-b}) + \sqrt{\frac{b}{1-b}} \sin(V\sqrt{1-b}) \right] e^{-V\sqrt{a+b}(x-h)/h}, \quad x \geq h \tag{1.94}$$

The first 3 guided TE modes in the above waveguide for V = 9.0 are plotted in Fig. 1.9.

The dispersion curves and field distributions for TM modes can be derived in the same manner as we did for TE modes.

1.4 Optical Fibers

Glass was one of the first materials used in the earliest days of optics to make a dielectric waveguide because of the ease of forming it into different shapes and the availability of a large number of dopants that change its optical materials. Much of this knowledge came from thousands of years of glass works in human history. Glass is a transparent, amorphous solid. Due to its transparency in the visible and near-IR spectral regimes, glass was mostly investigated for communications applications. The basic component of glass is fused silica (SiO_2). Proper amounts of other oxides can be added to adjust the refractive index as well as change other characteristics, such as the mechanical strength and chemical durability of the glass. Figure 1.10 shows the refractive index of various silica-based forms of glass as the result of adding ZrO_2, TiO_2, Al_2O_3, or GeO_2. These dopants result in an increase of the glass refractive index; however, the index can also be decreased by adding B_2O_3 or F.

The optical properties of fibers also depend on external environmental conditions. For example, pure bulk silica has a thermal linear expansion coefficient, $\alpha_{\text{Therm}} = 5.5 \times 10^{-7}/^{\circ}\text{K}$[3], and a temperature-dependent refractive index change, $n_{\text{Therm}} = 1.2 \times 10^{-5}/^{\circ}\text{K}$[10]. Therefore, changes in the height of the guiding layer due to temperature can be written as:

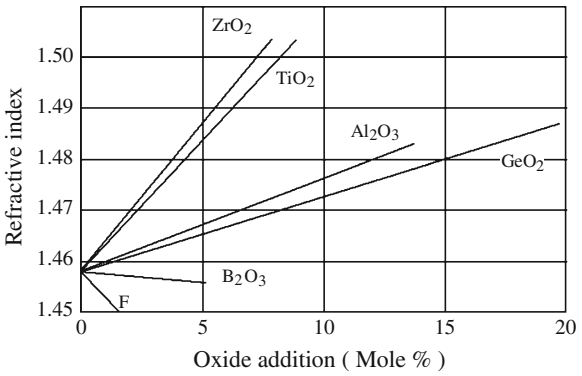

Fig. 1.10 Refractive index of doped silica glasses at 0.5893 μm (calculated based on C. L. Chen, Elements of Optoelectronics and Fiber Optics, p 467 1996)

$$\Delta h = \alpha_{Them} h \Delta T. \tag{1.95}$$

Similarly, the index change of the guiding or cladding layer due to temperature change can be written as:

$$\Delta n_{core,\, cladding} = n_{Them} \Delta T \tag{1.96}$$

Although this appears to be a small value, this index change due to temperature can be comparable to the changes in index the fiber optic sensor is designed to detect and can greatly affect the detection sensitivity.

Mechanical stress can also change the refractive index and dimension of the fibers. The photo elastic coefficient of the silica glass is $C = -4.2 \times 10^{-12}/\text{Pa}$. Therefore, giving a pressure, P, the change of the index of the core or cladding layers can be written as: $\Delta n_{core,\, cladding} = CP$.

Here, we assume the effect is isotropic, which is true for bulk glass; however, for non-isotropic waveguide structures, the effect is anisotropic.

Changes in the surrounding temperature and pressure change the refractive index and the physical dimensions of the fiber, which changes the effective indices for the TE and TM modes.

The basic geometry of optical fiber is shown in Fig. 1.11. To confine light so that the light travels in guided mode in a fiber, the core has to have a higher refractive index, n_1, than the cladding index, n_2, as shown in Fig. 1.12. Note that the plastic jacket has been omitted in the index profile. Different dopants can be

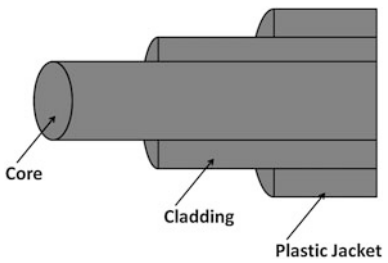

Fig. 1.11 Cross section of an optical fiber

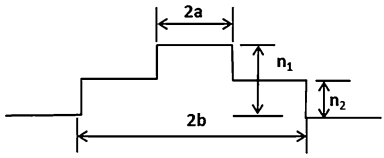

Fig. 1.12 Index profile of a step-index fiber

used to modify the refractive index of silica: P_2O_5 and GeO_2 are commonly used to increase the refractive index of SiO_2 glass, while B_2O_3 and F are used to reduce the index of SiO_2 glass.

1.4.1 Fiber Modes

Fiber modes can be calculated by solving the wave equation as shown in (1.97):

$$\left(\nabla^2 + k^2\right) E_z\left(z, r, \varphi, t\right) = 0 \tag{1.97}$$

where z is the direction along the fiber, E_z is the longitudinal electrical field; $\bigtriangledown^2$ indicates the Laplacian operator and $k^2 = \omega^2 n^2 / c^2$. ω is the angular frequency of light wave; n is the refractive index of medium, which equals n_1 for the core and equals n_2 for the cladding; c is the speed of light in vacuum. Although we only show E_z in (2, 1), E_z can be simply replaced with H_z to find out the magnetic field distribution in a fiber; the following analysis is similar for electrical field and magnetic field.

Due to the cylindrical shape of the fiber and because the refractive index profile of fiber is cylindrically symmetric, it is more convenient to use a cylindrical coordinate for describing the mode field profiles of optical fibers. In a cylindrical coordinate, the transverse electric and magnetic fields of light wave can be expressed as E_r, E_φ, H_r, and H_φ. Transverse components can be obtained with Maxwell's curl equation, given that the axial components E_z and H_z are known by solving (1.97). For reference electric and magnetic fields, a Laplacian operator in a cylindrical coordinate can be written as:

$$\begin{aligned} E &= E_r \hat{r} + E_\phi \hat{\phi} + E_z \hat{z} \\ H &= H_r \hat{r} + H_\phi \hat{\phi} + H_z \hat{z} \\ \nabla &= \frac{\partial}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial}{\partial \phi}\hat{\phi} + \frac{\partial}{\partial z}\hat{z} \\ \nabla^2 &= \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2} \end{aligned} \tag{1.98}$$

If a wave propagates along the fiber in z direction and the field is cylindrically symmetric, a general solution of the Wave Equation in a cylindrical coordinate can be written as:

$$E_z(z,r,\varphi,t) = E_z(r)\exp[i(\omega t - \beta z)]\exp[\pm il\varphi] \tag{1.99}$$

In (1.99), l is an integer; β is the propagation constant, as in the case of planar waveguides.

Expanding $\bigtriangledown^2$ in cylindrical coordinate and using (1.98), we may write (1.97) more explicitly as (1.100), which is a Bessel equation of order l:

$$\frac{\partial^2 E_z(r)}{\partial r^2} + \frac{1}{r}\frac{\partial E_z(r)}{\partial r} + \left(k^2 - \beta^2 - \frac{l^2}{r^2}\right)E_z(r) \tag{1.100}$$

Confined propagation requires:

$$\frac{\omega n_2}{c} < \beta < \frac{\omega n_1}{c} \tag{1.101}$$

Considering the behavior of fields when $r \to 0$ and $r \to \infty$, we may express the field in core and cladding as:

$$\begin{aligned} E_{core}(r) &= AJ_l(hr), \text{ for } r < a \\ E_{cladding}(r) &= BK_l(qr), \text{ for } r > a \end{aligned} \tag{1.102}$$

In (1.102), A and B are constants, J_l is first kind of Bessel function, K_l is the first kind of modified Bessel function where wave vectors h and q are defined as:

$$h^2 = (\omega n_1/c)^2 - \beta^2; q^2 = \beta^2 - (\omega n_2/c)^2$$

To make sure the solution satisfies the boundary condition, that E_φ, E_z, H_φ, and H_z are continuous at the boundary where $r = a$, after a lengthy derivation, we can obtain the following characteristic equation that determines the propagation constant β of the guided modes:

$$\left[\frac{J_l'(ha)}{haJ_l(ha)} + \frac{K_l'(qa)}{qaK_l(qa)}\right]\left[n_1^2\frac{J_l'(ha)}{haJ_l(ha)} + n_2^2\frac{K_l'(qa)}{qaK_l(qa)}\right] = \left[\frac{1}{(qa)^2} + \frac{1}{(ha)^2}\right]^2\left(l\frac{\beta c}{\omega}\right)^2 \tag{1.103}$$

(1.103) is a quadratic equation in $J_1'/[haJ_1(ha)]$ and this equation has two roots corresponding to two classes of solutions. With $l = 0$ in (1.103), the modal conditions can be separated into (1.104) and (1.105), corresponding to TE and TM modes, respectively:

$$\frac{J_1(ha)}{haJ_0(ha)} = -\frac{K_1(qa)}{qaK_0(qa)} \tag{1.104}$$

$$\frac{J_1(ha)}{haJ_0(ha)} = -\frac{n_2^2K_1(qa)}{n_1^2qaK_0(qa)} \tag{1.105}$$

The wave vectors h and q have the following relationship:

$$(ha)^2 + (qa)^2 = \left(a\frac{\omega}{c}\right)^2 \left(n_1^2 - n_2^2\right) = V^2 \tag{1.106}$$

V is called the normalized frequency, which determines how many modes the fiber can support and is defined as:

$$V = \frac{2\pi a}{\lambda}\sqrt{n_1^2 - n_2^2} \tag{1.107}$$

Also, the numerical aperture, NA, and the normalized propagation constant, b, can be defined as:

$$NA = \sqrt{n_1^2 - n_2^2} \tag{1.108}$$

$$b = \frac{n_{eff}^2 - n_2^2}{NA^2} \tag{1.109}$$

where, $n_{eff}\ (= \beta/K)$ is the effective refractive index that corresponds to each guided mode in the fiber.

When NA is large, V is large: therefore, many modes are allowed to propagate as guided modes of the fiber. The number of guided modes, M, that an optical fiber supports, can be approximated by (1.110) given $V > 5$. M is also called mode volume.

$$M = \frac{4V^2}{\pi^2} \tag{1.110}$$

We can also define the normalized index difference Δ as:

$$\Delta = \frac{n_1 - n_2}{n_1} \tag{1.111}$$

The Eqs. (1.104) and (1.105) can be solved graphically. Figure 1.13 shows the left-hand side of (1.104) plotted in blue and the right-hand side of (1.104) with different values of V plotted in red curves. The intersections of blue and red curves correspond to allowed propagation constants that satisfy the characteristic equation. For $V = 2$, no intersection exists; therefore, this fiber does not support TE mode. For $V = 5$, only one intersection exists, which indicates that only one TE mode will be propagating in the fiber. With $V = 7$, two intersections can be found in Fig. 1.13, which means that two TE modes are supported.

The fundamental mode for a fiber with a circular cross section is the HE11 mode. For $V < 2.4048$, the fiber only supports the HE11 mode. Therefore, fibers having circular cross sections are single mode fibers if V is less than 2.4048. The transverse electric field associated with the HE11 mode is linearly polarized.

For weakly guided fibers, i.e. $n_1 \approx n_2$, having the core radius large compared to λ, the characteristic equation can be simplify to [1]:

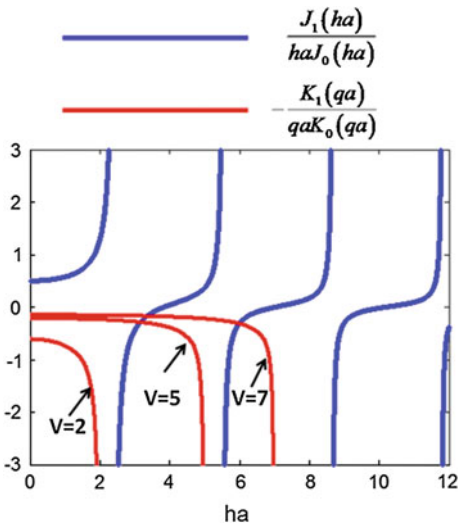

Fig. 1.13 Graphical solution of (1.104)

$$V\sqrt{1-b}\frac{J_{\ell-1}(V\sqrt{1-b})}{J_\ell(V\sqrt{1-b})}+V\sqrt{b}\frac{K_{\ell-1}(V\sqrt{b})}{K_\ell(V\sqrt{b})}=0 \tag{1.112}$$

Roots to this equation correspond to linear polarized modes and are labeled as LP modes [1]. The relation between a traditional mode designation and the LP mode designation is shown in Table 1.1. Figure 1.14 shows the b-V curves of several LP modes for V less than 6.

Table 1.2 also shows the cut-off values of normalized frequency V for LP modes l and m from 0 to 4:

The field distribution of LP modes with $l = 0$ can be deduced as:

$$E_x(r,\phi,z)=E_0e^{-jkn_{eff}z}\begin{cases}\frac{J_0(V\sqrt{1-br/a})}{J_0(V\sqrt{1-b})}, & 0\le r\le a\\ \frac{K_0(V\sqrt{br/a})}{K_0(V\sqrt{b})}, & r\ge a\end{cases} \tag{1.113}$$

$$E_z(r,\phi,z)=E_0\frac{j}{k_0an_{eff}}e^{-jk_0n_{eff}z}\cos\phi\begin{cases}\frac{V\sqrt{1-b}J_1(V\sqrt{1-br/a})}{J_0(V\sqrt{1-b})}, & 0\le r\le a\\ \frac{V\sqrt{b}K_1(V\sqrt{br/a})}{K_0(V\sqrt{b})}, & r\ge a\end{cases}. \tag{1.114}$$

$$H_z(r,\phi,z)=E_0\frac{j}{\eta_0k_0a}e^{-jk_0n_{eff}z}\sin\phi\begin{cases}\frac{V\sqrt{1-b}J_1(V\sqrt{1-br/a})}{J_0(V\sqrt{1-b})}, & 0\le r\le a\\ \frac{V\sqrt{b}K_1(V\sqrt{br/a})}{K_0(V\sqrt{b})}, & r\ge a\end{cases} \tag{1.115}$$

As shown in Fig. 1.15 and Table 1.1, the lowest order mode, LP01, has no cutoff. The higher modes are cut off for V less than 2.4048.

Table 1.1 Traditional versus LP mode designation

Normalized frequency V	Traditional mode designation	LP mode designation	Number of modes	Total number of modes
0 ~ 2.4048	HE_{11}	LP01	2	2
2.4048 ~ 3.8317	TE_{01}, TM_{01}, HE_{21}	LP11	4	6
3.8317 ~ 5.1356	EH_{11}, HE_{31}	LP21	4	10
	HE_{12}	LP02	2	12
5.1356 ~ 5.5201	EH_{21}, HE_{41}	LP31	4	16
5.5201 ~ 6.3802	TE_{02}, TM_{02}, E_{22}	LP12	4	20

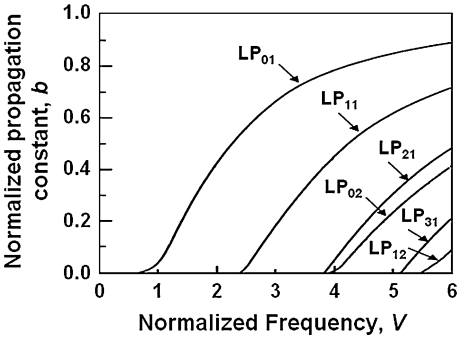

Fig. 1.14 Normalized propagation constant (b) as function of normalized frequency (V) for guided modes in step-index optical fiber

Table 1.2 Cutoff values of normalized frequency (V) for LP mode designation

V	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$l = 0$	0	3.832	7.016	10.173
$l = 1$	2.405	5.520	8.654	11.792
$l = 2$	3.832	7.016	10.173	13.323
$l = 3$	5.136	8.417	11.620	14.796
$l = 4$	6.379	9.760	13.017	16.224

1.4.2 Multi-Modal Effects

When light is launched into a fiber with a larger number of guided modes, it would mostly couple into a mode with a profile that mostly resembles the incident beam. If we define $|a_{lm}|^2$ as the power of each mode, the total power of all the modes can be written as:

$$P_{total} = \sum_{l'm'} |a_{l'm'}|^2 = \sum_{l'm'} \left| \iint E_{l'm'}(x,y) E^*_{l'm'}(x,y) dxdy \right|^2 \tag{1.116}$$

Therefore, mode coupling efficiency to a particular (l, m) transverse mode field distribution can be calculated as,

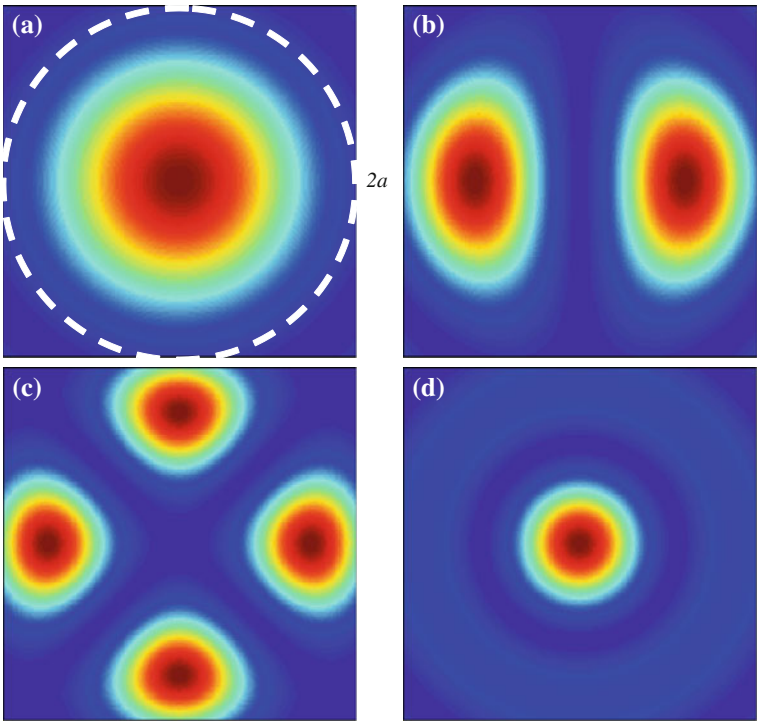

Fig. 1.15 Sample intensity profiles of several lower-order modes: **a** $LP_{01,}$ **b** $LP_{11,}$ **c** $LP_{21,}$ **d** LP_{02} ($2a$ diameter of fiber core)

$$\eta_{lm} = \frac{\text{Power coupled into mode lm}}{\text{Total incident power}} = \frac{|a_{lm}|^2}{\sum_{l'm'} |a_{l'm'}|^2} = \frac{\left|\iint E_{lm}(x,y)E^*_{lm}(x,y)dxdy\right|^2}{\sum_{l'm'} \left|\iint E_{l'm'}(x,y)E^*_{l'm'}(x,y)dxdy\right|^2}. \quad (1.117)$$

Mode matching is generally hard to do and the incident beam evolves into a large number of guided modes. Coupling into a large number of modes in a multimode fiber—unlike a single mode fiber with only a fundamental LP_{01} mode—introduces a large modal dispersion and significantly affects the pulse shape as the pulse propagates in the fiber. However, the exact solutions of coupled mode equations for all possible modes are very complicated, so they can be treated by assuming a modal continuum rather than by considering a number of individual modes where the mode coupling problem can be explained by a diffusion process. Thus, the impulse response of the multimode fiber can be written as:

$$P(z,t) = \Theta_0^2 \sqrt{\frac{\pi}{Tt}} \left(\frac{1}{2} + \frac{t}{\gamma_\infty z T}\right)^{-1} \exp\left(-\frac{t}{T} - \frac{\gamma_\infty^2 z^2 T}{4t}\right) \quad (1.118)$$

where $T = \frac{n}{2cA} = \frac{n\Theta_0^2}{2c\gamma_\infty}$, Θ_0 is the initial angular width at $z = 0$, γ_∞ is the steady-state attenuation coefficient for a very long fiber $(z \to \infty)$, n is the refractive index

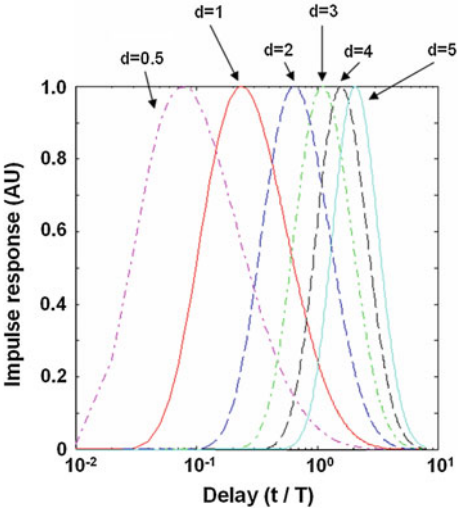

Fig. 1.16 Normalized impulse response as a function of normalized time delay (t/T) in a multimode fiber over a normalized distance ($d = \gamma_\infty z$)

of the fiber, and A is the loss coefficient measured in m^{-1} rad^{-2}. The results are plotted as a function of the normalized delay time (t/T) for various normalized lengths ($d = \gamma_\infty z$) and the impulse responses are normalized for equal peak value in Fig. 1.16.

For a multimode fiber, the mean pulse delay can be derived as a function of the propagation distance (z), where the mode coupling predominates and is based on the same diffusion model, the time delay is given as:

$$\tau_\theta = \frac{T}{2}\left[\gamma_\infty z + \left(\frac{\theta^2}{\Theta_0^2} - \frac{1}{2}\right)\left(1 - e^{-2\gamma_\infty z}\right)\right]. \tag{1.119}$$

The resultant Fig. 1.17a shows the variation of τ_∞ as a function of $\gamma_\infty z$ for $\theta = 0$, $\theta = \infty$, and in between possible modes for a few-mode fiber for $V < 5$. The results imply that for a relatively short fiber length, the delay between each mode is proportional to the propagation distance. Here, propagation angle of each mode is obtained by guided mode analysis and the corresponding result is shown in Fig. 1.17b.

1.4.3 Losses in Optical Fibers

Light guiding in optical fibers is similar to that of slab waveguides. The early fibers were extremely lossy, with a typical loss of 1,000 dB/km. The use of optical fibers in communications was not available until 1970 when fibers with a loss of 20 dB/km near 1 μm wavelength region were introduced. Current technology makes it possible to manufacture fibers with a loss as low as 0.2 dB/km.

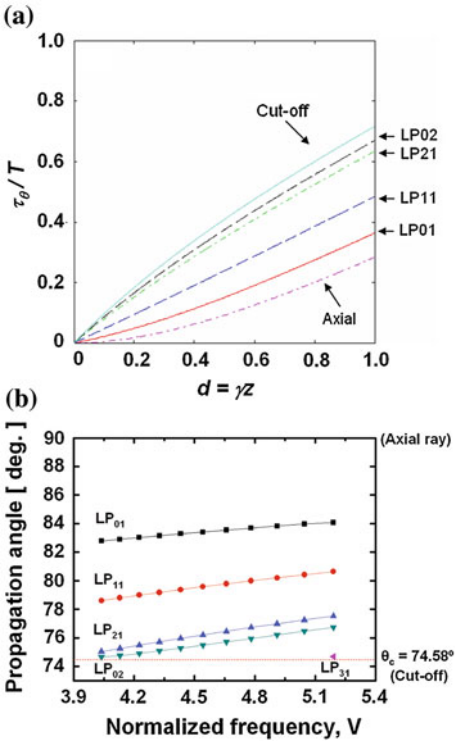

Fig. 1.17 Delay and propagation angle in multimode fiber: **a** time delays; **b** propagation angles

For both multimode and single-mode fibers, the losses depend strongly on wavelength (Fig. 1.18). Absorption and scattering of light traveling through a fiber are the main sources of the losses. Fused silica glass (SiO_2) exhibits two strong absorption bands: a middle-infrared absorption band resulting from vibrational transitions and an ultraviolet absorption band due to electronic and molecular transitions. Figure 1.18 shows the absorption tails of glass due to electronic transitions in the ultraviolet (UV) region and molecular vibrational mode transitions in the infrared (IR) region. In addition to the intrinsic absorptions there are extrinsic absorptions due to impurities, mainly OH^- ion vibrations associated with water vapor dissolved in the glass. The fundamental OH^- ion stretching vibrational mode is at 2.72 μm, and the harmonic overtones, 1.38, 0.95, 0.72 and 0.6 μm, are in the near-IR and visible region.

Rayleigh scattering is another intrinsic effect that contributes to the attenuation of light in glass. The random localized variations of the molecular positions in glass create random in homogeneities of the refractive index that act as tiny scattering centers. These microscopic variations of refractive index cause Rayleigh scattering. The scattered intensity is wavelength dependent and proportional to ω^4 or to $1/\lambda^4$. On the short wavelength side, less than 1.5 μm, losses are mainly determined by the intrinsic Rayleigh scattering. On the long wavelength side, loss

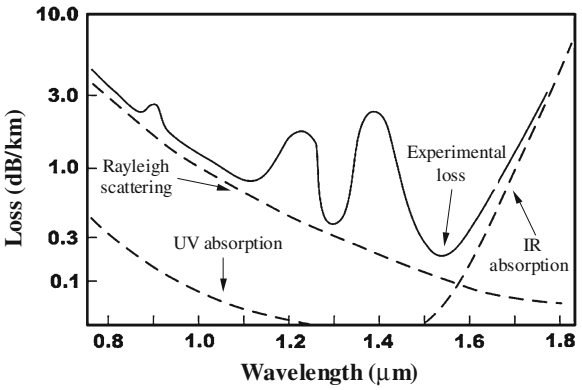

Fig. 1.18 Loss spectrum of silica fibers (graphed based on Miya [4]

is caused by the intrinsic molecular vibration absorption. Although current technology is able to decrease the OH^- concentration level extremely low, the absorption peaks can still be noticed.

1.4.4 Nonlinearity of Optical Fibers

From the two curl equations of the Maxwell Eqs. (1.75–1.78), we have

$$\nabla^2 E = \mu_0 \varepsilon_0 \frac{\partial^2 E}{\partial t^2} + \mu_0 \frac{\partial^2 P}{\partial t^2} \tag{1.120}$$

$$P = \varepsilon_0 \chi E = \varepsilon_0 \left(\chi^{(1)} + \chi^{(2)} E + \chi^{(3)} EE + \ldots \right) E = \varepsilon_0 \chi^{(1)} E + P_{NL} \tag{1.121}$$

where $P_{NL} = \varepsilon_0 \left(\chi^{(2)} E + \chi^{(3)} EE + \ldots \right) E$ and $\chi^{(j)}$ $(j = 1, 2, \ldots)$ is the jth-order susceptibility of the material. We rewrite the above equation to be:

$$\nabla^2 E = \frac{n^2}{c^2} \frac{\partial^2 E}{\partial t^2} + \mu_0 \frac{\partial^2 P_{NL}}{\partial t^2} \tag{1.122}$$

where $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$ is the light speed in vacuum and $n = \sqrt{1 + \chi^{(1)}}$ is the refractive index.

Glass is an isotropic centro-symmetric material, so $\chi^{(2)}$ vanishes; therefore, we can concentrate on the third order term $\chi^{(3)}$ only. Assuming the applied field is a monochromatic wave $E(t) = E_0 \cos(\omega t)$, plugging it into the expression of P_{NL} in (1.122), we have:

$$P^{(3)}(t) = \frac{1}{4} \varepsilon_0 \chi^{(3)} E_0^3 \cos(3\omega t) + \frac{3}{4} \varepsilon_0 \chi^{(3)} E_0^3 \cos(\omega t) \tag{1.123}$$

The first term generates third harmonic, frequency at 3ω, due to the applied field at frequency ω. The conversion efficiency is usually very low without the careful arrangement of phase matching. The other term is responsible for an optical Kerr effect, where the nonlinear refractive index change can be written as:

$$\Delta n = n_2 I \tag{1.124}$$

where n_2, the nonlinear refractive index coefficient is defined as $n_2 = \frac{3\eta_0}{n_0^2 \varepsilon_0}\chi^{(3)}$, where $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$, and I is the optical intensity of the incident wave defined as $I = \frac{|E(\omega)|^2}{2\eta}$. The nonlinear coefficients in silica fibers are quite small. It has been measured that n_2 in silica fibers is approximately $2.2 \sim 3.4 \times 10^{-20}$ m^2/W [2], which is 2 orders of magnitude smaller than most other nonlinear media. However, the nonlinear effect in fibers can be significantly enhanced due to the characteristics of the single mode fibers, i.e., small effective core area and low loss over a long propagation length. The first characteristic leads to high intensity and the second one enables long interaction length. In order to compare to other nonlinear bulk materials, we can define a figure of merit which is the product of the intensity and effective length, IL_{eff}. For the bulk materials, assuming beam is Gaussian,

$$IL_{eff} = \frac{P}{\pi w^2}\frac{\pi w^2}{\lambda} = \frac{P}{\lambda} \tag{1.125}$$

For fibers,

$$IL_{eff} = \int_0^L \frac{P}{\pi w^2} e^{-\alpha z} dz = \frac{P}{\pi w^2}\left(\frac{1 - e^{-\alpha L}}{\alpha}\right) \tag{1.126}$$

Therefore, the enhancement factor is:

$$\frac{IL_{eff-fiber}}{IL_{eff-bulk}} = \frac{\lambda}{\pi w^2 \alpha} \tag{1.127}$$

which can be as high as 10^9 around a wavelength of 1.55 μm.

1.4.5 Birefringence

Even a single-mode fiber can support two degenerate modes polarized in two orthogonal directions. The two modes would not couple to each other in a fiber with perfectly cylindrical geometry and isotropic material. However, the inevitable non-perfect cylindrical geometry or small fluctuations in material break the mode degeneracy and result in modal birefringence. The mode propagation constants become slightly different for the modes polarized in x and y directions [3].

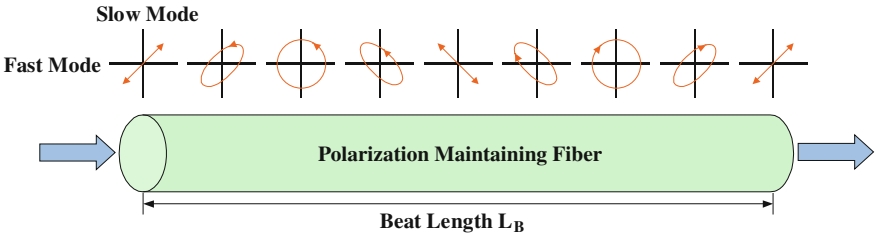

Fig. 1.19 Evolution of polarization along a PMF when input light is linearly polarized at an angle with the fast axis

$$B = \frac{|\beta_x - \beta_y|}{k_0} = |n_x - n_y| \tag{1.128}$$

where n_x and n_y are the effective indices in the two orthogonal polarization states. The axis along which n is larger is called the slow axis while the other one is called the fast axis. The two modes exchange energy with each other periodically as they propagate along the fiber with the period of one beat length, L_B, defined as:

$$L_B = \frac{2\pi}{|\beta_x - \beta_y|} = \frac{\lambda}{B} \tag{1.129}$$

Because conventional single-mode fibers have random birefringence fluctuations along the fiber, light launched into the fiber with linear polarization loses polarization within a short distance. To overcome the problem, polarization-maintaining fibers (PMF) are made in which a large amount of birefringence is introduced intentionally by making either the core or the cladding elliptical in shape, or by inserting stress-applying elements beside the core. The birefringence achieved is in the order of 10^{-6} to 10^{-4}. If the polarization of the incident light is coincident with the slow or the fast axis of the PMF, the polarization remains unchanged during propagation. If the polarization axis has an angle with these axes, the polarization changes periodically with a period equal to the beat length L_B, as shown in Figs. 1.19.

References

1. Gloge D (1971) Weakly Guiding Fibers. Appl Opt 10(10):2252–2258
2. Agrawal GP (1989) Nonlinear fiber optics. Academic Press, San Diego, CA (2nd edn, 1995; Chinese translation, 1992; Russian translation, 1996; Japanese translation, 1997)
3. Kaminow IP(1981) Polarization in optical fibers. IEEE J Quantum Electron (QE) 17:15–22
4. Miya T, Terunume Y, Hosaka T, Miyashita T (1979) Ultimate low-loss single-mode fiber at 1.55 pm, Electron Lett 15:106–108

Chapter 2
Fiber Optic Interferometric Devices

Utkarsh Sharma and Xing Wei

2.1 Introduction

Fiber optic interferometry can be broadly explained as the techniques that utilize the fundamental principles of optical interference to measure physical sample properties or detect changes via sensing systems that are partially or completely realized using fiber optic components. While the field of optical interference dates back to second half of seventeenth century, the advent of fiber optic interferometry technology is rather recent as it stemmed out of advances in the fiber optics in late 1970's and early 1980's. In the next two sections, we will briefly review the major advancements made in the field of optical interferometry and fiber optic interferometric devices.

2.1.1 Developments in Optical Interferometry: A Journey towards Understanding Light

Historically, the experiments based on optical interference over the last three centuries have played a key role in helping the physicists and scientists to make breakthrough advancements in Physics and led to wide range of applications involving highly sensitive metrology and sensing [1–3]. One of the unique advantages that the phenomenon of optical interference offered to the physicists was the capability to realize experimental optical setups and analyze the results that could be used to either verify or question their hypothesis and theories. The understanding of fundamentals of light and optical interference has come a long way since physicists first started exploring this subject matter more than three centuries ago. In one of the earliest experiments dating back to 1665, it was

U. Sharma (✉) · X. Wei
Carl Zeiss Meditec, Inc., 5160, Hacienda Drive, Dublin CA-94568, USA
e-mail: Utkarsh.1980@gmail.com

J. U. Kang (ed.), *Fiber Optic Sensing and Imaging*,
DOI: 10.1007/978-1-4614-7482-1_2,

F. M. Grimaldi's careful observations regarding the phenomenon of diffraction that first hinted at the wave-nature of the light. R. Hooke and C. Huygens furthered the wave theory of light but it was Newton's assertion of corpuscular theory of light that found broader acceptance for most of the seventeenth and eighteenth century. It was the remarkable contributions of E. Young, A. Fresnel and J. C. Maxwell that emphatically put the wave theory of light in the forefront again in the nineteenth century. Finally, in the era of the birth of modern physics and quantum mechanics in the twentieth century, the path-breaking works of M. Planck, A. Einstein, L. deBroglie, N. Bohr, W. Heisenberg and other notable physicists helped in establishing the wave-particle dual nature of light. It may be fitting to say that mankind's continually evolving understanding of the fundamentals of light and phenomenon such as optical interference has mirrored with the key advances made in the field of Physics over the last three centuries.

2.1.2 Fiber Optic Interferometry

While the impact of optical interferometry on modern physics cannot be overstated, it has been the advancements in the fiber optic interferometry technology that can be credited for utilizing this phenomenon for a wide range of commercial and industrial applications including strain sensors for structural monitoring, medical imaging, remote sensing and precise measurement applications to name a few [4–6]. There have been three critical factors that have been instrumental towards development of this technology. Firstly, it was the invention of laser that provided a high intensity light source with strong spatial and temporal coherence properties. Secondly, it was the advancements in optoelectronics industry that facilitated development of optical detectors to realize sophisticated low noise detection techniques to record optical interference effects with high sensitivity. And finally, it was the development of low loss single mode fibers that made it possible guide light in a flexible fiber optic waveguide and realize compact, low-cost, robust and versatile fiber optic interferometers that would otherwise be impractical to achieve with bulk-optic components.

Two of the major contributions of fiber optic interferometry have been in fiber optic sensors and fiber optic devices. In this chapter, we will largely focus on fiber optic sensors and its applications.

2.1.2.1 Advantages of Fiber Optic Technology

There are some unique advantages of fiber optics technology that make it especially attractive for sensing applications [5].

1. Compact and lightweight: Fiber optic interferometers and sensors can be made highly compact and hence find use in airborne and space-based sensing applications.
2. High accuracy and sensitivity compared to typical mechanical or chemical based sensors.
3. Flexibility for customized applications: Single mode fibers (SMF) provide a loss-less and flexible optical wave-guide for transmission, delivery and collection of optical signals. For example, one can use a flexible fiber to deliver or collect light from internal body organs by integrating an optical fiber with an endoscope. One can also utilize the nearly loss-less transmission properties of SMFs for remote sensing operations.
4. Silica material properties: Optical fibers are made from silica (glass) and hence carry some inherent advantages such as usability in harsh, high temperature and rugged environments, and immunity to electromagnetic interference. Silica is also a chemically passive material and hence it is not affected by corrosive factors that might be present in the environment.
5. Multiplexing capabilities that allow distributed sensing applications.

2.1.2.2 Fiber Optic Interferometry Based Sensing: Technology Trends and Applications

Fiber optic sensor technology was touted to be a huge breakthrough in 1980's and early 1990's. Although it led to several successful commercial applications and its impact on research and specialized engineering applications has been invaluable, it has still failed to achieve the widespread commercial penetration in the sensors market that it had once promised. The main reason for this failure is the stiff competition it has faced due to existence of low cost conventional alternatives such as electronic/mechanical or chemical sensors. The sensors industry is highly fragmented with various independent market sectors where each has very specialized needs [6]. Although fiber optic technologies could offer higher resolution and precision, such high sensitivities are often not required for most commercial applications.

It was the steep growth of telecommunication industry in the 1990's that fueled the research and development in fiber optics technology and led to availability of less costly, efficient and more sophisticated fiber optic components. However, the cost of fiber optic sensor technology still needs to be reduced significantly to offer a compelling incentive to sensors industry for adoption of the newer technology. The prices of fiber optic and optoelectronic components such as laser diodes and SMFs have seen a steep decrease by a factor of around hundred-folds over the last 30 years [5]. It is likely that this downward trend in prices will continue as the technology matures further, and this may open up newer possibilities for optical fiber sensor technology.

There is also an alternative route towards progressive commercial success for fiber optic sensor technology. The key to this approach would be to identify and develop niche applications that can benefit from the unique advantages of the fiber optic sensors. Environmental and atmospheric monitoring, industrial chemical processing, advanced manufacturing processes, biotechnology, and defense are some of the areas that can utilize the improved performance of the newer technology. With increased emphasis on monitoring the environmental changes, fiber optic sensors could play a pivotal role in real-time measurements of pollutant levels and contamination in the environment. The multiplexing and distributed sensing capabilities of fiber optic sensors make them an excellent candidate for structural health monitoring in airborne and satellite applications where weight is a major consideration. It can also be used to remotely monitor chemical processes in otherwise rugged and hazardous conditions. In medicine, fiber optics has been found extremely useful in realizing optical imaging techniques such as optical coherence tomography and delivering therapeutic light into an internal organ via optical fiber conduit. Overall, the field of fiber optic interferometry is bound to prosper and have a wider commercial impact in future as the technology evolves and finds newer applications.

In the following section, we will discuss the basic fundamentals of fiber optic interferometry and sensors. In Sect. 2.4, we will review basic principles of operations of some of the selected fiber optic components that are routinely used to realize fiber optic interferometric sensors and devices. In Sect. 2.3, several interferometric architectures will be covered along with a few selective examples of applications that use these interferometers.

2.2 Fundamentals of Optical Interferometry and Fiber Optic Sensors

In this section, we will review the basic fundamentals about the optical interference, and generic descriptions of fiber optic sensors. In the topics describing fiber optic concepts and applications, the analysis will be limited to single mode optical fibers for the scope of this chapter.

2.2.1 Optical Interference and Wave Representation of Light

As per the scalar wave model, monochromatic light constitutes of oscillating electric and magnetic field components that can be mathematically represented by the wavefunction $\psi(r,t)$, where it is a function of time, t and spatial position, $r = (x, y, z)$. As derived from Maxwell's equations in the earlier chapter, the light wavefunction satisfies the following scalar wave equation in vacuum,

$$\nabla^2\psi(r,t) = \frac{1}{c^2}\frac{\partial^2\psi(r,t)}{\partial t^2}, \tag{2.1}$$

where ∇^2 is the Laplacian operator, and c is the speed of light in vacuum. Any function that satisfies the above equation could be a possible representation of optical waves. A characteristic property of the scalar wave equation is that it is linear, and hence the principle of superposition can be applied if there is an interaction of two or more wavefunctions. If the two optical fields, $\psi_1(r,t)$ and $\psi_2(r,t)$, have an overlap then the resulting optical field is given by the linear superposition of the two waves as shown in the following equation:

$$\psi_{Tot}(r,t) = \psi_1(r,t) + \psi_2(r,t). \tag{2.2}$$

It is often mathematically convenient to represent the light waves as complex fields, although the electric or magnetic fields that carry the energy are represented by the real part of this complex field. Monochromatic plane wave solution is often used to describe various fundamentals of light as it is one of the simplest solutions for the scalar wave equation. A monochromatic plane wave can be mathematically represented as following:

$$\begin{aligned}\psi(\vec{r},t) &= E_0 \exp\left(i\left(\vec{k}.\vec{r} - 2\pi\upsilon t\right)\right)\\ &= E_0 \exp\left(i\left(k_x x + k_y y + k_z z - 2\pi\upsilon t\right)\right),\end{aligned} \tag{2.3}$$

where E_0 is the field amplitude, $\vec{k}$ is the propagation wavevector, and υ is the frequency of light. Often in a simplified form, a plane wave propagating in the z direction is represented as $E_0 \exp(i(kz - \omega t))$, where $\omega(= 2\pi\upsilon)$ is the angular frequency of the light. In context to this chapter, a modified plane wave based solution is also quite suitable to model light propagation in single mode fibers when weakly guiding approximation is used [Chap. 1].

2.2.1.1 Interference of Monochromatic Planar Waves

The resulting electric field due to interference of two co-propagating plane waves of the same frequency can be given as:

$$E_{Tot} = E_1 \exp(i(kz - \omega t + \phi_1)) + E_2 \exp(i(kz - \omega t + \phi_2)), \tag{2.4}$$

where, E_1, E_2 and ϕ_1, ϕ_2 are the amplitudes and the phases of the two waves respectively. The frequency of optical light is very high ($\sim 10^{14}$ Hz) and hence one can only detect the intensity of the light using optical detectors. The intensity is proportional to the time-averaged value of squared electric field: $I\alpha\langle E.E^*\rangle$ (refer Chap. 1 for more details). The intensity of the resulting field in Eq. (2.4) at the detector can be given as follows:

$$\begin{aligned} I_{Tot} &= \langle (E_1 + E_2).(E_1 + E_2)^* \rangle = |E_1|^2 + |E_2|^2 + 2\mathrm{Re}\left(E_1 E_2^*\right) \\ &= |E_1|^2 + |E_2|^2 + 2|E_1||E_2| cos(\phi_1 - \phi_2) \end{aligned} \tag{2.5}$$

Hence, the resulting intensity due to interference of two waves is simply not the sum of the individual intensities but is rather given by:

$$I_{Tot} = I_1 + I_2 + 2\sqrt{I_1 I_2} cos(\phi_1 - \phi_2). \tag{2.6}$$

It should be noted that the effective phase difference ($\phi_1 - \phi_2$, in this case) plays an important role and determines if it is a constructive ($\phi_1 - \phi_2 = 2n\pi$; $n = 0, \pm 1, \pm 2 \ldots$) or destructive interference ($\phi_1 - \phi_2 = (2n-1)\pi$; $n = 0, \pm 1, \pm 2 \ldots$). If two identical optical waves (assuming zero phase offsets, i.e. $\phi_1 = \phi_2$) undergo different path lengths before interference occurs, then it will result in phase difference, $\Delta\phi = k(z_1 - z_2) = k\Delta z$, where Δz is the path difference between two waves. If the phase difference changes continuously, it will result in interference fringes and the measured intensity will vary from a maximum of $I_{\max}\left(= I_1 + I_2 + 2\sqrt{I_1 I_2}\right)$ to a minimum of $I_{\min}\left(= I_1 + I_2 - 2\sqrt{I_1 I_2}\right)$. A useful parameter of interest is the fringe visibility contrast or fringe modulation depth and is defined as: $\dfrac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \dfrac{2\sqrt{I_1 I_2}}{I_1 + I_2}$.

2.2.1.2 Interference of Planar Waves of Different Frequency

In the last sub-section, we discussed the interference between two waves with same frequency but different optical path length or phase difference. Another important interference phenomenon is the resulting electric field due to interference of plane waves of the different frequency. This mathematical analysis is especially useful because in reality light waves are not strictly monochromatic but have a finite spectral bandwidth. Having a finite bandwidth allows the ability to modulate the field amplitude. Availability of high speed optical modulators (electro-optic and acousto-optic) make it possible realize many useful technologies that have had far reaching impact in fiber optic industry. Mathematically the sum of two waves of equal amplitude but with different angular frequency can be represented as following (assuming zero phase offsets for simplicity, i.e. $\phi_1 = \phi_2 = 0$):

$$\begin{aligned} E_{Tot} &= E_0 \exp(i(k_1 z - \omega_1 t)) + E_0 \exp(i(k_2 z - \omega_2 t)) \\ &= E_0 \exp(i(k_1 z - \omega_1 t)) + E_0 \exp(i((k_1 + \Delta k)z - (\omega_1 + \Delta\omega)t)), \end{aligned} \tag{2.7}$$

where the angular frequency and the propagation wavevector of the two waves differ by $\Delta\omega$ and Δk, respectively. If we simplify for the real part (electric field value) of the Eq. (2.7) above then we get:

$$\begin{aligned}\mathrm{Re}(E_{Tot}) &= 2E_0\cos\left(\left(k_1+\frac{\Delta k}{2}\right)z-\left(\omega_1+\frac{\Delta\omega}{2}\right)t\right)\cos\left(-\frac{\Delta k}{2}z+\frac{\Delta\omega}{2}t\right)\\ &= E_{Mod}(t)\cos\left(\left(k_1+\frac{\Delta k}{2}\right)z-\left(\omega_1+\frac{\Delta\omega}{2}\right)t\right),\end{aligned} \tag{2.8}$$

where the resulting wavefield has a modulated amplitude $E_{Mod}(t) = 2E_0\cos\left(-\frac{\Delta k}{2}z+\frac{\Delta\omega}{2}t\right)$.

The resulting intensity measured by detector has modulations with a beat frequency, $\Delta\omega$, as shown in the equation below:

$$\begin{aligned}I_{Tot} &= \langle(E_1+E_2).(E_1+E_2)^*\rangle = |E_1|^2+|E_2|^2+2\mathrm{Re}\left(E_1E_2^*\right)\\ &= 2I_0(1+\cos(-\Delta kz+\Delta\omega t)) = 4I_0\cos\left(\frac{-\Delta kz+\Delta\omega t}{2}\right)^2.\end{aligned} \tag{2.9}$$

This result has a huge significance as beat frequencies can be orders of multitude smaller than the optical frequency ($\sim 10^{14}$ Hz) which cannot be detected using optoelectronic detectors. Instead one can rely on measurement of beat frequency (typically kHz–GHz range) for highly sensitive phase measurements. Figure 2.1 shows linear sum of two monochromatic waves of different frequency and different amplitude. It should be noted that the resulting modulation depth is not the maximum because the two waves have different amplitude.

Note: It must be noted that interference or linear superposition of the scalar electric field applies only if the two waves have same polarization. Otherwise, the addition of the two differently polarized light waves would result in an output field of new polarization. Another important point to note is that both temporal and spatial coherence are required for two fields to manifest interference effects. Spatial coherence is a measure of correlation between electric field or wavefunctions at two different positions in space. Temporal coherence is a measure of

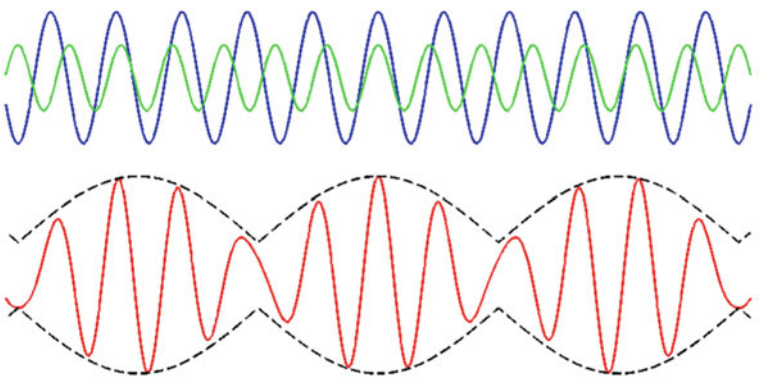

Fig. 2.1 Superposition of two monochromatic waves of different frequency and amplitude (in *blue* and *green*). The resulting electric field is given by the *red* graph below, with amplitude modulation showed with *dashed black lines*

correlation or predictable phase relationship between fields or wavefunctions observed at different points in time. As per the mathematical definition, monochromatic plane waves that we have used so far in our analysis have infinite coherence time and coherence lengths.

2.2.1.3 Quasi-monochromatic, Polychromatic, and Broadband Light

Monochromatic plane wave representation is a great tool to understand some of the basic fundamentals of wave theory of light [2, 3]. However, in reality, there is no such thing as a monochromatic plane wave because a truly monochromatic plane wave has infinite energy and infinite time duration (time from $-\infty$ to $+\infty$). In fact, all the travelling light waves that carry energy have finite bandwidths and hence for simplicity, we would call it polychromatic light. Polychromatic waves can be expressed by using weighted sum of monochromatic waves of different frequencies using the famous principle of superposition. The use of Fourier methods greatly simplifies the representation of broadband light. For example, an arbitrary wavefunction, $E(t)$, can be defined for a given location (say at $z = 0$) using the superposition integral of monochromatic waves of different frequencies, amplitude and phases as shown in the equation below:

$$E(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(\omega) \exp(-i\omega t) d\omega. \tag{2.10}$$

Here $E(\omega)$ can be obtained by carrying out the Fourier transform as shown below:

$$\tilde{E}(\omega) = \int_{-\infty}^{\infty} E(t) \exp(i\omega t) dt = F.T.\{E(t)\}. \tag{2.11}$$

The spectrum of the wavefunction, $E(t)$, is defined as the absolute value of the square of the Fourier transform of the wavefunction: $S(\omega) = |F.T.\{E(t)\}|^2 = |E(\omega)|^2$.

2.2.2 Phase and Group Velocity

The concept of group velocity and optical path delay is inherently central to any practical interferometric setup. So far we have described monochromatic waves that travel in vacuum with the phase velocity, $v_{Ph} = \frac{\omega}{k} = c$. When the waves travel in a dispersive medium with refractive index n, the phase velocity is given as: $v_{Ph} = \frac{c}{n}$. We have already discussed that monochromatic waves are ideal

representations and hence do not carry any energy. For a monochromatic wave, this is often called phase velocity, and in vacuum, EM waves of different frequency travel with the same phase velocity, i.e. of light. In reality, however, energy is carried by wave packets or wave pulses and any media of propagation has dispersion (see Chap. 1). Unlike, monochromatic waves that have indefinite extent in time, a wave packet consist of a localized wavefunction (spatial or temporal) that travels with a group velocity $\left(v_g = \frac{\partial\omega}{\partial k}\right)$ in a dispersive medium. Figure 2.2 shows the resulting wave packet as a result of superposition of multiple monochromatic waves of variable frequency and amplitudes. In reality, a light pulse of localized energy is made up of continuum of frequencies (that essentially can be explained by the sum of Fourier frequencies of the light pulse). In the Fig. 2.2, we can see that even the discrete number of wavelengths or frequencies can add up to form a modulated electric field envelope or wave packet.

In Sect. 2.2.1.2, we saw an interesting result that when two monochromatic waves of different frequency superimpose, it results in modulated electric field amplitude. One of the key insights from this result is that superposition of many monochromatic light waves with varying frequencies around a central frequency can result in a localized light wave packet or a light pulse. In Eq. (2.8), the simplified expression of the combined waves depicted a monochromatic wave with average wavevector and frequency $\left(\bar{k} = k + \frac{\Delta k}{2};\ \&\ \bar{\omega} = \omega + \frac{\Delta\omega}{2}\right)$, but with amplitude modulation. If we look at the modulated envelope, we realize that it has a phase of its own and it travels with a group velocity of:

$$v_g = \frac{\Delta\omega}{\Delta k} = \frac{\partial\omega}{\partial k} \tag{2.12}$$

Hence for propagation of light in any optical media, the group refractive index (n_g) is defined as: $n_g = \frac{c}{v_g}$. It must be noted that group refractive index should not be confused with the refractive index (n), which is defined as: $n = \frac{c}{v_{Ph}}$.

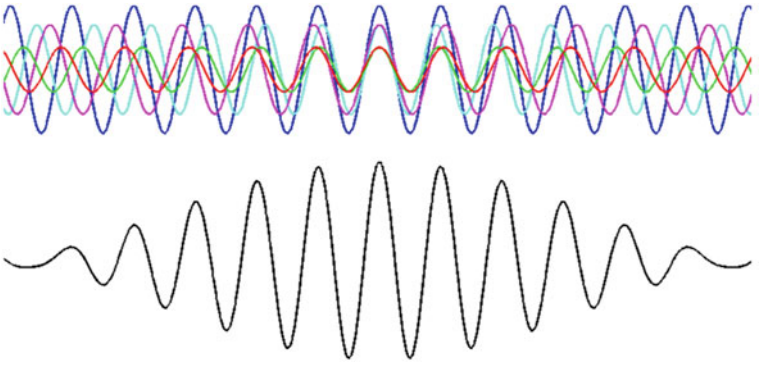

Fig. 2.2 Superposition of multiple monochromatic waves of varying frequency and amplitude results in formation of a wave packet below (*black*)

When light travels through a fiber or any other optical media, the phase term of the associated electric field is modified as follows: $E_{out} = E_{in}\exp(ik_0nz)$, where n is the refractive index, k_0 is the wavenumber in vacuum, and z is the distance travelled in the media. One of the often used term in interferometry is the optical path length (OPL), which is described as the product of the physical distance travelled by an electric field and the refractive index of the medium ($OPL = nz$). A related and equally useful term is called optical path difference (OPD), which is described by the differential OPL between two light waves of the same origin, but that travelled through different media or different distances. OPD is often calculated in two beam interferometry and is given by: $OPD = n_1z_1 - n_2z_2$. For many applications in interferometry, it is often convenient to use the group refractive index to calculate the OPD or OPL.

2.2.3 Fiber Optic Interferometric Sensors

In a typical fiber optic interferometric sensor, the light is divided in at least two parts and at least one part of the light interacts with the measurand (a quantity or physical effect that is intended to be measured). The interaction of the measurand with the light field would result in a phase shift or phase modulation, which can be detected when the modified light field interferes with the reference light (Fig. 2.3). Fiber optic interferometric sensors typically offer high sensitivities due to several reasons including low propagation loss in fiber, and interferometric detection. As shown in Fig. 2.3, the light can be modified by multitudes of environmental perturbations or measurable quantities such as temperature, strain, heat, humidity, force, pressure, flow, acoustic, vibrations, acceleration, velocity, electric or

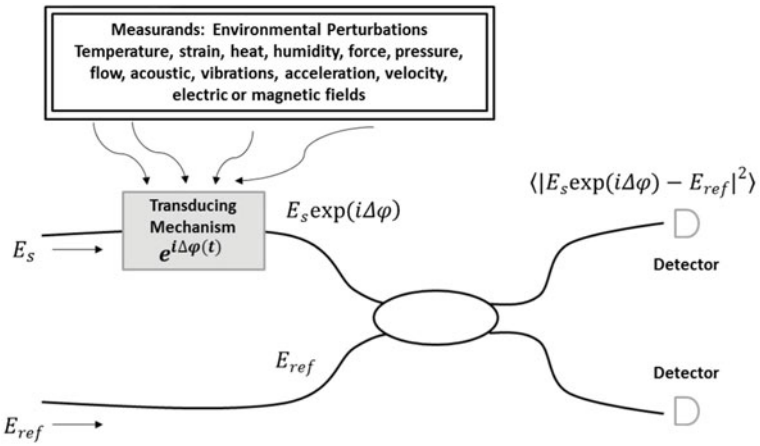

Fig. 2.3 Schematic to show a generic interferometric fiber optic sensor. Environmental perturbations such as temperature, strain, pressure etc. interact with the light field and impart a phase shift or modulation that can be analyzed by processing the interference term at the detector

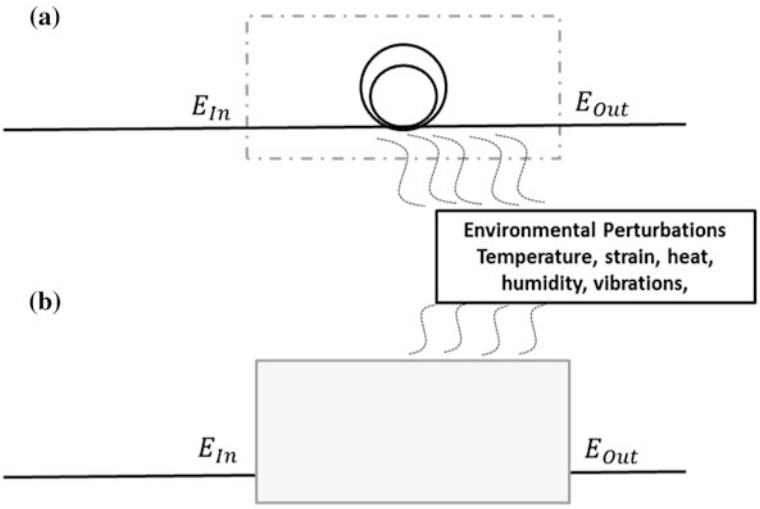

Fig. 2.4 Schematic to show a generic intrinsic and extrinsic fiber optic sensors. **a** intrinsic fibre sensors. **b** extrinsic fiber sensors

magnetic fields etc. These environmental perturbations interact with the electric field through a transducing mechanism and impart a change or modulation in the phase term: $\Delta OPL = \Delta(k_0 nz)$.

Broadly speaking, fiber optic sensors can be divided into two categories (Fig. 2.4):

- Intrinsic fiber sensors: The light is modified inside the fiber and the fiber itself acts as a transducer, in part or as a whole. Often the fiber is attached to a material that acts as a transducer in tandem with fiber. The intrinsic fiber sensors have the advantage of compact and efficient design with high sensitivity.
- Extrinsic or hybrid fiber sensors: The light is carried by a fiber to a location where the light is modified by environmental perturbations or measurand, and the modified light is then collected back by the same or another fiber and directed to a detector where it is processed and analyzed. One of the major advantages of the extrinsic fiber sensors is that the fiber acts as a flexible and rugged dielectric conduit of light and enables delivery and collection of light for measurement purposes, which otherwise would had been prohibitive due to harsh environmental conditions.

In the intrinsic fiber sensors, these environmental perturbations impart a change in the physical property of the fiber itself; such as temperature or strain induced changes in length or refractive index which in turn imparts a change in the optical phase of the light travelling through the fiber. On the other hand, extrinsic fiber sensors could have much wider range of applications because it is then feasible to expand the transducing mechanisms as the light is not necessarily confined within the fiber when interacting with the perturbation field.

2.3 Fiber Optic Interferometer Architectures

In this section, we will briefly discuss some of the common fiber optic interferometer architectures topologies and draw parallels with their bulk optic counterpart interferometer architectures. The four interferometer architectures are Michelson, Mach–Zehnder, Sagnac and Fabry–Perot interferometers. The goal for this section is to provide the reader with an overview of different architectures and a more detailed treatment for most of these designs is either provided in the later sections (Michelson and Mach–Zehnder interferometers in Sect. 2.4 of this chapter) or later in this book (see the chapter on Sagnac interferometer later in the book). In principle, any of the interferometric architecture (with the exception of Sagnac interferometer) can be alternatively realized to measure any of the physical properties. However, different architecture may present different design, cost and performance trade-offs and one or another architecture may be preferred depending upon the unique requirements of a specific application.

2.3.1 Michelson Interferometer

Michelson interferometer is probably the most commonly known optical configuration that is used for interferometry. One of the most significant applications of this interferometer was in the famous Michelson-Morley experiment that was carried out in 1887 to detect the relative motion of earth and aether in the universe. While the unexpected results of the experiments confounded scientists at the time, it was one of the key findings that inspired Einstein's special theory of relativity.

Figure 2.5 shows the typical schematics of bulk-optic and fiber optic realizations of Michelson interferometers. In a Michelson interferometer, an incident light field

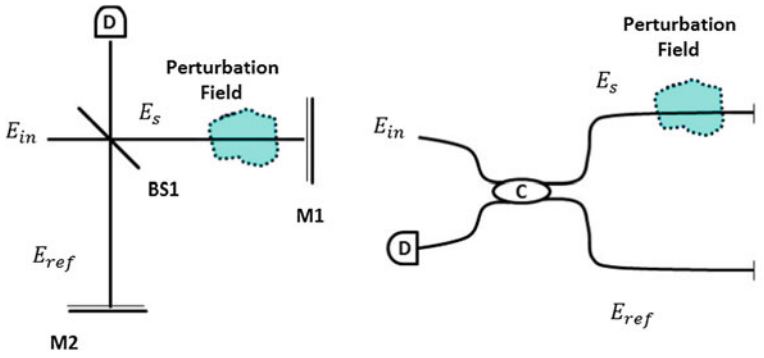

Fig. 2.5 Schematic design for bulk-optic (*left*) and fiber optic Michelson interferometer architecture. The electric fields, optical and fiber optic components shown in the figure are as follows: E_{in}: incident electric field; E_s: electric field in the sample arm; E_{ref}: electric field in the reference arm; *BS1*: beam splitter; *M1* and *M2*: mirrors; *C*: fiber optic coupler; and *D*: detector

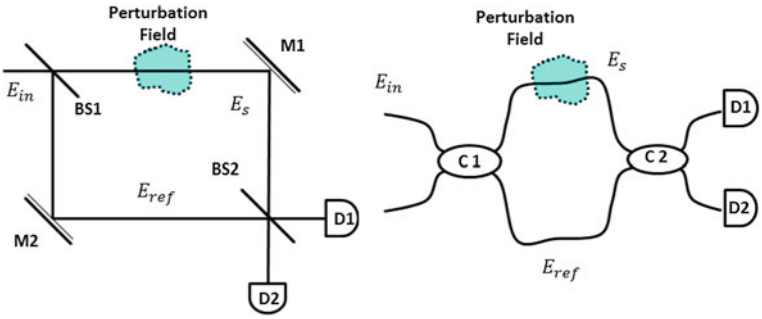

Fig. 2.6 Schematic design for a bulk-optic (*left*) and fiber optic Mach–Zehnder interferometer architecture

is divided into two parts (sample and reference) by a beam splitter or fiber optic coupler. The light in the reference and sample paths are reflected back and combined using the same splitter and the interference patterns of the two lights are measured using a detector. The light in the sample arm accrues perturbation field dependent phase shifts that can be measured using interference with the reference light.

2.3.2 Mach–Zehnder Interferometer

In a Mach–Zehnder interferometer, the incident light field is divided into two parts (sample and reference) by a beam splitter or fiber optic coupler. Unlike Michelson, the light in the reference and sample paths is not reflected; rather it is directed to a second splitter/combiner where the two lights combine and the interference patterns are measured using detectors (Fig. 2.6). In fact, if a Mach–Zehnder interferometer is cut in half and folded back, the topology will become similar to that of Michelson interferometer.

2.3.3 Fabry–Perot Interferometer

A Fabry–Perot (FP) interferometer of etalon consists of two reflective surfaces that are often parallel. The incident light is reflected back and forth and transmitted multiple times at the two partially reflective surfaces. The superposition of these multiply reflected and transmitted beams at the two surfaces results in interference effects that determine the transmission and back-reflected light characteristics of the light. Figure 2.7 shows various bulk-optic and fiber optic realizations for FP interferometers. The two examples corresponding to extrinsic and intrinsic fiber optic FP interferometers are also shown. In the intrinsic fiber optic FP sensor the reflective surfaces inside the fiber can be created by micro-machining, fiber Bragg gratings (FBG), or thin film deposition.

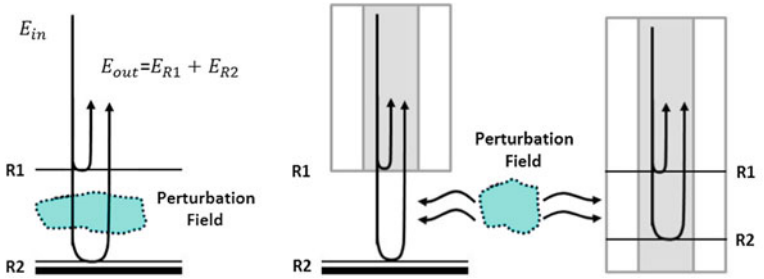

Fig. 2.7 Schematic design for a bulk-optic (*left*), extrinsic fiber optic (*middle*), and intrinsic fiber optic (*right*) Fabry–Perot interferometer architecture. *R1* and *R2* are the two reflective surfaces forming the Fabry–Perot cavity

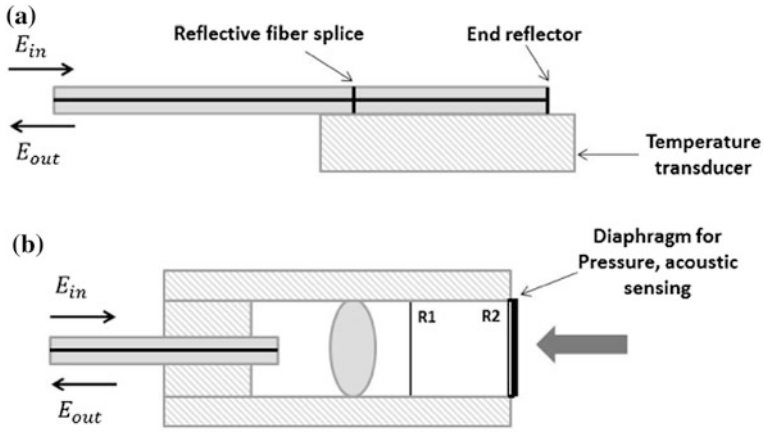

Fig. 2.8 **a**. Intrinsic FP interferometer based sensor for thermal measurements. **b**. Extrinsic FP interferometer based sensor for pressure acoustic sensing

In Sect. 2.2.3, we discussed that environmental perturbations could modify the medium of propagation of light and result in phase shifts that is proportional to effective change in *OPL* as shown here: $\Delta\phi = \Delta(nk_0z) = k_0z\left(\Delta n + \frac{n\Delta z}{z}\right)$. Several environmental perturbations such as temperature, pressure, strain and acoustic waves could impart these phase shifts via interaction with fiber or external medium. For example, Fig. 2.8a shows an intrinsic fiber optic temperature sensor based on FP interferometer. Here the composite material attached to the fiber introduces temperature dependent strain to the fiber, thereby resulting in phase shifts. Figure 2.8b shows another FP interferometer based implementation that can be used to measure acoustic waves or pressure changes. The diaphragm with a reflective surface on one side acts as a transducer which results in displacement with respect to pressure or acoustic changes.

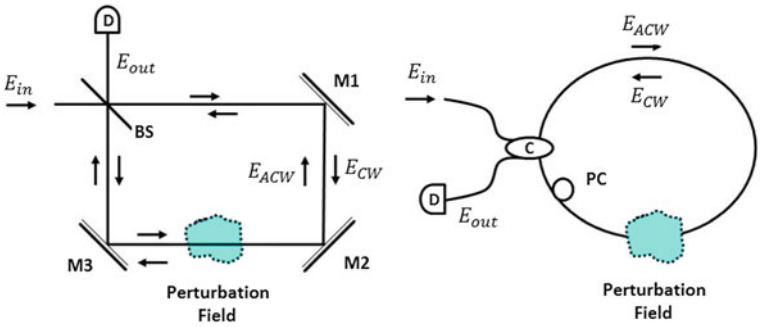

Fig. 2.9 Schematic design for a bulk-optic (*left*), and intrinsic fiber optic (*right*) Sagnac interferometer architecture. E_{CW} and E_{ACW} are the electric field traveling in the clockwise and anticlockwise direction in the Sagnac ring or loop

2.3.4 Sagnac Interferometer

In Sagnac or ring interferometer, the incident light field is split into two parts. While the path travelled by the two beams is the same, the two beams travel in opposite angular directions (i.e. clockwise and anti-clockwise). After completing the loop trajectory, the two beams combine at the point of entry and undergo interference (Fig. 2.9). Because the two beams move in opposite angular directions, the interference signal at the coupler is highly sensitive to the angular motion of Sagnac loop itself. Sagnac interferometers have found applications in fiber optic gyroscopes and a more detailed discussion on theory and applications will be provided in the later chapter.

2.4 Basic Operation Principles of Fiber Optic Components and Assemblies

It is important to understand the basic operation principles of some of the selected fiber optic components as majority of fiber optic sensors and interferometric setups can be realized using these components. This section focuses on providing the basic mathematical tools to model these components for interferometry applications.

2.4.1 Ideal 2 × 2 Fiber Optic Couplers

A very basic fiber optic component is a 2 × 2 fiber optic coupler that has two input ports and two output ports, each port being a single-mode fiber that supports only the LP01 mode of light at the wavelength of interest (Fig. 2.10). The "single-mode"

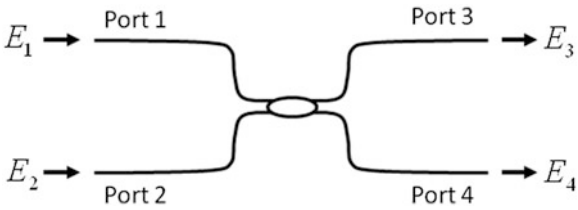

Fig. 2.10 Coupling of light through a 2 × 2 fiber optic coupler

fiber actually supports two degenerate modes of orthogonal polarizations. But for simplicity the polarization effects are ignored for the moment. The light signals entering the two input ports are mixed in the coupler and then redistributed at the two output ports as illustrated in figure 2.10.

To describe the physics of light wave coupling we start from a simple mathematical model regardless of the light coupling mechanism. Following the convention introduced in Chap. 1, we describe the time dependence of the light signals at the 4 ports as

$$E_n(t) = \frac{1}{2} E_n \exp(-i\omega t) + \text{c.c.} \tag{2.13}$$

where E_n ($n = 1,\ 2,\ 3,$ and 4) are the complex amplitudes of the electric fields, and c.c. stands for the complex conjugate.

For an ideal loss-less 2 × 2 coupler, the output is a unitary transformation of the input, which can be conveniently expressed in the form a unitary matrix U

$$\begin{bmatrix} E_3 \\ E_4 \end{bmatrix} = U \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}. \tag{2.14}$$

Here, a and b are complex values (a^* and b^* are their complex conjugate), which satisfy the following normalization constraint

$$|a|^2 + |b|^2 = 1. \tag{2.15}$$

This unitary transformation guarantees that the total optical "power" is conserved (loss-less)

$$|E_3|^2 + |E_4|^2 = |E_1|^2 + |E_2|^2. \tag{2.16}$$

The values of $|a|^2$ and $|b|^2$, usually expressed in percentages, are the coupler's power splitting ratios for the parallel and cross paths. For example, a 50/50 coupler and an 80/20 coupler can be described with $U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ and $U = \begin{bmatrix} \sqrt{0.8} & \sqrt{0.2} \\ -\sqrt{0.2} & \sqrt{0.8} \end{bmatrix}$, respectively.

People familiar with electron spin in quantum mechanics may recognize that matrix U has the same properties as a spin-1/2 rotation operator. It also resembles the Jones matrix for a loss-less polarization component (to be discussed in Sect. 2.4.5). In fact the underlying mathematics is very much the same for these distinctly different physical subjects (It would be an insightful exercise to expand matrix U in terms of the elegant Pauli matrices. But this is beyond the scope of this book and we leave it to the interested readers).

It should be noted that matrix U is not unique, and the same coupler may be described with other equivalent forms of U. For example, a 50/50 coupler could as well be represented by $U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ or $U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$. Depending on the exact locations where the input and output amplitudes E_n ($n = 1,\ 2\ ,\ 3,$ and 4) are defined, the matrix U could absorb some additional phase factors. For instance, if E_4 for the 4th port is re-defined at a different location along the fiber such that E_4 has a phase shift of π or $E_4' = E_4 \exp(i\pi) = -E_4$, Eq. (2.14) can then be rewritten as

$$\begin{bmatrix} E_3 \\ E_4' \end{bmatrix} = \begin{bmatrix} E_3 \\ -E_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} a & b \\ b^* & -a^* \end{bmatrix}\begin{bmatrix} E_1 \\ E_2 \end{bmatrix}. \quad (2.17)$$

Note the transformation matrix has changed after absorbing a phase factor matrix. A global common phase factor has no significance in the analysis, but relative phases are critical for the interference outcome. It is very important to keep track of the phase factors in a consistent manner.

For non-monochromatic light waves, the ω-dependence of U must be considered. In general, all matrix elements of U vary with ω, not only the amplitude but also the phase. Variation of amplitudes $|a|$ and $|b|$ means the coupler splitting ratios are wavelength dependent. Variation of the phases, on the other hand, indicates time delay and group velocity dispersion.

2.4.2 Mach–Zehnder Interferometer

The matrix representation is very convenient for the analysis of a fiber optic interferometric system. Figure 2.11 shows a simple layout of a Mach–Zehnder interferometer containing two 50/50 couplers and a phase shift element in between

As discussed earlier, the two 50/50 couplers can be described with

$$U_1 = U_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (2.18)$$

The phase shifter in the middle section is represented by a diagonal matrix:

$$U_2 = \begin{bmatrix} \exp(i\phi) & 0 \\ 0 & 1 \end{bmatrix}. \quad (2.19)$$

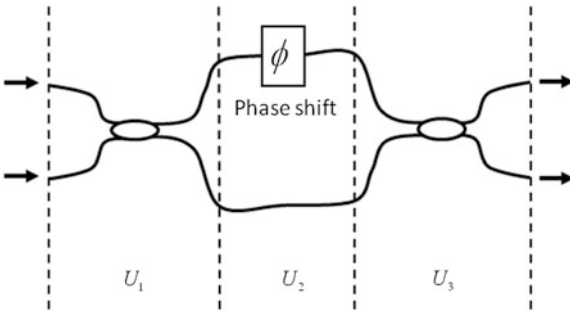

Fig. 2.11 Mach-Zehnder interferometer comprising 3 segments described by unitary matrices U_1, U_2, and U_3

For convenience we multiply a common phase factor $\exp(-i\phi/2)$ and express U_2 in a more symmetric form (only the relative phase matters)

$$U_2 = \begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix}. \tag{2.20}$$

The overall transformation matrix for the entire system can then be computed by concatenating the three matrices

$$\begin{aligned} U_{\text{total}} = U_3U_2U_1 &= \left(\frac{1}{\sqrt{2}}\right)^2 \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} e^{i\phi/2} + e^{-i\phi/2} & e^{i\phi/2} - e^{-i\phi/2} \\ e^{i\phi/2} - e^{-i\phi/2} & e^{i\phi/2} + e^{-i\phi/2} \end{bmatrix} = \begin{bmatrix} \cos(\phi/2) & i\sin(\phi/2) \\ i\sin(\phi/2) & \cos(\phi/2) \end{bmatrix} \end{aligned} \tag{2.21}$$

This result indicates that the system behaves like an optical switch. When $\phi = 0$ the resulting matrix $U_{\text{total}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ represents a 100 % parallel-state, and when $\phi = \pi$ the resulting matrix $U_{\text{total}} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$ represents a 100 % cross-state.

2.4.3 Michelson Interferometer

In Sect. 2.3.1 we have also introduced an important interferometer topology, the Michelson interferometer, which involves reflecting the two output signals of the 2×2 coupler back into the same coupler to achieve optical interference (Fig. 2.5). To model such an interferometer, it is necessary to know the transformation matrix of the same coupler for light waves propagating in the opposite

direction (input at ports 3 and 4, and output at ports 1 and 2 in Fig. 2.10). In the following, we show that the matrix for the reverse propagating light is U^T, the transpose of U. In some sense this is a trivial result: it simply means the matrix element that links a specific input port to a specific output port is reciprocal and remains the same regardless of light transmission direction.

The derivation is based on time reversal symmetry, which the ideal loss-less 2×2 fiber optic coupler obeys. (Not all optical components obey time reversal symmetry, as will be discussed later in Sect. 2.4.6). By definition the unitary matrix U has the property $U^{-1} = U^+$ ("U dagger", conjugate transpose of U), Eq. (2.14) can be easily reversed

$$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = U^+ \begin{bmatrix} E_3 \\ E_4 \end{bmatrix} = \begin{bmatrix} a^* & -b \\ b^* & a \end{bmatrix} \begin{bmatrix} E_3 \\ E_4 \end{bmatrix}. \tag{2.22}$$

But U^+ is not the matrix for the reverse propagating light waves. There is one extra step. Time reversal means replacing t with $-t$, and Eq. (2.13) becomes

$$\begin{aligned} E_n(-t) &= \frac{1}{2} E_n \exp[-i\omega(-t)] + \text{c.c.} \\ &= \frac{1}{2} E_n^* \exp(-i\omega t) + \text{c.c.} \end{aligned} \tag{2.23}$$

This means under time reversal, the complex amplitudes E_n ($n = 1,\ 2\ ,\ 3,$ and 4) all need to be replaced with their complex conjugate. Applying complex conjugate to both sides of Eq. (2.22), we have

$$\begin{bmatrix} E_1^* \\ E_2^* \end{bmatrix} = U^T \begin{bmatrix} E_3^* \\ E_4^* \end{bmatrix} = \begin{bmatrix} a & -b^* \\ b & a^* \end{bmatrix} \begin{bmatrix} E_3^* \\ E_4^* \end{bmatrix}. \tag{2.24}$$

This subtle difference between U^+ and U^T should not be overlooked. For example, let us consider a Michelson interferometer with perfectly reflective mirrors at ports 3 and 4. What is the overall transformation U_{total} after the light waves are reflected back into the coupler, interfere, and reemerge at ports 1 and 2. Knowing the transformation in the reverse direction is U^T, we can perform a straightforward calculation (ignoring a common phase factor introduced by the mirrors):

$$U_{\text{total}} = U^T U = \begin{bmatrix} a & -b^* \\ b & a^* \end{bmatrix} \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix} = \begin{bmatrix} a^2 + b^{*2} & ab - a^* b^* \\ ab - a^* b^* & a^{*2} + b^2 \end{bmatrix}. \tag{2.25}$$

If U^+ were used instead of U^T for the counter-propagating waves, the answer would be a unit matrix regardless of the nature of the coupler, which is obviously incorrect.

2.4.4 Other Fiber Optic Couplers

The loss-less 2 × 2 coupler discussed so far is only an ideal physical model. In reality, couplers all have finite optical losses. More realistic 2 × 2 couplers can be modeled using an ideal 2 × 2 coupler by adding attenuations to individual ports. Once the attenuation is included, the transformation matrix U is no longer unitary and some of the symmetry properties of an ideal 2 × 2 coupler no longer exist. For a realistic 2 × 2 coupler in general, the coupling ratios for the two through paths (from port 1 to port 3 and from port 2 to port 4) can be different. So are the coupling ratios for the two cross paths (from port 1 to port 4 and from port 2 to port 3). Nevertheless, one can often still use Eq. (2.14) as an approximation for a lossy 2 × 2 coupler by lowering the value of $|a|^2 + |b|^2$ to below 1. The value of $-10\log\left(|a|^2 + |b|^2\right)$ is called the "excess loss" (in dB) of the coupler. The excess loss of a good quality fiber optic coupler is typically less than 0.1 dB.

There are other kinds of fiber optic couplers such as a 1 × 2 Y-coupler or a more sophisticated 3 × 3 coupler. A 1 × 2 Y-coupler can be simply modeled with a 2 × 2 coupler having one of the 4 ports not used (terminated with an index-matching absorber). An ideal 3 × 3 coupler, on the other hand, could be modeled theoretically with a 3 × 3 unitary matrix. A 3 × 3 coupler is more difficult to fabricate and it is not nearly as common as the 2 × 2 coupler in fiber optic interferometry.

2.4.5 Fiber Birefringence and Polarization Controllers

The polarization effect has been ignored in the theoretical model presented above. As we mentioned earlier, the "single-mode" fiber actually supports two degenerate modes of orthogonal polarizations. For a realistic interferometric system based on single mode fibers, the two polarization modes must be considered and polarization controllers are frequently used.

As we take into account the effect of polarization, a simple section of optical fiber itself becomes a 2 × 2 coupler. Fiber birefringence causes coupling between the two polarization modes in a way very similar to light wave coupling in a 2 × 2 coupler discussed in the previous section. When polarization is considered, a 2 × 2 fiber optic coupler in some sense becomes a 4 × 4 coupler and the transformation matrices will have to be expanded.

The polarization state of light in a single-mode optical fiber is very similar to that of a plane light wave in free space. It could be linear, circular, or elliptical. In the following, we shall briefly describe the mathematical tools for modeling polarization. We do not intend to go into all the details of polarization optics since they are not the focus of this book. There are many good text books on this subject, for example [2, 3].

To model polarization, we first define two orthogonal polarization states along the fiber as the basis vectors. These two polarizations do not have to be linear, although linear polarizations are often chosen for convenience. Once the basis polarization states are defined, the electric field of the light is projected to these two basis vectors, resulting in two complex amplitudes E_x and E_y (we use subscripts "x" and "y" here assuming horizontal and vertical linear polarizations are the chosen basis). The complex vector $\begin{bmatrix} E_x \\ E_y \end{bmatrix}$ is called the Jones vector.

Since a complex number has a real part and an imaginary part, a Jones vector contains total 4 independent parameters. The polarization information is actually fully described by just 2 independent parameters (a point on the 2-dimensional surface of the Poincare sphere). The other 2 independent parameters in the Jones vector are the magnitude $\sqrt{|E_x|^2 + |E_y|^2}$ and a common phase, which can be factored out if we are interested in the polarization only.

As light travels through a fiber optic system, the Jones vector is transformed by a 2×2 matrix known as the Jones matrix

$$\begin{bmatrix} E_{2x} \\ E_{2y} \end{bmatrix} = J \begin{bmatrix} E_{1x} \\ E_{1y} \end{bmatrix} = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{bmatrix} \begin{bmatrix} E_{1x} \\ E_{1y} \end{bmatrix}. \tag{2.26}$$

Note the similarity of Jones matrix J to the U matrix for the ideal 2×2 fiber optic coupler discussed in Sect. 2.4.1. In fact, for a non-absorbing (loss-less) polarization control device (a wave plate, for example), J is also a unitary matrix and it takes the same general form as described in Eq. (2.14). Regardless of the construction of the device, a unitary Jones matrix always has two orthogonal eigenvectors with two phase-only eigenvalues (norm = 1). These two eigenvectors correspond to two principal optical "axes" (generalized to include circular and elliptical polarizations), which do not change during the transformation. The two eigenvalues determine the phase retardation between the two optical axes. Therefore, any non-absorbing polarization control device can be generally considered a phase retarder.

The most commonly used fiber optic polarization controller is realized by bending the single-mode fiber into a loop (it could also be partial or multiple loops). Bending induces birefringence through anisotropic stress at the fiber core. Alternatively, one can apply anisotropic pressure on a section of fiber without bending to achieve similar effects. By symmetry, the optical axes are the two directions parallel and perpendicular to the plane of the loop. The amount of phase retardation of a fiber loop is inversely proportional to the loop diameter: the smaller the diameter, the higher the phase retardation. The Jones matrix of such a fiber loop oriented at angle θ with a phase retardation of δ is

$$J = \begin{bmatrix} \cos\frac{\delta}{2} + i\sin\frac{\delta}{2}\cos 2\theta & i\sin\frac{\delta}{2}\sin 2\theta \\ i\sin\frac{\delta}{2}\sin 2\theta & \cos\frac{\delta}{2} - i\sin\frac{\delta}{2}\cos 2\theta \end{bmatrix}. \tag{2.27}$$

For example, a fiber loop oriented at a 45° angle between horizontal and vertical ($\theta = \pi/4$) with a phase retardation $\delta = \pi/2$ (equivalent to a $\lambda/4$ waveplate) has the following Jones matrix

$$J = \begin{bmatrix} \cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\cos\frac{\pi}{2} & i\sin\frac{\pi}{4}\sin\frac{\pi}{2} \\ i\sin\frac{\pi}{4}\sin\frac{\pi}{2} & \cos\frac{\pi}{4} - i\sin\frac{\pi}{4}\cos\frac{\pi}{2} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}. \tag{2.28}$$

Consider a horizontally polarized light input $\begin{bmatrix} E_{1x} \\ E_{1y} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, after passing this $\lambda/4$ waveplate, the output becomes circularly polarized

$$\begin{bmatrix} E_{2x} \\ E_{2y} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}. \tag{2.29}$$

A manual fiber optic polarization controller usually consists of 2 or 3 fiber loops with different diameters and adjustable orientations. The total effect can be modeled by multiplying the Jones matrices of the individual fiber loops, each being a linear waveplate. Excluding a common phase factor, any non-absorbing polarization controller has only 3 independent parameters (3 degrees of freedom): 2 for defining the principal optical axes, and 1 for the retardation. In principle, a polarization controller consisting of 3 fiber loops is sufficient to produce any desired Jones matrix for polarization adjustment.

Again, similar to the ideal 2 × 2 fiber optic coupler, non-absorbing polarization devices based on mechanical stress induced birefringence obey time reversal symmetry. As a result, the Jones matrices for the forward and backward directions are the transpose of each other.

Although much of the discussion in this section is about non-absorbing polarization devices, the Jones matrix is also used to describe polarization components that do absorb light. For example, a polarizer that blocks vertically polarized light but passes horizontally polarized light can be described by a non-unitary Jones matrix

$$\begin{bmatrix} E_{2x} \\ E_{2y} \end{bmatrix} = J\begin{bmatrix} E_{1x} \\ E_{1y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} E_{1x} \\ E_{1y} \end{bmatrix}. \tag{2.30}$$

2.4.6 Fiber Optic Circulators and Isolators

The fiber optic circulator is an interesting and very useful device. A typical circulator has 3 ports and the light wave propagates in a circular fashion as shown in Fig. 2.12

Such a device is non-reciprocal. It appears to violate time reversal symmetry: the light path from port 1 to port 2 can be theoretically loss-less and yet it is unidirectional and irreversible. There are different designs for optical circulators, and all of them involve the non-reciprocal property of Faraday effect.

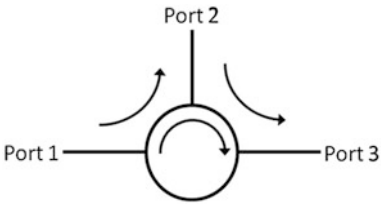

Fig. 2.12 Three-port optical circulator

The Faraday effect is a magneto-optic effect, in which the polarization of light is rotated in an optical medium with a magnetic field applied along the light propagation direction (Fig. 2.13).

Polarization rotation can also occur without magnetic field. The general effect is called optical activity, for example, in crystalline quartz that exhibits chirality. Optical activity causes the refractive indices for right-hand and left-hand circularly polarized light waves to differ. The Faraday effect can be viewed as optical activity induced by magnetic field. However, there is a fundamental difference between the Faraday effect and optical activity caused by chirality. In the absence of magnetic field, optical activity is reciprocal: when light transmission direction is reversed, so is the polarization rotation. The Faraday effect, on the other hand, is non-reciprocal: the polarization rotates in the same direction regardless of the light transmission direction. The time reversal symmetry is not really violated because the magnetic field must flip its direction in a truly time-reversed world.

This peculiar property of Faraday effect makes optical circulator possible. Figure 2.14 shows one possible way to construct an optical circulator.

It contains two polarization splitters/combiners, a half-waveplate, and a Faraday rotator. The optical axis of the half-waveplate is at 22.5° from the x-axis, and it rotates the both x and y polarizations by 45° in a reciprocal manner. The Faraday rotator, on the other hand, also produces a 45° rotation, but non-reciprocal. As a result, in one direction (Fig. 2.14, schematic on the left), when light passes both the half-waveplate and the Faraday rotator, the polarization is unchanged. But in

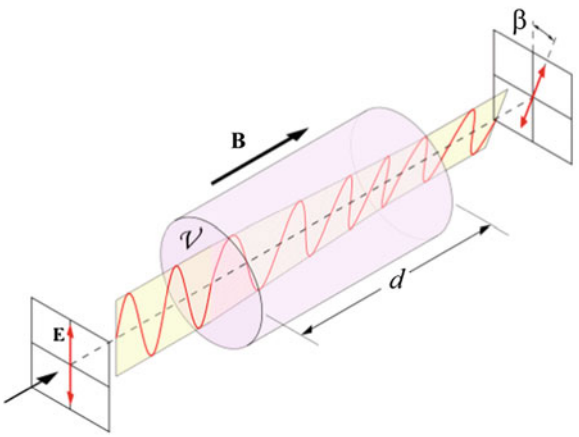

Fig. 2.13 Rotation of light polarization due to the Faraday effect [Image courtesy: Wikimedia commons image file database (drawing by: Bob Mellish)]

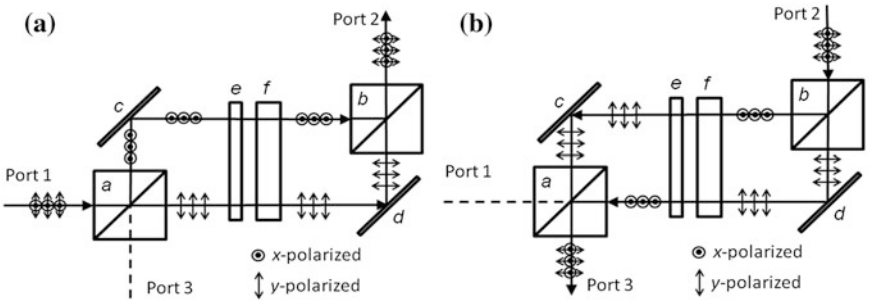

Fig. 2.14 Non-reciprocal light propagation based on Faraday rotation: (*a*) and (*b*) polarization beam splitters/combiners, (*c*) and (*d*) mirrors, (*e*) half-waveplate, and (*f*) Faraday rotator

the opposite direction (Fig. 2.14, schematic on the right), the polarization is rotated by 90°. Therefore, in the forward direction light travels from port 1 to port 2, but in the backward direction light entering port 2 would come out at port 3.

A fiber optic isolator is a simpler device with only two ports, and it is based on the same operation principle as the circulator. Light can travel in one direction but not the other. An optical isolator is often used to prevent back-reflection that tends to cause light source instability and degradation.

2.5 Conclusion

Fiber optic interferometric sensors have found several industrial applications including fiber optic gyros for navigation in airplanes and space-based systems, high-precision process control and manufacturing, structural health monitoring of bridges, dams, ships etc., oil and gas exploration activities, medical applications and so on.

While the sensitivity of the fiber optic interferometric sensors is typically very high, an important challenge lies in the capability to differentiate the environmental perturbations such as temperature, strain, pressure etc. In principle, all environmental perturbations could be converted to optical signals by applying appropriate transducing mechanisms. However, sometimes multiple effects could contribute simultaneously and modify the light in the fiber in a similar manner. For example, changes in temperature, strain, pressure or any mechanical perturbation could all impact the light in fiber by changing fiber lengths and refractive index such that it is difficult to differentiate one perturbation to another. The right solution to such challenges is the design of appropriate transducing mechanism that is selective in nature towards the measurement of a desired environmental perturbation.

Although the fiber optic sensors have already proven superior performances, their relatively higher cost compared to conventional low cost sensors remains a barrier towards its widespread adoption for wider range of applications. The semiconductor industry has continued to progress steadily over the last few

decades and as a result we have seen steady decrease in the cost of laser diodes, detectors and other optoelectronic devices. It is expected that the further decrease in optoelectronic components cost in future could be the right catalyst for the continued progress of this field.

References

1. Born M, Wolf E (1999) Principles of optics, 7th edn. Cambridge University Press, Cambridge
2. Hecht E (2002) Optics, 4th edn. Addison Wesley, San Francisco
3. Saleh BEA, Teich MC (1991) Fundamentals of photonics. Wiley Interscience, New York
4. Udd E (1991) Fiber Optic Sensors. Wiley Interscience, New York
5. Udd E, Spillman Jr WB (2011) Fiber optic sensors: an introduction for engineers and scientists, 2nd edn. Wiley, New York
6. Culshaw B (2008) Fiber-optic sensing: a historical perspective. J Lightwave Technol 26(9):1064–1078

Chapter 3
Fiber Optic Imagers

Do-Hyun Kim and Jin U. Kang

3.1 Basic Fiber Optic Imager

In most fiber optic imagers, their main differences from their bulk optic imaging counterparts are that the free-space beam path and bulk optics components are replaced by optical fibers and fiber optic components. These seem to be relatively trivial differences; however, they greatly improve the functionality and practicality of the fiber optic imaging systems compared to bulk systems for a wide range of applications—especially for endoscopic, in situ imaging. An example of a simple fiber optic imager is shown in Fig. 3.1. It consists of a light source, a fiber coupler, a light detector, an imaging fiber, and an imaging probe. Depending on the applications, the fiber coupler can be a simple power splitter (normally a 50/50 power splitter), wavelength-division-multiplexing (WDM) coupler, or other type of wavelength- or polarization-dependent coupler.

The imaging fiber can be single-mode, multi-mode, or multi-core, also known as a fiber bundle. The imaging probe typically consists of either an optical lens, graded index lens, or ball lens.

D.-H. Kim (✉)
Center for Devices and Radiological Health, US Food and Drug Administration, 10903 New Hampshire Ave, Silver Spring, MD 20993, USA
e-mail: do-hyum.kim@fda.hhs.gov

J. U. Kang
Department of Electrical and Computer Engineering, Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218, USA
e-mail: jkang@jhu.edu

J. U. Kang (ed.), *Fiber Optic Sensing and Imaging*,
DOI: 10.1007/978-1-4614-7482-1_3,

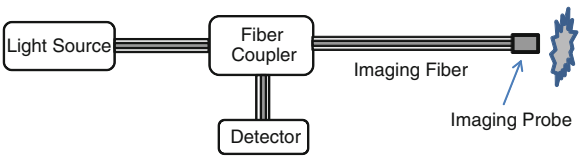

Fig. 3.1 Schematic of a simple fiber optic imager

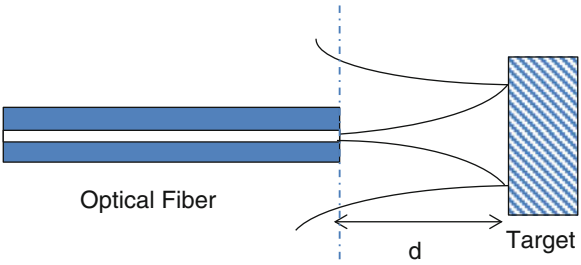

Fig. 3.2 Bare fiber probe

3.1.1 Simple Fiber Optic Imaging Probes: Single-Mode

The simplest imaging/sensing fiber optic probe is a bare fiber without any imaging components, as shown in Fig. 3.2. This is typically used when the working distance, d, is very small, less than a few hundred microns. In general, a Gaussian beam can be used to describe a beam profile exiting the optical fiber where it can be written as:

$$E(x, y, z) = E_0(z)e^{ik\frac{\rho^2}{2q(z)}} \tag{3.1}$$

where $\rho^2 = x^2 + y^2$, and the complex beam parameter, q, can be written in terms of the beam radius of curvature, R(z), and the beam width, w(z), as:

$$\frac{1}{q(z)} = \frac{1}{R(z)} - j\frac{\lambda}{\pi w^2(z)} \tag{3.2}$$

The Gaussian beam parameters can be analysed by a simple ABCD ray-transfer matrix where a new beam parameter, q_2, from the initial beam parameter, q_1, is related as:

$$q_2 = \frac{Aq_1 + B}{Cq_1 + D} \tag{3.3}$$

By combining Eqs. 3.2 and 3.3, one can derive the new beam width, w_2, as:

$$w_2^2 = \frac{B^2 + A^2\frac{\pi^2}{\lambda^2}w_0^4}{(AD - BC)(\frac{\pi^2}{\lambda^2}w_0^2)} \tag{3.4}$$

If we assume that the target is a perfect mirror, the overall ABCD matrix for beam propagation in free space for distance, d, reflected off of a target and propagating distance, d, back to the fiber can be written as:

$$\mathrm{M}_{Final} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \mathrm{M}_{Return}\mathrm{M}_{FlatMirror}\mathrm{M}_{Incident} = \begin{bmatrix} 1 & \mathrm{d} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & \mathrm{d} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2\mathrm{d} \\ 0 & 1 \end{bmatrix} \tag{3.5}$$

By combining 3.4 and 3.5, we can simplify Eq. 3.4 as:

$$w_2^2 = \frac{4d^2 + \frac{\pi^2}{\lambda^2} w_0^4}{\frac{\pi^2}{\lambda^2} w_0^2} \tag{3.6}$$

By defining the Rayleigh range or confocal parameter, z_0, as: $z_0 = \frac{\pi}{\lambda} w_0^2$ and the overall beam propagation distance, z, as z = 2d, the equation reduces to a well-known Gaussian beam width equation, which is in function of propagation distance as:

$$w_2^2 = w_0^2 \frac{z^2 + z_0^2}{z_0^2} \tag{3.7}$$

One of the main issues with the fiber probe is the coupling efficiency, T, of the returning beam back into the fiber. Due to the small aperture of the fiber, for a large z, the coupling efficiency is poor; this plays an important role in the quality of the images that can be obtained. The fiber coupling efficiency, T, assuming the system is lossless, can be calculated simply by an overlap integral between the fiber mode profile and the imaging beam profile. This can be written as:

$$T = \frac{\left| \iint F_r(x,y) W'(x,y) dxdy \right|^2}{\iint F_r(x,y) F_r'(x,y) dxdy \iint W(x,y) W'(x,y) dxdy} \tag{3.8}$$

where $F_r(x,y)$ is the imaging beam field profile at the distal end of the fiber as it returns back, and W(x,y) is the fiber mode field profile which is the beam profile of the imaging beam as it exits the fiber. From this model it is clear that the fiber does not "see" the full imaging beam as it returns back to the fiber. In general, the returning imaging field seen by the fiber, or W(x,y), is much larger than the beam diameter at the sample surface and the one exited from the fiber. Only when the target is in contact with the fiber end, the $F_r(x,y) = W(x,y)$ and T = 1 can be achieved, which describes the perfect coupling case. However, as z increases, $F_r(x,y)$ increases and this results in decreased coupling efficiency, T.

If a Gaussian beam profile is assumed, as in Eq. 3.1, the optical power collected by the fiber tip—which is the integral of the beam intensity over the fiber core area at the fiber tip plane—can be written as:

$$P_2 \propto \iint_s \frac{P_0}{w_2^2} \exp\left(-\frac{x^2+y^2}{w_2^2}\right) dxdy \propto \left[erf\left(\frac{r_0}{w_2}\right)\right]^2 \tag{3.9}$$

where P_0 is the power carried by the beam, s is the circular fiber core area with radius r_0, and $erf(u)$ is the error function defined as $erf(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt$. For estimation it is assumed that $r_0 = w_0$. There T is calculated by dividing the right-hand side of Eq. (3.9) to $\left[erf\left(\frac{r_0}{w_0}\right)\right]^2$, which results in

$$T = \left[erf\left(\frac{r_0}{w_2}\right)\right]^2 / \left[erf\left(\frac{r_0}{w_0}\right)\right]^2 = \left[erf\left(\frac{w_0}{w_2}\right)\right]^2 / \left[erf(1)\right]^2 \tag{3.10}$$

The result above shows that the working distance for the bare fiber probe is less than 0.5 mm. However, this also means that the bare fiber probe works very well, better than the ones with lensed imagers if the working distance is less than 0.1 mm, with the best image occurring when the probe is in contact with the target, i.e., z = 0 (Fig. 3.3).

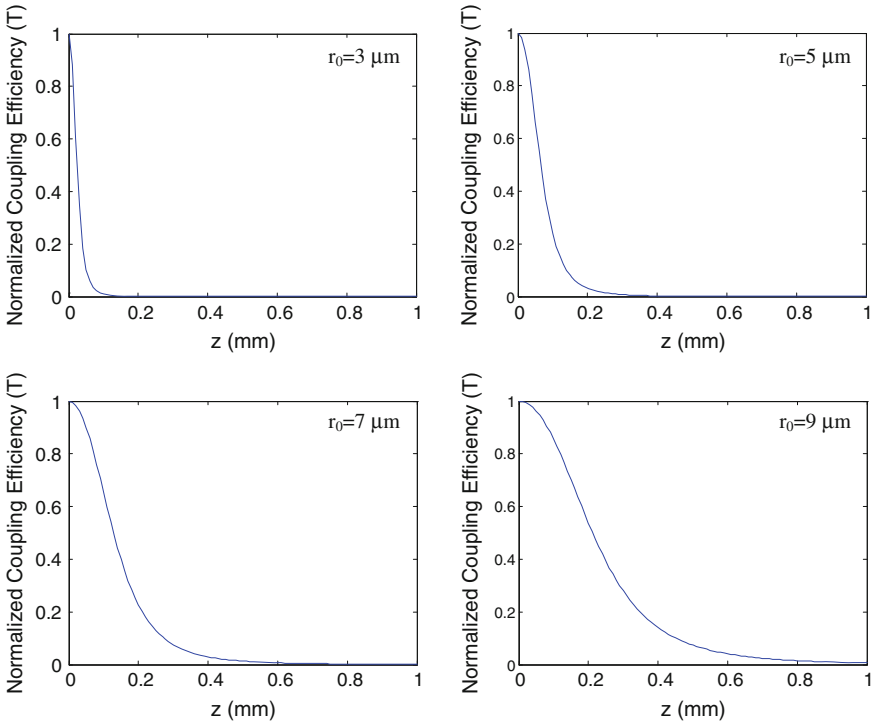

Fig. 3.3 Normalized coupling efficiency of a simple fiber probe as a function of total propagation distance, z, which is twice the target distance for 4 different fibers with varying core radius, r_0

3.1.2 Simple Fiber Probe Transverse Resolution

For all imaging systems, it is automatically assumed that the resolution of the imaging system is limited by the beam spot size. Even for a bare fiber probe, as shown in Fig. 3.2, it is often assumed that the resolution follows the Gaussian beam propagation where the beam width changes as (rewriting 3.7):

$$w_2(z) = w_0\left(1 + \frac{z^2}{z_0^2}\right)^{1/2} \tag{3.11}$$

Rayleigh range, z_0, is defined as the distance at which the beam width increases by $\sim$1.4 times the original beam width. However, the lack of imaging system makes the bare fiber probe-based imaging system non-reciprocal and the finite aperture size of the fiber affects the effective resolution of the probe. In other words, the resolution of such imaging system is limited by the imaging beam coupling back to the fiber.

3.1.2.1 Resolution Measurement

To measure the resolution of the bare fiber imaging probe, one can use a simple test setup shown in Fig. 3.4. This particular experiment was performed using an 800 nm source and a fiber having a core diameter of $\sim$5 microns. It measures the received signal power as the fiber probe is scanned across a United States Air Force (USAF) target containing high/low reflectivity boundary at some height above the target.

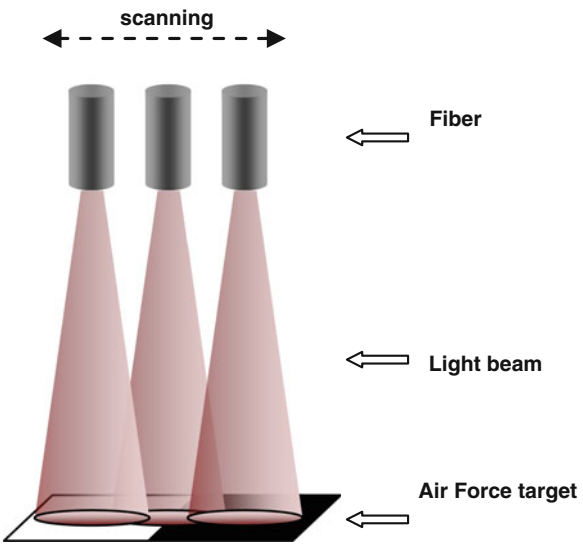

Fig. 3.4 Experimental setup to measure the transverse resolution of a fiber probe

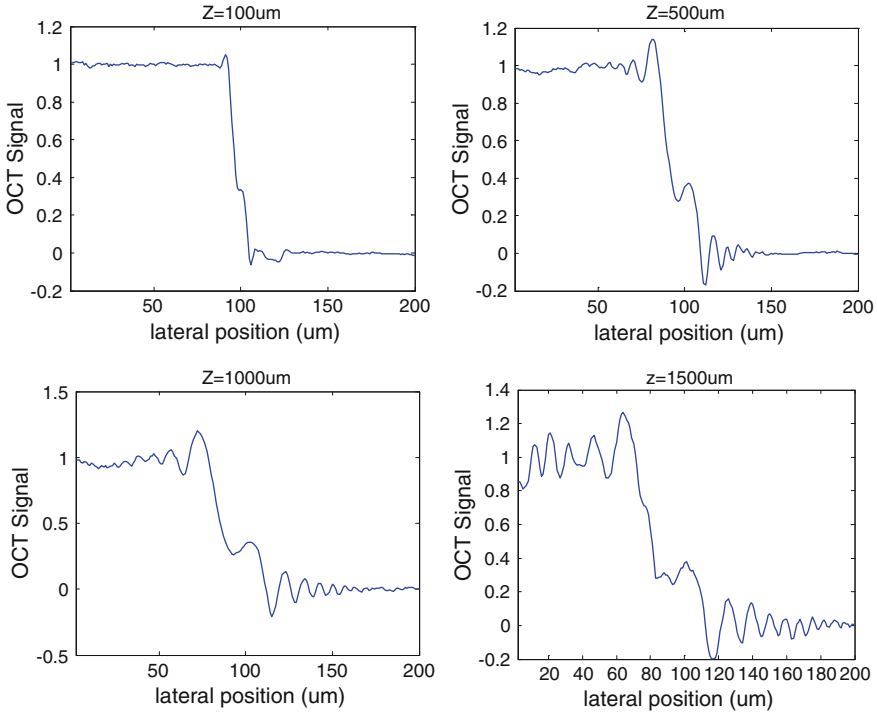

Fig. 3.5 Experimentally obtained, normalized signal for 4 different imaging distances between the fiber probe and the target, showing the transverse response function of the probe for a target having a sharp boundary at the lateral position of $\sim$100 microns

The test result is shown in Fig. 3.5 for 4 different imaging distances between the fiber probe and the target. The sharpness of the signal falling edge as the fiber is scanned from the high reflection to low indicates the resolution of the system. As can be seen from the figures, as z increases, the resolution of the bare fiber imaging system degrades. The ringing is the result of the edge diffraction.

The experimental result in Fig. 3.5 can be simulated using the signal coupling efficiency described by Eq. 3.8. The simulation result is shown in Fig. 3.6. From both the experimental and theoretical results, it can be clearly seen that the sharpness of the edge decreases significantly as the z increases. Again, the ringing around the edges is due to the edge diffraction, which will not be discussed in this book.

From these results, the system point-spread function **(PSF),** $p(x)$ of the bare fiber probe imaging system to an impulse input, $\delta(x)$, can be obtained. Since the input used in both the simulation and experiment is a step function, $u(x)$, the output of the system, $f(x)$, can be expressed as:

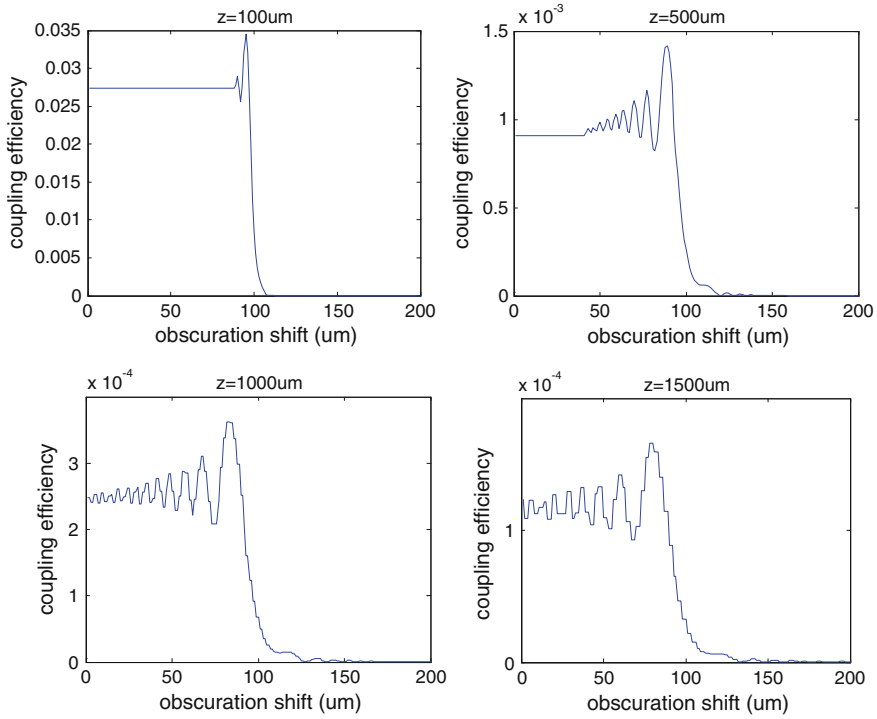

Fig. 3.6 Numerically calculated, normalized signal for 4 different imaging distances between the fiber probe and the target showing the transverse response function of the probe for a target having a sharp boundary at the lateral position of $\sim$100 microns

$$f(x) = p(x) \otimes u(x) = \int_{-\infty}^{+\infty} p(x')u(x - x')dx' \tag{3.12}$$

By taking the derivative of both sides, we can obtain an expression for the PSF as:

$$\frac{df(x)}{dx} = \int_{-\infty}^{+\infty} p(x') \frac{d[u(x - x')]}{dx} dx' = \int_{-\infty}^{+\infty} p(x')\delta(x - x')dx' = p(x) \tag{3.13}$$

Here the full-width half-maximum (FWHM) of the PSF is defined as the lateral resolution. Again we assumed the beam profile to be Gaussian, $f(x)$ and the corresponding error function is written as:

$$f(x) = A_0 \int_x^{\infty} \exp(-ax^2)\, dx = B_0 erfc(-bx) \tag{3.14}$$

The function, *erfc,* is called the complementary error function.

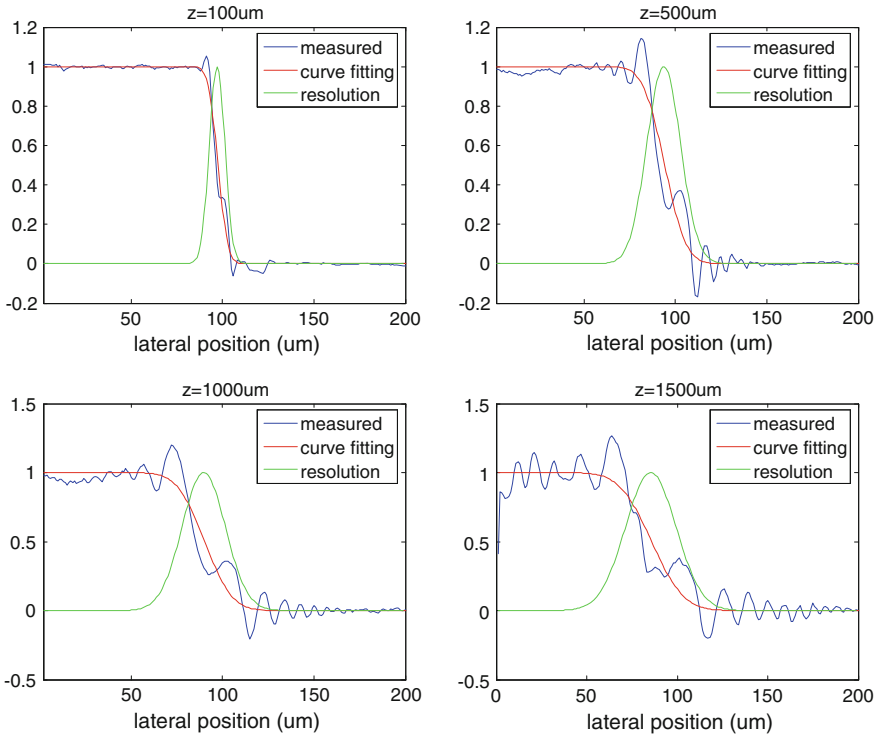

Fig. 3.7 Calculated PSF for 4 different imaging distances between the fiber probe and the target by fitting the experimental data obtained in Fig. 3.5

Using the experimental graphs in Fig. 3.5, the corresponding PSF can be extracted and Fig. 3.7 shows the PSF, or the resolution of the system for 4 different values of z. Clearly, PSF broadens with the increasing z.

Again using the theoretical graphs in Fig. 3.6, the corresponding PSF can be extracted and Fig. 3.8 shows the theoretical PSF result.

The resolution result is summarized in Fig. 3.9, where the transverse resolution obtained from both experimental and simulation results are plotted as a function of imaging distance. Note that the transverse resolution of the bare fiber probe is much better than that deduced by the spot size of the beam. For example, even when the target surface is 1.5 mm away from the end of the fiber probe, the resolution is still a respectable $\sim$33 um. In comparison, the beam size at the target is approximately 75 microns. To obtain <10 um resolution, the imaging distance has to be <100 um.

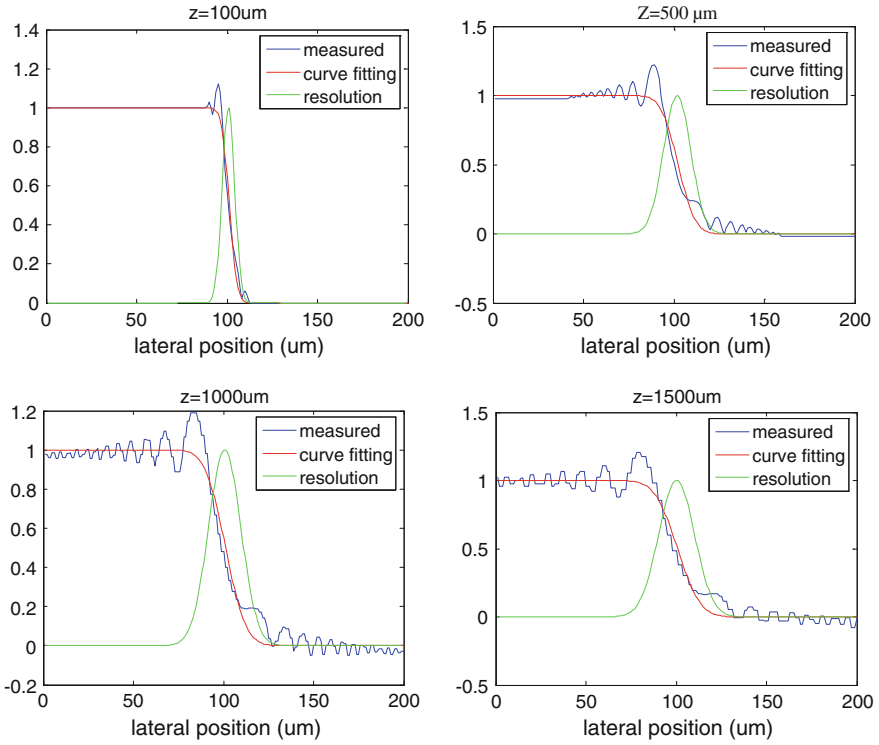

Fig. 3.8 Calculated PSF for 4 different imaging distances between the fiber probe and the target by fitting the calculated data obtained in Fig. 3.6

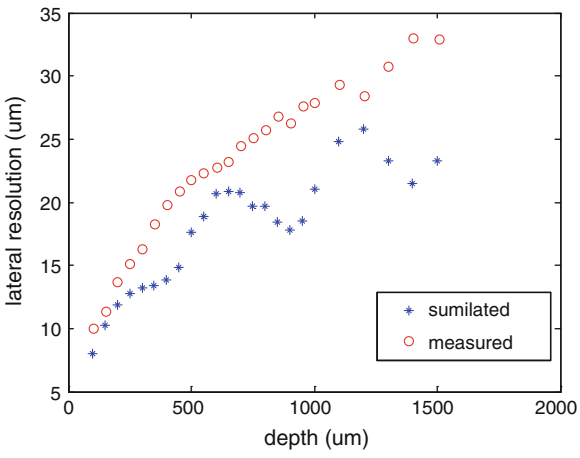

Fig. 3.9 Transverse (lateral) resolution of a bare fiber probe as a function of target depth

3.1.3 Simple Fiber Probe with Thin Lens

A fiber optic imager that requires a longer working distance >0.5 mm requires a lens to form an image at the distal end of the fiber, as shown in Fig. 3.10.

Following the same procedures shown in Sect. 3.1.1., the total ray-transfer matrix can be written as:

$$\begin{aligned} M_{Final} &= M_{Return} M_{FlatMirror} M_{Incident} \\ &= \begin{bmatrix} 1 & d_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} 1 & d_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & d_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} 1 & d_1 \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{3.15}$$

If the distance from the fiber to the lens, d_1, is set to 2f, i.e., $d_1 = 2f$, the equation reduces to:

$$M_{Final} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{2d_2}{f} - 3 & 2d_2 - 4f \\ -\frac{2}{f} + \frac{2d_2}{f^2} & \frac{2d_2}{f} - 3 \end{bmatrix} \tag{3.16}$$

By combining 3.4 and 3.16, we can simplify the Eq. 3.4 as:

$$w_2^2 = \frac{4(d_2 - 2f)^2 + (\frac{2d_2}{f} - 3)^2 \frac{\pi^2}{\lambda^2} w_0^4}{\frac{\pi^2}{\lambda^2} w_0^2} \tag{3.17}$$

The result above shows that the imaging depth is critically dependent on the size of the fiber mode. This effect is expected from the well-known confocal imaging where the depth resolution depends on the size of the aperture used. The detail of this effect and the fiber confocal imager is reviewed in detail in Sect. 3.2 (Fig. 3.11).

3.2 Fiber Optic Confocal Microscope

3.2.1 Review of Fundamentals of Microscope

A fiber optic confocal microscope (FOCM) is a type of confocal microscope in which the confocal pinhole is replaced by an optical fiber. Fundamentals of wide-field and conventional confocal microscopy apply to an FOCM, some of which

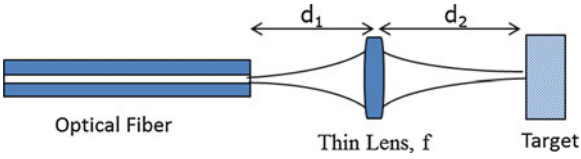

Fig. 3.10 Simple fiber probe with a thin lens having a focal length, f

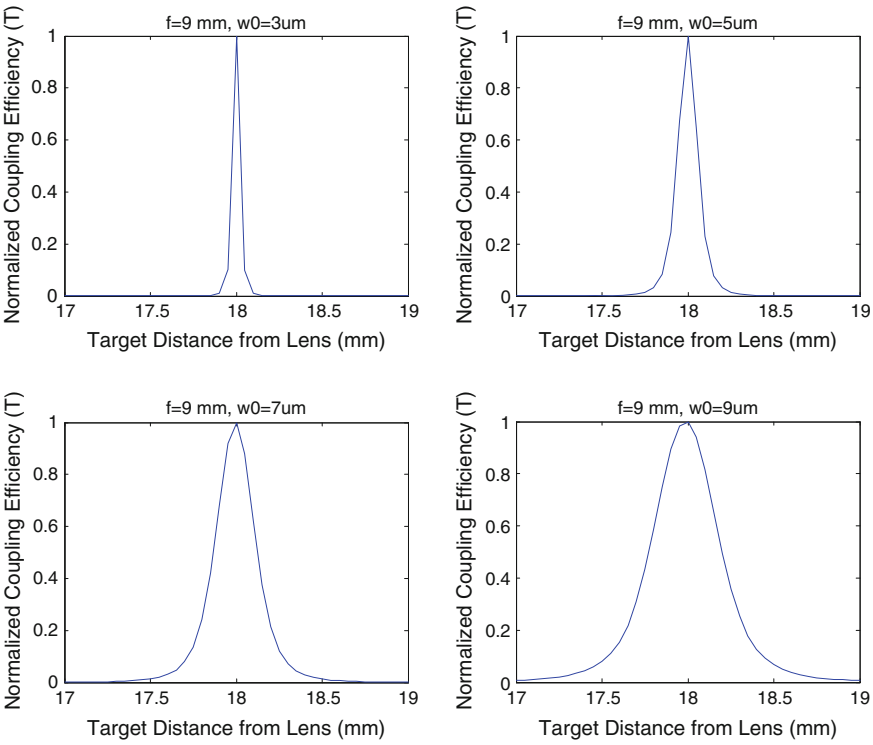

Fig. 3.11 Normalized coupling efficiency as a function of target distance from lens, d_2, for a simple fiber probe with a thin lens

will be introduced in this section briefly. More details of wide-field microscopy and confocal microscopy can be found in other books, for example, Mertz's [1] and Pawley's [2]. The principal goal of microscopy is to enlarge images to a certain extent, so that they can reveal details of objects not perceived by the unaided eye. The ability of a microscope to reveal details, in a more scientific way of description, is defined by the resolving power, which is an instrument property, specifying the smallest detail that a microscope can resolve in imaging an ideal specimen [3]. The related quantity—resolution—refers to the detail actually revealed in the image of a given specimen. Although resolving power refers to a property of the instrument and resolution refers to the level of detail obtained in the resulting image, both are quantified in terms of minimum resolvable distance, $d_{\min}$.

Ernst Abbé described the existence of a resolution limit in 1873 using the diffraction theory, which is now commonly known as Fourier optics. The key idea of Fourier optics is that the image formed from the specimen through an optical system is the superposition of diffracted images of each point of the specimen, and the image of the point itself is its Fourier transform–an Airy disk.

The size of the Airy-disk and its higher-order surrounding diffraction rings limit the highest resolution possible. There are many different ways to define the minimum resolvable separation. The most commonly adapted definition is the Rayleigh criterion: two overlapping Airy-disk images are resolvable when the central maximum of one coincides with the first minimum of the other (see Fig. 3.12a).

For quantitative analysis purposes, the image of a point observed through an optical system is expressed as the point-spread function (PSF). Image obtained by an optical system is a superposition of PSF of an extended object. The intensity PSF is related to both lateral and axial resolution of a microscope. The amplitude PSF can be calculated through a somewhat complicated procedure which can be found in many optics text books (for example, [4]), and is expressed as:

$$h(u,v) = -i\frac{2\pi nA\sin^2\alpha}{\lambda}e^{iu/\sin^2\alpha}\int_0^1 J_0(v\rho)e^{iu\rho^2/2}\rho d\rho \tag{3.18}$$

where λ is the vacuum wavelength, n is the refractive index, $n\sin\alpha$ is the numerical aperture of the lens, A is the area of the illumination, J_0 is the zeroth-order Bessel function. When the light path is in the z direction, v and u are normalized optical units perpendicular and parallel to z, respectively:

$$v = \frac{2\pi nr\sin\alpha}{\lambda} \tag{3.19}$$

$$u = \frac{2\pi nz\sin^2\alpha}{\lambda} \tag{3.20}$$

where, $r = \sqrt{x^2+y^2}$. The intensity PSF is given by $|h(u,v)|^2$.

As mentioned earlier, one way to define the resolution is to use the Rayleigh criterion. In this case, as can be seen in Fig. 3.12a, the resolution corresponds to

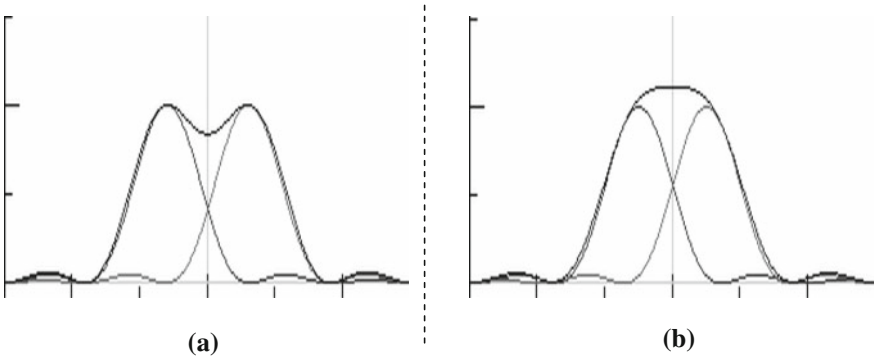

Fig. 3.12 Superposition of Airy disks. **a** Central maximum of one Airy disk coincides with the first minimum of the other. **b** Separation between two Airy disks are minimal to give non-zero superposition

the separation at which a 26 % contrast is achieved. Another commonly used way to define resolution is the separation at which the contrast is non-zero, any value between 0 and 100 % (see Fig. 3.12b). The Rayleigh criterion will be adapted throughout this section.

Under the Rayleigh criterion, we can define the lateral resolution for a conventional microscope as the distance between the central maximum and first minimum of the Airy disk, which is given by:

$$r_{xy} \approx \frac{1.22\lambda}{2NA} = \frac{0.6\lambda}{NA} \tag{3.21}$$

$$r_z \approx \frac{2\lambda n}{NA^2} \tag{3.22}$$

where λ is the wavelength of the light source and NA is the numerical aperture of the objective lens. This equation is convenient for calculating the lateral resolution of a conventional wide-field microscope, at which a reasonable contrast is obtained. In practice, it is difficult to measure the location of an intensity minimum; hence, it is quite common to measure the full-width at half-maximum (FWHM) of the microscope's PSF to determine the resolution experimentally. The measured value of FWHM is approximately 17 % smaller than the calculated value of r_{xy}.

3.2.2 Laser Scanning Confocal Microscopy (LSCM)

The word confocal means "of-same-foci." Confocal microscopy is different from wide-field microscopy because it restricts the detection of signals from the object by using a confocal aperture placed in front of the detector and on the conjugate plane of the image. Another aperture may be placed in front of the illuminator to increase spatial coherency and thus to produce a diffraction-limited illumination spot on the sample; however, this is not completely necessary. This is illustrated in Fig. 3.13.

In a scanning microscope, the illumination and/or focused image is confined to a very small region, usually a diffraction-limited spot; thus, scanning of small spots over the larger specimen area is necessary to reconstruct a usable image containing a region of interest larger than a diffraction-limited spot; hence, the name scanning microscope.

The advantage of a confocal microscope is prominent when a thick sample is being imaged. The sample must be thicker than the depth of field of the objective lens. Figure 3.13a illustrates wide-field illumination combined with wide-field imaging. The entire sample is illuminated and the signal—either scattered light or fluorescence emission from the sample—is projected onto the image plane with a certain thickness. Since the image is recorded by the detector, such as a focal plane array placed at the image plane, the light signal from a thick sample will contribute

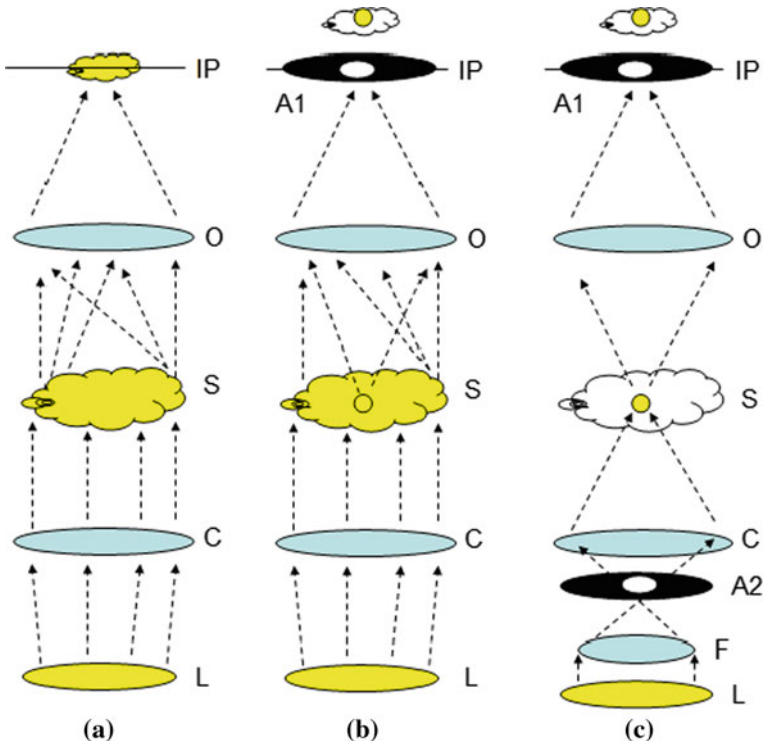

Fig. 3.13 (**a**) Illustration of a wide-field illuminating and wide-field imaging microscope; (**b**) wide-field illuminating and confocal imaging microscope; (**c**) point-illuminating and confocal-imaging microscope. *L* lamp, *C* condenser, *S* sample, *O* objective lens, *IP* image plane, *F* focusing lens, *A1* aperture at the image plane, *A2* aperture at the back-focal-plane of the condenser

to the recorded image in forms of blurriness or decreased contrast. If a confocal aperture is placed in front of the detector, as shown in Fig. 3.13b, then majority of the signal from the conjugate spot in the sample will be projected onto the aperture (thus confocal). Lights scattered or fluoresced from the spot other than the confocal spot will be rejected. The detector does not have to be a focal plane array, thus it can be replaced by a highly efficient photo-multiplier tube (PMT). When the sample or the aperture is raster scanned along the image plane (IP), a two-dimensional (2D) image of the sample only from the conjugate plane of IP is reconstructed. Although the arrangement shown in Fig. 3.13b, which is wide-field illumination combined with confocal detection, represents confocal microscopy, it does not utilize the illuminating power of the light source (L) entirely. As a matter of fact, this illumination scheme is highly inefficient because all the lights are lost except that which illuminates a very small spot in the sample. The arrangement shown in Fig. 3.13c solves this problem by using point illumination and point confocal detection, which is most common in modern confocal microscopy. It also

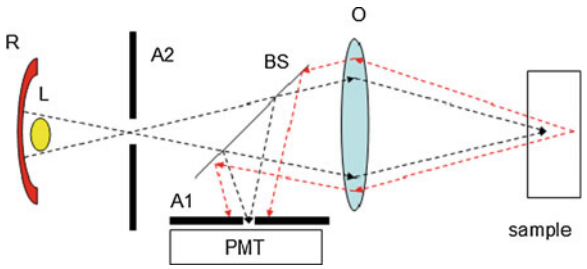

Fig. 3.14 Illustration of an epi-illumination confocal microscope. *R* focused reflector, *L* lamp, *O* objective lens, *A1* detector aperture, *A2* illuminator aperture, *BS* beam-splitter

provides higher rejection of scattered lights compared to Fig. 3.13b. An epi-illumination version of Fig. 3.13 leads to the famous illustration from Minsky's original invention report of confocal microscopy [5], which is shown in Fig. 3.14.

The advantage of epi-illumination is that one lens can play the role of condenser as well as that of objective lens. However, not all types of samples produce a decent epi-illuminated signal; thus, epi- or trans-illumination must be chosen carefully depending on the sample type. As can be seen in Fig. 3.14, the signal from the sample other than the focal spot is rejected by the aperture in front of the detector (A1). If the beam-splitter is replaced by a dichroic mirror that splits the excitation wavelength and emission wavelength, then the confocal microscope shown in Fig. 3.14 runs as a fluorescence confocal microscope.

For confocal microscopy, the FWHM extent of both lateral (xy-plane) and axial (z-direction) PSF are about 30 % narrower than that of a conventional wide-field microscope [6]. This improved PSF can be explained by Lukosz's principle [7]. The improvement is due to the rejection of scattered light from out-of-focus which contributes to the improvement of contrast. As was mentioned in Sect. 3.2.1, the resolving power of the optical system is determined by the optical elements (mostly by lenses), thus it remains the same for a wide-field microscope and confocal microscope when the same lenses are used. Resolution is the result induced from both resolving power and contrast; thus, resolution of a confocal microscope is slightly higher than that of a wide-field microscope, which is quantified as 30 % from experimental measurements and mathematical calculations [6].

$$r_{xy,confocal} \approx \frac{0.4\lambda}{NA} \tag{3.23}$$

For point-like objects, the same approach can be applied to determine a microscope's axial resolution. As in the lateral direction, the same Rayleigh criterion can be applied, resulting in:

$$r_{z,confocal} \approx \frac{1.4\lambda n}{NA^2} \tag{3.24}$$

Here, n stands for the refractive index of the specimen. As n is normally larger than 1 and NA is smaller than 1, the axial resolution is poorer than the lateral resolution. Practically, there are several technical concerns which reduce the

resolution of the system: the size of the pinhole and pixilation of the confocal microscope, noise of the detector, and the precision of the focusing mechanism.

Note that placing a pinhole with an arbitrary diameter will not guarantee the confocal operation of the system. If a pinhole is large enough that all the incoming signals—including that from the scattered light—pass through the pinhole, then the pinhole does not play any role; the system is simply a wide-field microscope with a limited field of view. Also, if a pinhole is too small, then the signal level will be too low to acquire optimal image quality. Diffraction from the pinhole will also affect the image quality negatively. Wilson et al. [8] showed that optimal confocality of the system is achieved when the optical coordinate v of Eq. 3.19 is less than 0.5.

$$v = \frac{2\pi n r \sin\alpha}{\lambda} < 0.5 \tag{3.25}$$

This can be rewritten as:

$$r < \frac{\lambda}{4\pi n \sin\alpha} \tag{3.26}$$

For example, when a 40 × NA 0.6 numerical aperture objective lens is used in the setup shown in Fig. 3.14, the diameter of the pinhole needs to be 6.7 μm or less for the system to be operated in a confocal mode with 632.8 nm He–Ne laser in air (n = 1). The smaller the pinhole diameter is, the higher the confocality of the system becomes; however, due to the signal loss, the pinhole does not have to be practically smaller than 6.7 μm. Improvement in resolution by using a pinhole smaller than 6.7 μm is also negligible.

3.2.3 Laser Scanning Confocal Microscope

A laser scanning confocal microscope (LSCM) is widely used; however, there are other types such as a spinning (Nipkow) disk confocal microscope [9] and a programmable array microscope [10]. A fiber-optic confocal microscope (FOCM) is a type of LSCM, thus only LSCM is considered in this chapter. Compared to other confocal microscopes, LSCM is unique because it uses a laser as the light source and the laser is scanned, not the sample or the detector. As mentioned previously, using laser as a light source enables achieving a diffraction-limited focal spot due to the high spatial coherence of laser. However, high coherence sometimes decreases the image quality of LSCM when interference is formed between the slide cover and objective lens. Scanning the sample—or sometimes, the detector—gives the highest image quality without the spherical aberrations induced in laser scanning. However, scanning the laser almost always achieves the highest scanning speed, thus the highest image acquisition rate. A typical LSCM setup is shown in Fig. 3.15.

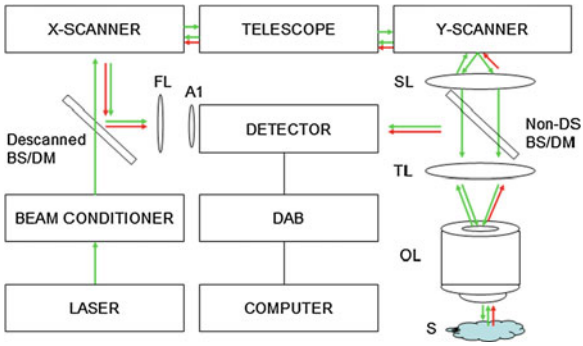

Fig. 3.15 Schematic of a typical LSCM. *BS* beam-splitter, *DM* dichroic mirror, *SL* scanning lens, *TL* tube lens, *OL* objective lens, *S* sample, *DAB* data acquisition board, *FL* focusing lens, *A1* confocal aperture

The example shown in Fig. 3.15 uses green (532 nm) laser as the light source. The beam conditioner usually contains a power adjustment element such as a variable neutral density filter or Pockel's cell, beam expander, laser shutter, polarizer, laser power monitor, spatial filter, regenerative amplifier, etc. Any of these components can be omitted if necessary. Beam-splitters are in general 50:50 splitters, and can be replaced by dichroic mirrors for fluorescence-mode operation. The X-scanner is in general a fast-scanner and can be a mirror scanned by galvo-motor, resonant-motor, or rotating mirror block. The Y-scanner is a slow-scanner, and a galvo-motor is commonly used. An acousto-optic tunable filter can also deflect the laser beam, and thus is used as the fast-scanner. A telescope between the X- and Y-scanners is not always necessary; however, it helps to reduce comma-errors in scanning by placing the pivotal point of the scanning mirrors at the back-focal plane of the scanning lens. When a telescope is not used, the X- and Y-scanners should be placed as close as possible. The relationship between SL, TL, and OL will be explained in more detail.

If the reflected light or fluorescence emission is captured by the detector using a non-descanned (DS) beam-splitter or dichroic mirror, the microscope is being operated under non-descanned mode. This is the common arrangement in two-photon microscopes. The microscope is not a confocal microscope, although it is still a laser scanning microscope. For this type of arrangement to be operated under a confocal regime, the signal must be collected using the descanned BS or DM, and the detector must be accompanied by the focusing lens (FL) and the aperture placed at the focal plane of the FL. IF FL and A1 are removed, the setup is no longer a confocal microscope, even though the signal is descanned. Descanned detection without a confocal aperture is almost meaningless, except that the illumination laser intensity can be higher than a wide-field microscope. Non-descanned confocal operation is impossible because the pinhole position cannot be fixed for a non-descanned (thus, still scanning) beam.

Figure 3.16 shows fluorescence microscopy images obtained by using both a wide-field fluorescence microscope and LSCM. The sample was a prepared test slide from Invitrogen (#F-24630), which is a mouse kidney section stained with Alexa Fluor 488 WGA, Alexa Fluor 568 Phalloidin, and DAPI. Figure 3.16a is a

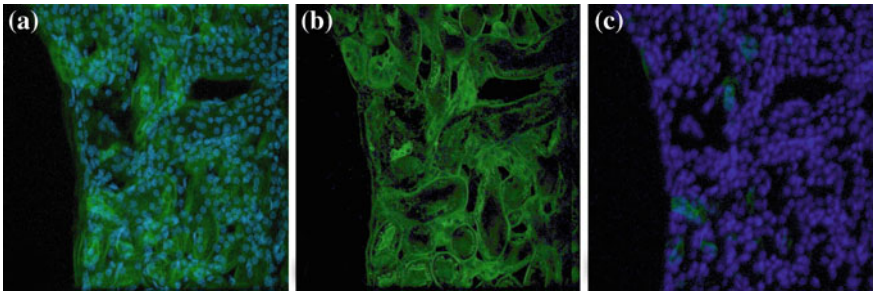

Fig. 3.16 (**a**) Wide-field fluorescence image; (**b**) Confocal fluorescence image from channel 1; and (**c**) channel 2. See text for more detail

wide-field image with 405 nm excitation. Emissions from both Alexa Fluor 488 (shown in green in the figure) and Alexa Fluor 568 (shown in blue in the figure) can be seen in one image. Figure 3.16b and c are confocal fluorescence images using 405 nm excitation and two PMT channels separated by a dichroic filter: Channel 1 is for >505 nm and Channel 2 is for <505 nm. Both channels are overlaid in Fig. 3.16b and c. The difference between Fig. 3.16b and c is their sample z-position. The depth discrimination capability of confocal microscopy can be clearly seen in both figures. At different depths—not like the wide-field image shown in Fig. 3.16a—only one emission can be dominantly seen in Fig. 3.16b and c. Cells and organelles in kidney are stained by different dyes, thus their depth positions in z-direction are separated by different emissions. Also, Fig. 3.16b and c demonstrate sharper boundaries compared to Fig. 3.16a, which is enhanced by rejection of scattered light. Depth discrimination and rejection of scattered light (thus a slightly higher resolution) are two major advantages of confocal microscopy over wide-field microscopy. Optical arrangement of scanning lens, tube lens, and objective lens is critical for successful operation of LSCM in confocal mode, along with achieving high confocality of the system.

Objective lenses are designed so that their highest diffraction-limited resolution is achieved when their back apertures are filled with collimated monochromatic light. Since the incoming light fills the back aperture of OL, scanning can be achieved only by changing the incident angle (θ) of the incoming light relative to the optical axis of OL, which is illustrated in Fig. 3.17a. Each OL has the maximum allowable incident angle (θ_{max}) which produces a usable image of the sample with minimal spherical aberration. Any incoming light with an incident angle larger than θ_{max} will result in a distorted or unevenly illuminated image. The back-aperture size of commonly used OL varies greatly: for example, Zeiss W Plan-APO 20x/1.0NA lens which is widely used for in vivo imaging has a back aperture with 18.0 mm diameter while Zeiss W Plan-APO 60x/1.0NA lens has a 6.0 mm aperture. Lasers usually produce a collimated Gaussian output beam with a diameter of 1.0–2.0 mm. Thus, SL and TL must provide a proper incident angle (θ) with a properly expanded beam diameter to OL. The less the SL-TL pair expands the laser beam, the less the distortion from the scanning; however,

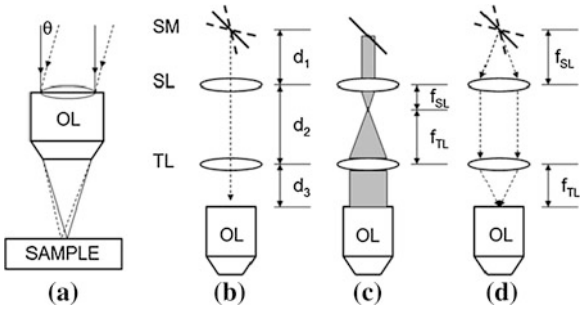

Fig. 3.17 (**a**) Illustration of the incident angle of an incoming scanning beam that fills the back aperture of an objective lens (*OL*); (**b**) Illustration showing the relative distance between scanning mirror (*SM*), scanning lens (*SL*), and tube lens (*TL*); (**c**) Illustration showing the function of SL and TL which expands the incoming collimating beam to fill the back aperture of the objective lens (*OL*). The sum of the focal lengths of SL (f_{SL}) and TL (f_{TL}) equals d_2; (**d**) Distance between the pivotal points of SM and SL needs to be f_{SL}, and distance between TL and back aperture of OL needs to be f_{TL}

scanning mirrors are limited for their sizes because their scanning speed is limited by the inertia of the mirrors. The role of beam expansion by an SL-TL pair is inevitable.

Figure 3.17b–d show the relative distances between SM, SL, and TL that are required for the optics to achieve proper beam expansion and scanning. Optimal performance of the scanning system critically depends on this optical arrangement. Focal lengths of SL (f_{SL}) and TL (f_{TL}) are chosen so that the pair provides the necessary beam expansion, as shown in Fig. 3.17c. For example, if the input beam diameter is 3 mm and the back aperture of OL is 6 mm, the SL-TL pair must provide 2-fold beam expansion; thus, $f_{TL}/f_{SL} = 2$ and $d_2 = f_{SL} + f_{TL}$. For the scanning system to achieve the most uniform illumination while providing incident angle to OL as large as θ_{max}, the distance between SL and the pivotal axis of SM needs to be f_{SL} ($= d_1$). Also, for the scanning beam to pass through the back aperture of OL with minimal translational shift, the distance between TL and the back aperture plane of OL needs to be f_{TL} ($= d_3$).

Measurement of the enhanced lateral (xy-plane) resolution of a confocal microscope is difficult, because there is no standardized resolution test-target that simulates a thick scattering biological specimen. If a generic USAF 1951 resolution target is used, the enhancement in lateral resolution may not be clearly detected; however, enhancement of axial (z-axis) resolution can be measured relatively easily. A simple method to measure the axial response is to scan a perfect reflector axially through its focus and measure the signal strength. Figure 3.18 shows a typical response.

A confocal system with less than an optimal optical arrangement will demonstrate a highly asymmetrical or Gaussian axial response. Practically, achieving perfect axial response, as shown in Fig. 3.18, is highly challenging. Asymmetry is

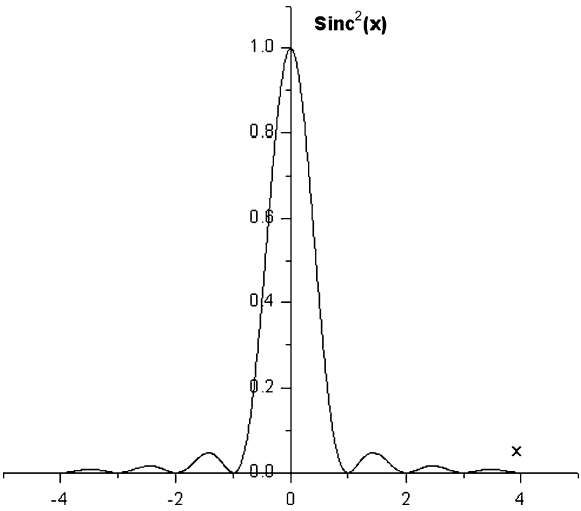

Fig. 3.18 A typical axial response of a confocal microscope–a squared Sinc function

hard to remove completely due to the optical design of the objective lens itself (mainly due to spherical aberration); however, it should be minimized. Other factors depending on the mirror defocus, such as system magnification, effective numerical aperture, and axial position also contribute to the asymmetric axial response.

3.2.4 Introduction to Fiber-Optic Confocal Microscopy

Due to their flexibility and robustness, optical fibers enable delivery of lights to hard-to-reach places as well as to collect them. Optical fibers also enable the separation of the optical components such as the light source and detectors from each other. Optical fibers have been adapted in confocal microscopy since the early 1990s [11] and can be found in many commercial wide-field microscope systems and endoscopic microscopy (endo-microscopy) systems. However, flexibility and robustness are not the only factors that make optical fibers unique in confocal microscopy applications. There are strong advantages of using optical fibers in confocal microscopy, including the following:

1. The output from a small-core single-mode fiber serves as an ideal Gaussian point-like light source, which contributes to achieving the diffraction-limited focal spot of the objective lens, thus achieving the highest possible resolution from given optics. This advantage is especially useful when a cost-effective diode laser is the light source, which does not have a Gaussian profile. However, this advantage is not very useful when a laser system with TEM_{00} mode output is available.

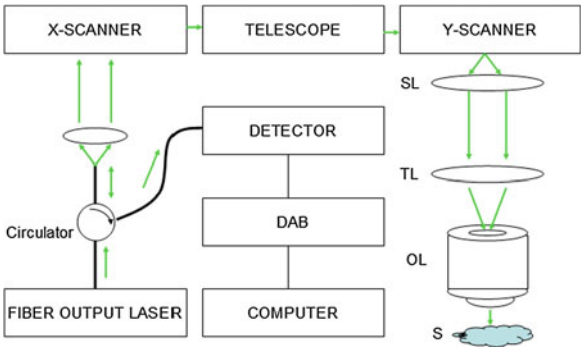

Fig. 3.19 Schematics of FOCM. Single-mode fiber tip serves as the point light source as well as the pinhole detector. Adaptation of an all-fiber circulator enables the entire construction to be simple and robust. Scanning and microscopic optics remain the same as the bulk optics version of Fig. 3.15

2. The core of an optical fiber is an ideal pinhole, in terms of shape and size. As described in Sect. 3.2.2, the pinhole diameter required for a NA = 0.6 lens with 632.8 nm laser in air is a few μm. A conventional pinhole with this size is very hard to produce: the shape of the pinhole can be irregular or the size can be incorrect. However, typical single-mode fibers have cores with a diameter in the order of a few μm. Once the signal is coupled into the detecting fiber, its delivery of light to the detector can be arranged in various ways so that the light collection efficiency can be maximized.
3. One of the disadvantages of using optical fiber as the pinhole is reduced collection efficiency. To efficiently couple the light into an optical fiber, not only the size of the focused signal light needs to be matched to the fiber core, but also the light field and numerical aperture must match those of the optical fiber. Adjusting the field and numerical aperture of signal light is almost impossible when the signal is a low-coherent fluorescence from biological samples. Also, since biological samples exhibit various levels of signal—whether it be fluorescence or reflectance—sometimes the size of the pinhole needs to be enlarged at the expense of reduced confocality. Many commercial confocal microscopes adapt multi-mode fiber with core size as large as 600 μm accompanied by a variable pinhole placed in front of the fiber tip. More details about collection efficiency will be discussed in a following section.
4. An all-fiber coupler or circulator makes it possible to use the fiber as the output source as well as the input source at the same time. This arrangement, as shown in Fig. 3.19, might be the best application of optical fibers in confocal microscopy in terms of utilizing the advantages of optical fiber over pinholes. One technical difficulty in such implementation stems from the fact that both ends of the fiber are in the image planes of the microscope. As a result, Fresnel reflection of the incident laser at the fiber tip can reduce the signal level. Careful management of reflections at the optical surfaces using angled fiber tips, anti-reflection coatings, and index-matching fluids helps reduce this unwanted effect [2].
5. Use of a miniature lens (such as a graded-index lens) attached to the tip of the optical fiber enables significant reduction of the optics size.

3.2.4.1 Coherent Detection

Once collimated using a collimating lens, the laser output from an optical fiber is no different than that from solid-state or gas lasers. Optical fiber as the light source in FOCM does not affect image characteristics due to the characteristics of the fiber. However, there is a significant difference between pinhole and optical fiber as detectors because optical fiber acts as a coherent detector. Single- and multi-mode fibers both serve as coherent detectors. Also, the spatial distribution of modes within a multi-mode detection fiber may also carry information regarding the source of the signal [2]. Not only does the size of the core of the optical fiber affect the confocality and signal level as the pinhole does, but also the mode of the incoming signal and the numerical aperture of the coupling lens affect the overall performance. For a pinhole, the signal is integrated over the area of the pinhole, while the signal is integrated over the modes at the entrance of the fiber. Since axial response is more affected by the coherent detection, the axial response of FOCM will be discussed in this section. Theoretical analyses used in this section mostly follow Sheppard and Wilson's excellent work in the 1990s.

When a pinhole is used in the confocal microscope detector, the detection is incoherent; thus, the axial intensity of the light is integrated over the area of the pinhole detector, such that:

$$I_i(z) = \int_{D_i} |E|^2 S_i dD_i \tag{3.27}$$

where D_i is the area and S_i is the sensitivity of the incoherent detector. The electric field is such that:

$$\begin{aligned} |E|^2 &= E_x E_x^* + E_y E_y^* + E_z E_z^* \\ &= |I_0|^2 + 2\Re(I_0 I_2^*)\cos\left[2(\gamma - \theta_p\right] + |I_2|^2 + 2|I_1|^2\cos^2(\gamma - \theta_p) \end{aligned} \tag{3.28}$$

where γ denotes the angle of incident polarization with respect to x-axis, θ_p is the angular component of the position vector, and $\Re$ represents real-part. I_0, I_1, and I_2 are intensity components with respect to the polarization vector, which are related to the electric field as such:

$$\begin{aligned} E_x &= -\frac{\kappa}{2}\left[\cos\gamma\left(I_0 + I_2\cos 2\theta_p\right) + \sin\gamma I_2 \sin 2\theta_p\right], \\ E_y &= -\frac{\kappa}{2}\left[\cos\gamma I_2 \sin 2\theta_p + \sin\gamma(I_0 - I_2\cos 2\theta_p)\right], \\ E_z &= -i\kappa\left[\cos\gamma I_1 \cos\theta_p + \sin\gamma I_1 \sin\theta_p\right] \end{aligned} \tag{3.29}$$

Using Eq. 3.27 in Eq. 3.28 for spherical coordinate, we obtain:

$$I_i(z) = \int_0^R \left(|I_0|^2 + 2|I_1|^2 + |I_2|^2\right)\rho d\rho \tag{3.30}$$

where R is the radius of the detector aperture.

For coherent detection using an optical fiber, the axial intensity is the integration over the field mode of the incident light which is given as:

$$I_c(z) = \left| \int_{D_c} ES_c dD_c \right|^2 \tag{3.31}$$

Since the terms of Eq. 3.29 consisting of $\sin\theta_p$, $\sin 2\theta_p$, $\cos\theta_p$, and $\cos 2\theta_p$ yield zero when D_c, Eq. 3.31 can be expressed as:

$$I_c(z) = \left| (\cos\gamma + \sin\gamma) \int_0^R I_0 \rho d\rho \right|^2 \tag{3.32}$$

Different from the incoherent detection of Eq. 3.30, the coherent detection shown in Eq. 3.32 exhibits explicit dependence on the polarization angle of the incident light, γ. When the radius of both the pinhole and the optical fiber core becomes point-like, the axial responses from both systems become identical.

3.2.4.2 Signal Collection

When the shot noise is the only source of the noise in the imaging system, then the signal-to-noise ratio (SNR) is given as [12]:

$$SNR = \frac{n_p}{\sqrt{n_p}} = \sqrt{n_p} \tag{3.33}$$

where the number of photons (n_p) is related to the incident power (P), wavelength (λ), incident time (t), Planck's constant (h), and speed of light (c), such that:

$$n_p = \frac{Pt\lambda}{hc} \tag{3.34}$$

For a detection system with quantum efficiency Q_E, SNR is given in terms of Q_E, n_p, and the number of involved electrons, n_n:

$$SNR = \frac{Q_E n_p}{\sqrt{Q_E n_p + n_n^2}} \tag{3.35}$$

SNR for a confocal microscope when assuming negligible electronic noise ($n_n = 0$) is given as:

$$SNR = \sqrt{Q_E n_p} \frac{F(v)}{\sqrt{F(v) + av^2/\sqrt{4}}} \tag{3.36}$$

where v is the normalized coordinate of Eq. 3.19, F(v) is the fraction of signal incident on the pinhole, and a is a constant representing the background noise

level. Background noise in majority of confocal microscopy systems is out-of-focus scattered light from thick specimen. For a circular pinhole and small v:

$$SNR = \sqrt{Q_E n_p}\frac{v}{2\sqrt{1+a}} \tag{3.37}$$

Equation 3.37 tells us that SNR increases as the pinhole diameter r, thus v, increases, and that as the background noise level—thus a—decreases.

In FOCM, SNR of the confocal microscopy expressed by Eq. 3.37 holds; however, the limited acceptance angle of the optical fiber reduces the signal level, which is one of the biggest differences from bulk confocal microscopy. This can be expressed by numerical aperture of optical fiber (NA_{fiber}). Normalized frequency, or V-number, of an optical fiber is given by:

$$V = \frac{2\pi r}{\lambda}\sqrt{n_1^2 - n_2^2} = \frac{2\pi r}{\lambda} NA_{fiber} \tag{3.38}$$

where r is the core radius, n_1 and n_2 are refractive indices of core and cladding, respectively. NA_{fiber} represents the numerical aperture of the fiber. Thus, for a given optical fiber and laser, confocality condition Eq. 3.26 can be re-written in terms of NA_{fiber}:

$$r = \frac{\lambda V}{2\pi NA_{fiber}} < \frac{\lambda}{4\pi n \sin\alpha} \tag{3.39}$$

After some arrangement, the requirement for the NA_{fiber} for confocal operation of a microscope becomes:

$$V 2n \sin\alpha < NA_{fiber} \tag{3.40}$$

This gives a good insight into how numerical apertures of the objective lens and the fiber play their role in confocal microscopy. V is a value smaller than 2.405 for single-mode operation of the fiber. Refractive index, n, is in general larger than 1. Thus, NA of the objective lens ($\sin\alpha = NA_{obj}$) must be much smaller than NA_{fiber}, or NA_{fiber} must be much larger than NA_{obj}. To achieve high efficiency of coupling light into an optical fiber, the focusing lens must provide high NA_{obj} so that the focused spot size can be comparable to the core size. High NA_{obj} may sacrifice confocality because of Eq. 3.40; thus, a compromise between coupling efficiency and confocality must be found. Lowered coupling efficiency causes lowered SNR. Since the detection fiber does not have to be operated in single-mode, a large-core high-NA optical fiber is more suitable for detection fiber, which gives both high coupling efficiency and confocality.

The same method used in Sect. 3.1.3 can be used to study the normalized coupling coefficient for different objective lenses in a single-fiber confocal setup. In this single-fiber confocal setup, an optical fiber is used for both illumination and detection. Assuming the profile of the input beam exiting from the fiber tip has a Gaussian distribution as before, the ray-transfer matrix can be calculated as [13]:

$$A = D = -\frac{f_2^2 - 2L_1f_2 + 2f_1f_2 - 2L_2f_2 + 2L_2L_1 - 2L_2f_1}{f_2^2}$$
$$B = -2\frac{f_1^2}{f_2^2}(f_2 - L_2) \tag{3.41}$$
$$C = -\frac{2(f_1 + f_2 - L_1)(f_1f_2 - L_1f_2 + L_2L_1 - L_2f_1 - L_2f_2)}{f_1^2f_2^2}$$

where f_1 and f_2 are the focal lengths of the two microscopic objective lenses, L_1 is the distance between the two lenses, and L_2 is the distance between the sample membrane and the f_2 lens (Fig. 3.20). Simplifications to the above expressions can be achieved if we assume $L_1 = f_1 + f_2$:

$$\mathrm{A} = \mathrm{D} = 1,\ \mathrm{C} = 0,\ \mathrm{B} = -2\frac{f_1^2}{f_2^2}(f_2 - L_2) \tag{3.42}$$

Because the input beam exiting from the fiber tip has a flat wavefront, q_1 can be written as:

$$\frac{1}{q_1} = \frac{1}{R_1} - i\frac{\lambda}{\pi w_0^2} = -i\frac{\lambda}{\pi w_0^2} \tag{3.43}$$

where w_0 and R_1 are the beam waist and the radius of curvature of the input Gaussian beam, respectively. By substituting Eqs. (3.42 and 3.43) into (3.40), the q-parameter of the return beam can be easily written as

$$q_2 = q_1 - 2\frac{f_1^2}{f_2^2}(f_2 - L_2) = i\frac{\pi w_0^2}{\lambda} - 2\frac{f_1^2}{f_2^2}(f_2 - L_2) \tag{3.44}$$

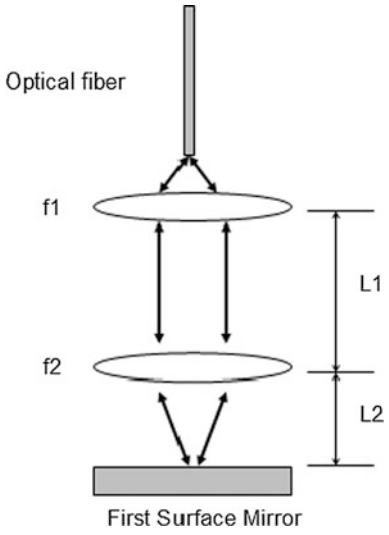

Fig. 3.20 Schematic of a single fiber confocal microscopy setup for studying power collection ratio

Since we also have

$$\frac{1}{q_2} = \frac{1}{R_2} - i\frac{\lambda}{\pi w_2^2} \tag{3.45}$$

by comparing the real and imaginary parts of Eq. (3.44) and (3.45), we can obtain the radius of curvature R_2 and the waist w_2 of the output beam, respectively:

$$R_2 = -\frac{(\frac{\pi w_0^2}{\lambda})^2 + 4(\frac{f_1}{f_2})^4 (f_2 - L_2)^2}{2\frac{f_1^2}{f_2^2}(f_2 - L_2)} \tag{3.46}$$

$$w_2^2 = (\frac{\lambda}{\pi w_0})^2 \left[(\frac{\pi w_0^2}{\lambda})^2 + 4(\frac{f_1}{f_2})^4 (f_2 - L_2)^2 \right] \tag{3.47}$$

The optical power collected by the fiber tip is the integral of the beam intensity over the fiber core area at the fiber tip plane, as was the case in Eqs. 3.9, and 3.10 can be used to calculate T, which shows the confocal effect of the system.

3.2.4.3 Variations of FOCM

The FOCM arrangement shown in Fig. 3.21 is a setup which utilizes the advantage of using optical fiber as a point-like light source and as a point coherent detector. The FOCM of Fig. 3.21 is especially useful in establishing a cost-effective confocal microscopy system because a low-cost diode laser can be used, and other optical components such as beam-splitter and pinholes can be avoided. However, it does not utilize the advantage of optical fiber in terms of flexibility and robustness, which enables compact remote sensing. Beam scanning at the distal end of the optical fiber needs a special design.

One of the commonly used arrangements of compact remote sensing FOCM is shown in Fig. 3.22. The bulk scanning optics of the arrangement in Fig. 3.21 were replaced with a miniature endoscopic fiber scanner. The micro actuator translates the fiber tip, which serves the role of a point-like light source and point detector at the same time in a raster or spiral scanning pattern in a small package. The descanned detector through the scanning fiber tip ensures confocal operation of the system. The objective lens can be a common spherical lens or graded index (GRIN) lens. The entire scanning section can be sealed in a compact package; thus, this arrangement is suitable for endoscopic application. However, mechanical scanning of the fiber tip in a small endoscopic package may limit the size of the entire package.

Another method for compact remote scanning is shown in Fig. 3.23 [14]. This arrangement uses an optical fiber bundle to deliver the scanning light signal from and to the scanning microscope. The scanning microscope does not have to operate under a confocal regime, but it has to be able to deliver an illumination laser into a single core of the fiber bundle while detecting a signal from the same core. Optical

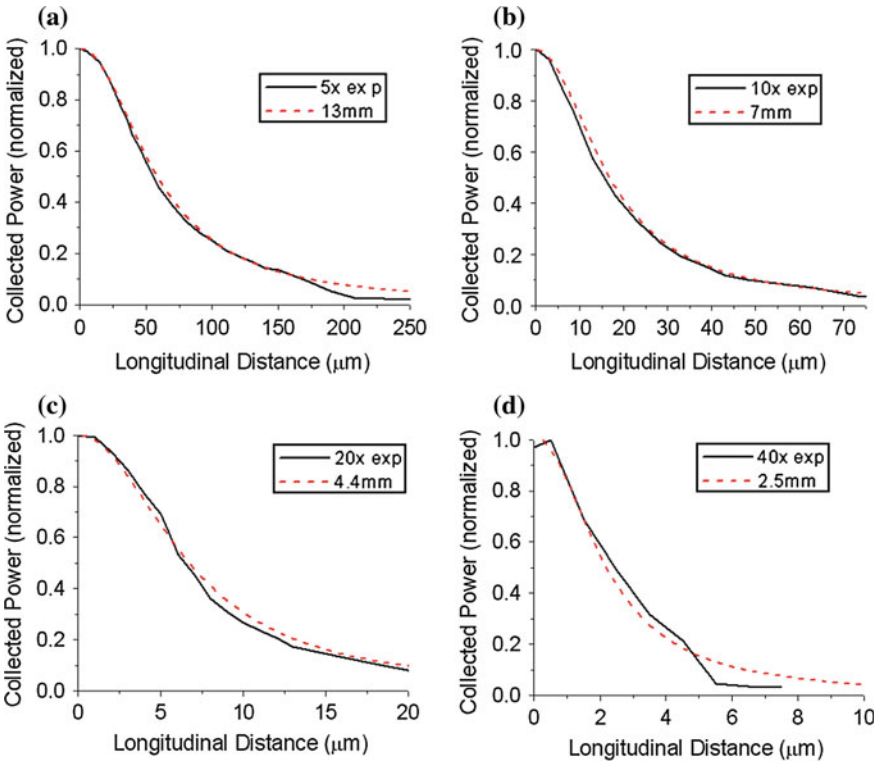

Fig. 3.21 Experimental (*solid*) and simulation (*dashed*) results showing the fiber optic confocal effect for four different microscopic objective lenses. The measured FWHM of the longitudinal confocal response was found to be (*a*) 5X: 102 μm, (*b*) 10X: 28.4 μm, (*c*) 20X: 11 μm, and (*d*) 40X: 4.4 μm, respectively

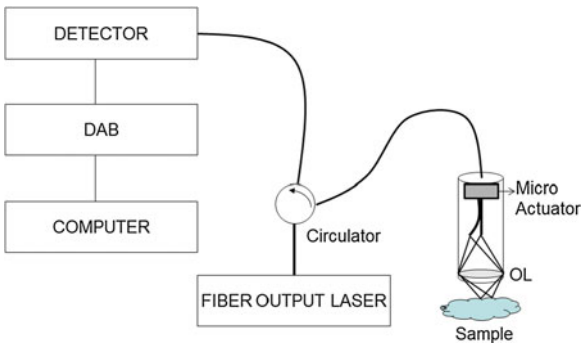

Fig. 3.22 FOCM with scanning implemented on the distal end of the optical fiber. The micro actuator introduces raster or spiral scanning of the fiber-tip. OL represents the objective lens

arrangement of the fiber bundle and objective lens should be such that the detection is achieved in high confocality. Using a GRIN lens will further minimize the size of the setup.

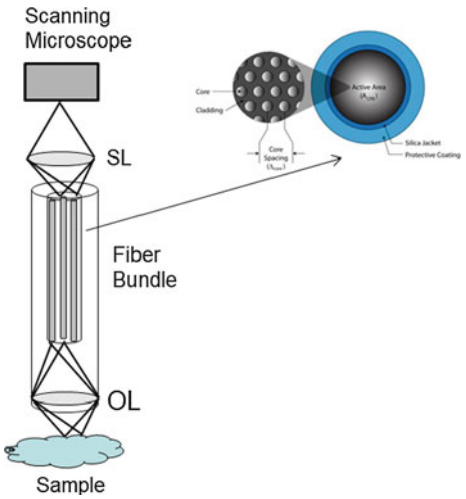

Fig. 3.23 Endoscopic FOCM based on fiber bundle. Inset shows the multiple core structure of the fiber bundle

3.3 Novel Application of FOCM

Confocal microscopy has been widely used in many practical fields ranging from basic science to industrial applications due to its depth discrimination property, which can be utilized for three-dimensional imaging of thick specimens [15–17]. The disadvantages of conventional pinholes used in normal confocal microscopes have been overcome by replacing the pinholes with optical fiber [11], resulting in the achievement of higher resolution [18]. Hollow-core fibers have been developed since the surgical application of mid-infrared (IR) lasers was established. However, typical hollow-core fibers have been made of metal-coated or dielectric-coated quartz or plastic tubes, and they have larger than a few hundred micrometers of diameter due to technical difficulties of production. Mid-IR applications and hollow-core fiber deliveries are well summarized in many chapters

Yablonovitch was the first to propose that the periodicity in dielectric materials could prevent the propagation of electromagnetic waves within a certain frequency range. The periodicity of the dielectric constant induces the removal of degeneracy of the free-photon states at Bragg planes and provokes a range of forbidden energies for the photons. This has led to the appearance of a photonic crystal also known as photonic bandgap material. There is a close analogy between electrons in semiconductors and photons in photonic crystals. Many concepts valid for electrons can now be extended to photons.

From the Maxwell's equations, we can eliminate the electric field E(r) and write the wave equation in terms of the magnetic field H(r):

$$\nabla \times \left(\frac{1}{\varepsilon(r)} \nabla \times H(r) \right) = \frac{\omega^2}{c^2} H(r) \tag{3.48}$$

And the transverse requirement needs to be satisfied:

$$\nabla \cdot H(r) = 0 \tag{3.49}$$

This is an eigenvalue problem. For a specific structure, ε(r) and H(r) are solved for some frequency. The electric field can be calculated by the following relation:

$$E(r) = \left(\frac{-i}{\omega\varepsilon(r)}\right)\nabla \times H(r) \tag{3.50}$$

Let us consider a one-dimensional (1-D) structure as shown in Fig. 3.24, which is constituted of an array of alternating dielectric materials with dielectric constants n_1 and n_2, and a period of a. From Bloch's theorem, the eigenstates of the translation operator in x direction, which are plane wave $e^{ik_x x}$, are also the eigenstates of the wavefunction Eq. (3.48). Thus the one-dimensional eigenstate of magnetic field can be expressed as:

$$H_k(x) = e^{ik_x x}u_k(x) \tag{3.51}$$

After a somewhat complex mathematical procedure, applying Eq. (3.51) into Eq. (3.48) leads to:

$$\Theta_k u_k(x) = \left(\frac{\omega(k)}{c}\right)^2 u_k(x) \tag{3.52}$$

where Θ_k is a Hermitian differential operator defined as

$$\Theta_k = (ik_x + \partial x) \times \left(\frac{1}{\varepsilon(x)}(ik_x + \partial x)\times\right) \tag{3.53}$$

We find from Eqs. (3.52) and (3.53) a very familiar structure of quantum physics: a Hamiltonian eigenvalue problem. Using the analogy of quantum physics, we can induce a dielectric version of Bloch's theorem

$$u_k(x) = u_k(x + a) \tag{3.54}$$

from which the dielectric version of the Brillouin zone is also predictable. Omitting all the details of quantum physics analogy, two conclusions can be made: (1) For an infinite periodic structure similar to Fig. 3.24, all the eigenfunctions can be expressed in terms of the eigenfunction of the first Brillouin zone (any one pair of n_1 and n_2 in Fig. 3.24); (2) Angular frequency ω(k) which appears in Eq. (3.52)

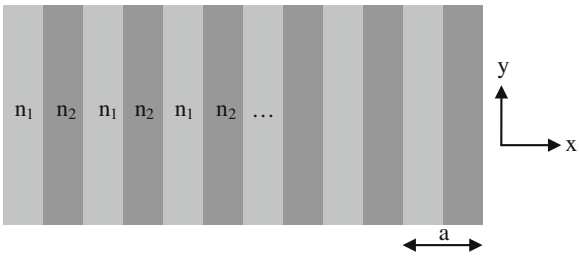

Fig. 3.24 One-dimensional periodic array of dielectric material with refractive index n_1 and n_2

can have only values satisfying the dispersion relation (solid line), which is graphically expressed in Fig. 3.25.

Figure 3.25 shows that some range of frequency values are not allowed for a certain value of wave vectors, which is marked as the shaded area named photonic bandgap. That is to say, in a periodic dielectric structure, light having a certain frequency (wavelength) cannot exist. By placing different photonic bandgap materials in a well-planned manner, a photonic well—an analogy of a quantum well in quantum physics—can be constructed and light can be confined only in a desired region. Extension of the above periodic 1-D structure to x–y plane, and placing the same structure along z-axis, a new type of light-guiding structure can be constructed–a photonic bandgap (or photonic crystal) fiber [19]. Photonic bandgap fibers have such versatility that virtually any size or shape of any wavelength of light-guiding core can be obtained by proper design of the bandgap structure and usage of bandgap materials. Photonic crystal fibers have recently been developed in various structures and have theoretically and experimentally been shown to have several advantages over conventional silica-based fibers, e.g., low temperature dependence [20], hollow-core mid-IR guidance [21]. Our main interest is to utilize hollow-core photonic bandgap fiber in single-fiber confocal microscopy, which can lead to real-time imaging while delivering mid-IR surgical laser power in the same single fiber.

A confocal microscope using a single optical fiber was built as shown in Fig. 3.26 [22]. A 532 nm solid state laser and Crystal Fiber's HC-530-01 hollow-core photonic bandgap fiber (HC-PBF, core diameter: 5 μm, center wavelength: 511 nm) were used for the experimental setup; however, other lasers and fibers—such as 632.8 nm He–Ne laser and gold-coated hollow-core fiber (HCF, core diameter 700 nm)—were also used for comparison. To obtain optimum results, we tried different objective lenses for coupling, collimating, and focusing (OL1-OL3),

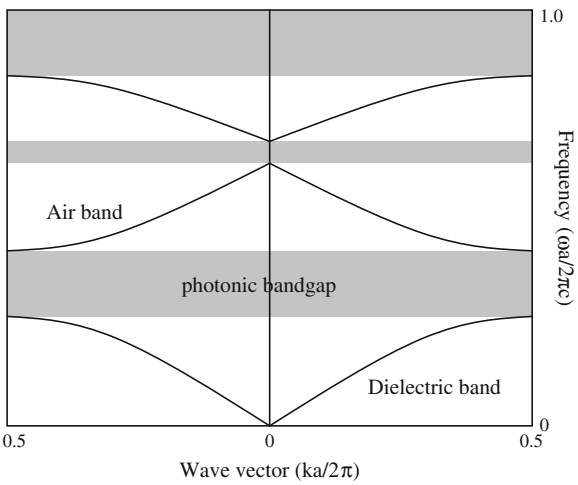

Fig. 3.25 Dispersion relation of 1-D photonic bandgap material in the first Brillouin zone

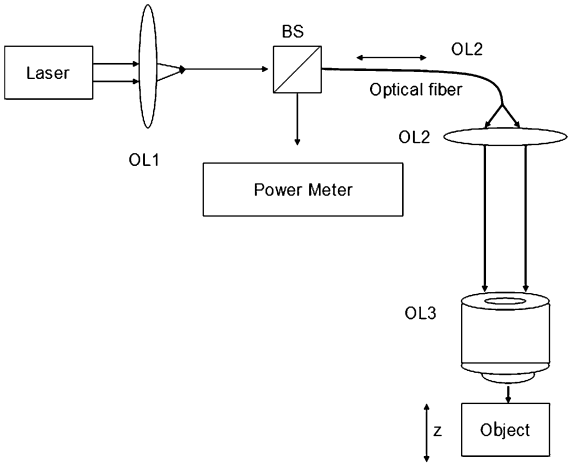

Fig. 3.26 Schematic of experimental setup. *BS* beamsplitter, *OL1* coupling lens, *OL2* collimating lens, *OL3* confocal lens, z defocusing displacement of a reflective object

ranging from 5× to 100×. The structure of the HC-PBF used for the experiment is shown in Fig. 3.27.

The axial response was measured and the experimental results with HC-PBF and 10×, 20×, and 60× focusing lenses are shown in Fig. 3.28. A solid state laser with a wavelength of 532 nm, an output power of 25 mW, and a beam diameter of 0.7 mm was used along with a beamsplitting cube. A dielectric total-reflection mirror was used as the object. The output power was in the microwatt range revealing very low coupling efficiency, which was mainly due to the transmission loss from the particular HC-PBF sample we used for the experiment.

The factory-specified transmission loss is <1 dB/m; however, the unit used has > 3 dB/m loss. FWHM were 14.7, 6.9, and 2.1 μm for the 10×, 20×, and 60× OL3's, respectively. Using the values of NA of the objective lenses as summarized in Table 3.2, the ratio of FWHM for confocal systems using 10×, 20×, and 60× lenses is 16:6.25:1.38, which is in close agreement with the ratio of the experimental data, 14.7:6.9:2.1, within experimental errors.

Figure 3.29 shows the experimental results with HCF and He–Ne laser with a wavelength of 632.8 nm, output power of 18.5 mW, and beam diameter of 1.8 mm which was used along with a beamsplitting cube.

As we expected from a larger core diameter, the confocal signal showed intensity as high as 2 mW. FWHM were 55 and 16 μm for 40× and 60×, respectively. Results from HCF showed a relatively large depth resolution, which is not suitable for high-resolution confocal microscopy; however, we expect normalized defocus ($\sim z/\lambda$) remains the same when defocusing displacement (z) and wavelength (λ) increase simultaneously, leaving the shape (i.e., FWHM) of the output signal the same.

Another advantage of confocal microscopy using hollow-core fiber is very a low Fresnel back-reflection at the air-fiber interface. We could not fully utilize this advantage at this time due to the fact that fiber couplers made with hollow-core or

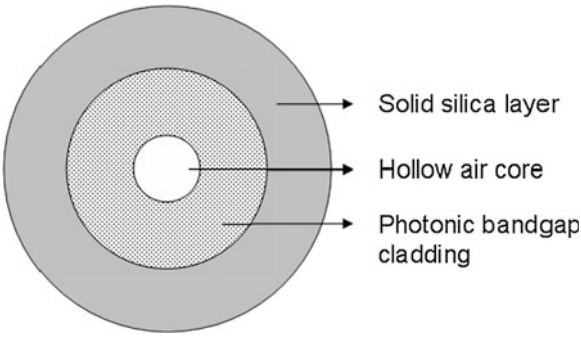

Fig. 3.27 Schematic of cross-section of HC-PBF

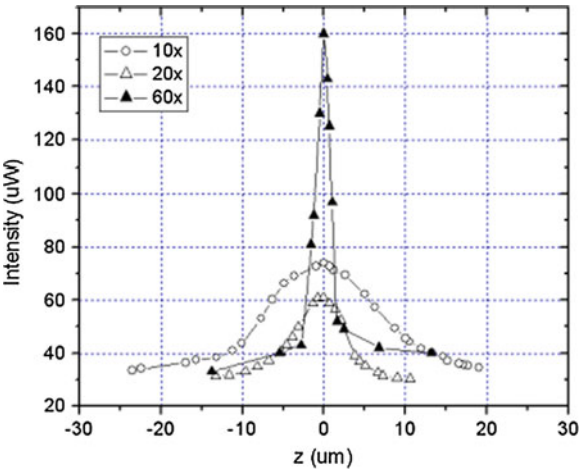

Fig. 3.28 Experimental axial responses obtained with 1 m of HC-PBF, OL1 = 20×, OL2 = 10×, and different OL3's marked as legends in the graph

Table 3.2 Specifications of objective lenses used for the experiment

	NNA	Focal length, f (mm)	WD (mm)
5×	0.10	25.4	13.0
10×	0.25	16.5	5.5
20×	0.40	9.0	1.7
40×	0.65	4.5	0.6
60×	0.85	2.9	0.3

photonic bandgap fibers have not been developed yet; thus, a single fiber configuration using a beamsplitting cube—another source of signal decrease from back-reflection—was adapted. Total back-reflection through the system was measured at 30 μW when HC-PBF was used and 200 μW when a solid-core single-mode fiber was used. That is an 85 % reduction of background level obtainable without anti-reflection coating or angle cleaving of the fiber tips. To reduce the back-reflection at the surfaces of the beamsplitting cube, we

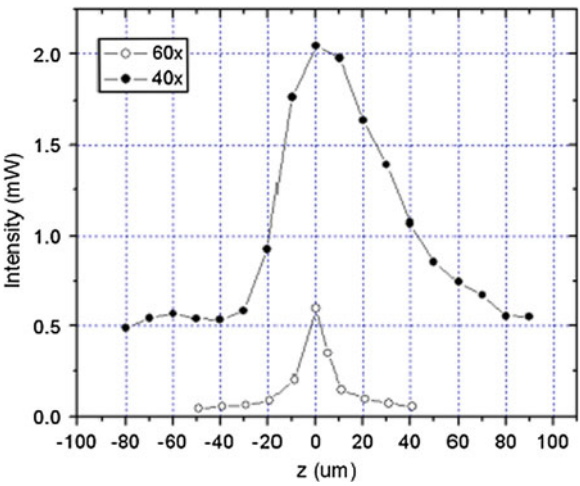

Fig. 3.29 Experimental axial responses obtained with 1 m of HCF, OL1 = 5×, OL2 = 10×, and OL3 was 40× and 60×

investigated a mirror with a hole. A small hole with diameter d1 was drilled in 45 degree through the mirror, where d_1 is the beam diameter of the laser (Fig. 3.30).

When the confocal signal coming out from the fiber is collimated through the lens, OL1, it will have a larger diameter than the input laser according to the NA of fiber and working distance of OL1. For a fixed value of NA of a fiber, the larger the WD, the larger the signal reflected at the mirror back into the power meter. For a fixed d1 and a given OL1, however, a larger WD means a longer focal length, which will result in a larger spot size (d_3) of the input laser at the fiber tip and will lead to poor coupling efficiency. Using a 532 nm solid state laser with $d_1 = 0.7$ mm, the focused spot size was calculated using the formula $w = 4\lambda f/(3\pi w_0)$ to be 21.3, 11.6, and 3.74 μm for 10×, 20×, 60× lenses, respectively.

Although the spot size is larger than the core diameter (d_0) of HC-PBF, 20× showed the best input coupling due to a smaller NA of 0.40 than that of 60× lens. The beam-splitting efficiency, $\eta_1 = (1 - d_1^2/d_2^2)$, was not physically meaningful for 20× as d_2 is smaller than d_1. Note that d_2 can be calculated from the relation, $d_2 = 2WD\tan\alpha$. The proper choice of d1 and OL1 is needed for a given wavelength and fiber for better total efficiency. Assuming a simple coupling efficiency as $\eta_2 = d_0^2/d_3^2$, total efficiency ($\eta_1 \times \eta_2$) was calculated and plotted in Fig. 3.31 for different lasers, (a) for 5 μm core fiber and (b) for 8 μm core fiber.

The optimal result is predicted from the calculation to be when a lens with f = 4 mm (40×) was used in corporation with solid state laser, or when a lens with f = 7.5 mm (20×) was used with He–Ne leaser; however, $d_2 < d_1$ for both lenses, thus 10× lens was used for the measurement. The background level was as low as 15 μW compared to 30 μW when a beam-splitting cube was used. Figure 3.31b when compared to Fig. 3.31a shows that the efficiency can be improved when a fiber with a larger core diameter is used. Note that most of mid-IR hollow-core fibers have >50 μm core diameter.

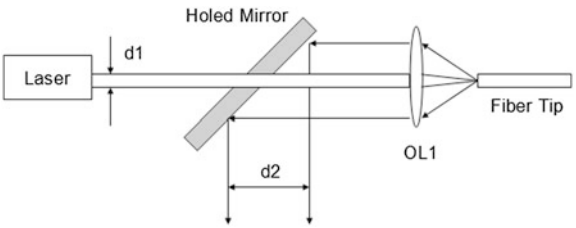

Fig. 3.30 Detailed schematic of the beam-splitting section using a mirror with a hole. *WD* working distance of OL1, d_1 diameter of laser beam, d_2 diameter of collimated output

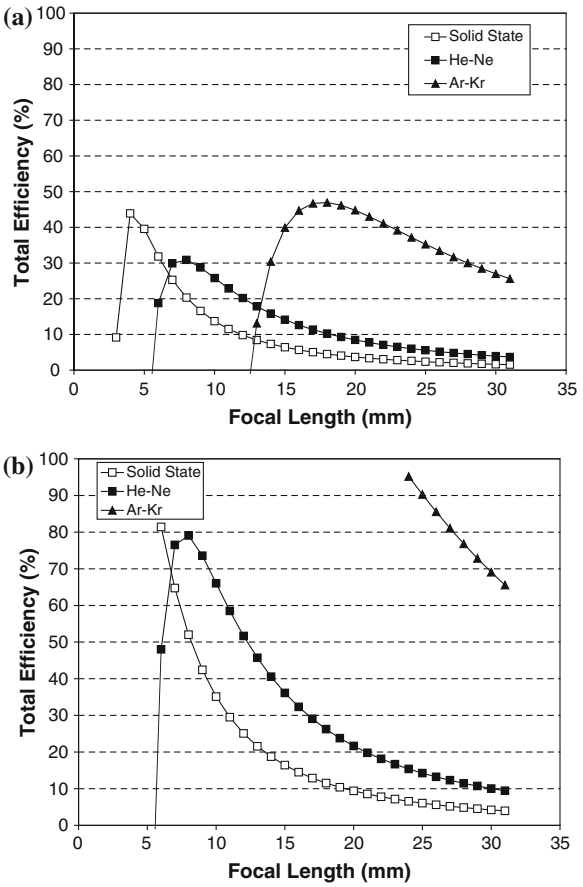

Fig. 3.31 Theoretical calculation of total detection efficiency for different lasers. Beam diameter is 0.7, 1.3, 3.0 mm, and the wavelength is 532, 632.8, 514.5 nm for the solid state laser, He–Ne laser, Ar-Kr laser, respectively. The fiber is set to have core diameter of **a** 5 μm and **b** 8 μm. NA is assumed to be 0.12 for both cases

Figure 3.32 shows the experimental results of the axial resolution measurement using the designed holed-mirror type of beamsplitter. There is a hole with a 0.7 mm diameter, at an angle of 45 degrees to the mirror surface and the input laser passes through the hole before being coupled to the HC-PBF. For precise measurement of the axial resolution, a nano-positioning piezo-electric translational

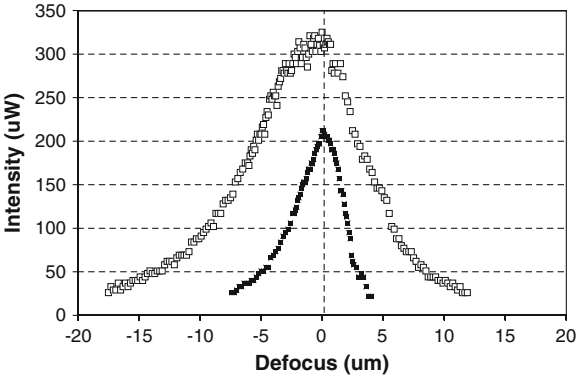

Fig. 3.32 Experimental axial responses obtained with 50 cm of HC-PBF, OL1 = 10×, OL2 = 10×. OL3 was 60× for solid squares and 40× for open squares

stage was used to move the object. The precision of the movement is as low as ~100 nm. The actual movement step used for the experiment was 200 nm. Using 50 cm of HC-PBF, OL1 = 10×, OL2 = 10×, FWHM was measured to be 4.70 μm for 60× OL3, and 10.4 μm for 40× OL3. Reduction of the background level by usage of a holed-mirror beamsplitter was 50 % lower than that of a cubic beamsplitter.

References

1. Mertz J (2009) Introduction to optical microscopy. Roberts and Company publishers, Colorado
2. Pawley JB (ed) (1995) Handbook of biological confocal microscopy. Plenum Press, New York
3. Slayter EM, Slayter HS (1992) Light and electron microscopy. Cambridge University Perss, Cambridge
. Born M, Wolf E (1999) Principles of optics, 7th edn. Cambridge University Press, Cambridge
5. Minsky M (1961) US Patent 3,013,467
6. Wilson T, Sheppard CJR (1984) Theory and practice of scanning optical microscopy. Academic Press, London
7. Lukosz W (1996) Optical systems with resolving powers exceeding classical limit, j. Opt. Soc Am 56:1463
8. Wilson T, Carlini AR (1987) Size of the detector in confocal imaging systems. Opt Lett 12(4):227–229
9. Nakano A (2002) Spinning-disk confocal microscopy: a cutting-edge tool for imaging of membrane traffic. Cell Structure and Function 27(5):349–355
10. Verveer PJ, Hanley QS, Verbeek PW, Van Vliet LJ, Jovin TM (1998) Theory of confocal fluorescence imaging in the programmable array microscope. J Microscopy 189:192–198
11. Dabbs TP, Glass M (1992) Fiber-optic confocal microscope: FOCON. Appl Opt 31:3030–3035
12. Sheppard CJR, Gu M, Roy M (1992) Signal-to-noise ratio in confocal microscope systems. J Microsc 168:209–218
13. Sharma U, Chen G, Kang JU, Ilev I, Waynant RW (2005) Fiber optic confocal laser Doppler velocimeter using an all-fiber laser source for high resolution measurements. Opt Expr 13(16):6250–6258

14. Lee CM, Engelbrecht CJ, Soper TD, Helmchen F, Seibel EJ (2010) Scanning fiber endoscopy with highly flexible 1mm catheterscopes for wide-field, full-color imaging. J Biophotonics 3(5–6):385–407
15. Carlsson K, Aslund N (1987) Confocal imaging for 3-D digital microscopy. Appl Opt 26:3232–3238
16. Hamilton DK, Wilson T (1982) Surface profile measurement using the confocal microscope. J Appl Phys 53:5320–5322
17. Wilson T (1990) Confocal microscopy academic. San Diego, Calif
18. Bae JH, Kim KH, Hong MH, Gim CH, Jhe W (2000) High-resolution confocal detection of nanometric displacement by use of a 2 × 1 optical fiber coupler. Opt Lett 25:1696–1698
19. Knight JC (2003) Photonic crystal fibers. Nature 424:847–851
20. Kim DH, Kang JU (2004) Sagnac loop interferometer based on polarization maintaining photonic crystal fiber with reduced temperature sensitivity. Opt Express 12:4490–4495
21. Shephard JD, MacPherson WN, Maier RRJ, Jones JDC, Hand DP, Mohebbi M, George AK, Roberts PJ, Knight JC (2005) Single-mode mid-IR guidance in a hollow-core photonic crystal fiber. Opt Expr 13:7139–7144
22. Kim DH, Kang JU, Ilev IK (2007) An advanced confocal microscope using a single hollow-core photonic bandgap fiber design. Electonic Letters 43(11):608–609

Chapter 4
Optical Fiber Gratings for Mechanical and Bio-sensing

Young-Geun Han

4.1 Theoretical Analysis of Fiber Gratings

In general, optical fiber gratings can be classified as either fiber Bragg gratings (FBGs) or long-period fiber gratings (LPGs) depending on whether the periodic variation in the refractive index is ranged in the submicron or in the hundreds of microns scale [1, 2]. Figure 4.1 shows the operating principle of FBGs and LPGs. The periodic modulation in the refractive index induce, in the case of FBGs, mode coupling between two counter-propagating modes, whereas in the case of LPGs, it causes coupling between core and cladding modes. FBGs can be used to perform a wide range of functions, such as reflection, filtering, and sensing. Depending on the configuration of their physical structures such as index profile, grating period and tilt, FBGs can be categorized as either uniform FBGs, apodized FBGs, blazed FBGs, or chirped FBGs. Theoretical analysis of fiber gratings can be readily obtained by solving Maxwell's equation and considering small index perturbation [3–5].

In order to analyze optical characteristics of LPGs and FBGs, let's drive the coupled mode equation starting from Maxwell's equation. Assume that each field can be decomposed into spatial and temporal terms as:

$$\begin{aligned} \vec{E} &= \vec{E}(\vec{r},t) = \vec{E}(\vec{r})e^{-i\omega t} \\ \vec{H} &= \vec{H}(\vec{r},t) = \vec{H}(\vec{r})e^{-i\omega t} \end{aligned} \tag{4.1}$$

Assuming that the medium is linear, homogeneous and isotropic, ε μ, σ are independent of $\vec{H}$, $\vec{E}$, space or time and direction, the Maxwell's equation for a monochromatic wave can be written as:

Y.-G. Han (✉)
Department of Physics, Hanyang University, 17 Haengdang-dong, Seongdong-gu, Seoul 133-791, Korea
e-mail: yghan@hanyang.ac.kr

J. U. Kang (ed.), *Fiber Optic Sensing and Imaging*,
DOI: 10.1007/978-1-4614-7482-1_4,

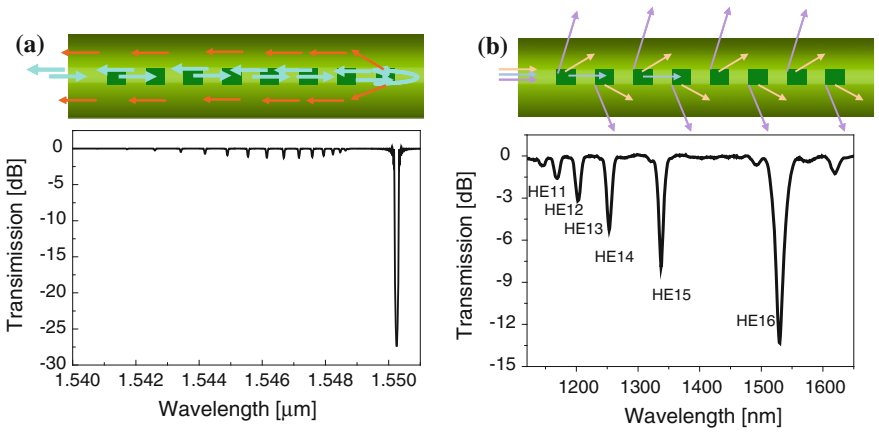

Fig. 4.1 Operating principle of (**a**) fiber Bragg gratings and (**b**) long-period fiber gratings

$$\begin{aligned} \nabla \times \vec{H}(\vec{r}) &= -i\omega\varepsilon\vec{E}(\vec{r}) \\ \nabla \times \vec{E}(\vec{r}) &= i\omega\mu\vec{H}(\vec{r}) \end{aligned} \tag{4.2}$$

$$\begin{aligned} \nabla \times \vec{H}(\vec{r}) &= -i\omega\varepsilon\vec{E}(\vec{r}) \\ \nabla \times \vec{E}(\vec{r}) &= i\omega\mu\vec{H}(\vec{r}) \end{aligned} \tag{4.3}$$

Assuming that the wave propagates in z-direction, and the special term of fields can be decomposed into transverse and longitudinal components as

$$\begin{aligned} \vec{E}(\vec{r}) &= \vec{E}_t(\vec{r}) + \vec{E}_z(\vec{r}) \\ \vec{H}(\vec{r}) &= \vec{H}_t(\vec{r}) + \vec{H}_z(\vec{r}) \end{aligned} \tag{4.4}$$

Then the Maxwell's equation can be decomposed into transverse and longitudinal components.

Transverse components:

$$\begin{aligned} \nabla_t \times \vec{H}_z(\vec{r}) + \nabla_z \times \vec{H}_t(\vec{r}) &= -i\omega\varepsilon\vec{E}_t(\vec{r}) \\ \nabla_t \times \vec{E}_z(\vec{r}) + \nabla_z \times \vec{E}_t(\vec{r}) &= i\omega\mu\vec{H}_t(\vec{r}) \end{aligned} \tag{4.5}$$

Longitudinal components:

$$\begin{aligned} \nabla_t \times \vec{H}_t(\vec{r}) &= -i\omega\varepsilon\vec{E}_z(\vec{r}) \\ \nabla_t \times \vec{E}_t(\vec{r}) &= i\omega\mu\vec{H}_z(\vec{r}) \end{aligned} \tag{4.6}$$

The transverse mode equation can be obtained by applying $\hat{e}_z\times$ on Eq. (4.5) which results in:

$$\begin{aligned}\frac{\partial \vec{H}_z(\vec{r})}{\partial z} - i\omega\varepsilon\hat{e}_z \times \vec{E}_t(\vec{r}) &= \nabla_t \vec{H}_z(\vec{r}) \\ \frac{\partial \vec{E}_z(\vec{r})}{\partial z} + i\omega\mu\hat{e}_z \times \vec{H}_t(\vec{r}) &= \nabla_t \vec{E}_z(\vec{r})\end{aligned} \tag{4.7}$$

The longitudinal mode equations from Eq. (4.6) can be obtained as

$$\begin{aligned}\vec{E}_z(\vec{r}) &= \tfrac{i}{\omega\varepsilon}\nabla_t \times \vec{H}_t(\vec{r}) \\ \vec{H}_z(\vec{r}) &= \tfrac{-i}{\omega\mu}\nabla_t \times \vec{E}_t(\vec{r})\end{aligned} \tag{4.8}$$

By substituting Eq. (4.8) into Eq. (4.5), the transverse mode equation can be expressed by

$$\begin{aligned}\nabla_t \times \frac{-i}{\omega\mu}\nabla_t \times \vec{E}_t(\vec{r}) + \nabla_z \times \vec{H}_t(\vec{r}) &= -i\omega\varepsilon\vec{E}_t(\vec{r}) \\ \nabla_t \times \frac{i}{\omega\varepsilon}\nabla_t \times \vec{H}_t(\vec{r}) + \nabla_z \times \vec{E}_t(\vec{r}) &= i\omega\mu\vec{H}_t(\vec{r})\end{aligned} \tag{4.9}$$

For an ideal waveguide, the mode profile can be written as:

$$\begin{aligned}\vec{E}(\vec{r},t) &= \vec{E}_l(x,y)e^{i\beta_l z - i\omega t} \\ \vec{H}(\vec{r},t) &= \vec{H}_l(x,y)e^{i\beta_l z - i\omega t},\end{aligned} \tag{4.10}$$

where β_l is the propagating constant of the lth order mode.

Since $\mu = \mu_0, \varepsilon = \varepsilon_0 n_0^2(x,y)$, where n_0 is a refractive index in an ideal waveguide, then, Eq. (4.5) in the ideal mode case can be modified as:

$$\begin{aligned}\frac{-i}{\omega\mu_0}\nabla_t \times \nabla_t \times \vec{E}_{l,t}(x,y) + i\beta_l\hat{e}_z \times \vec{H}_{l,t}(\vec{r}) &= -i\omega\varepsilon_0 n_0^2(x,y)\vec{E}_{l,t}(x,y) \\ \frac{i}{\omega\varepsilon_0}\nabla_t \times \left(\frac{1}{n_0^2(x,y)}\nabla_t \times \vec{H}_{l,t}(\vec{r})\right) + i\beta_l\hat{e}_z \times \vec{E}_{l,t}(\vec{r}) &= i\omega\mu_0\vec{H}_{l,t}(\vec{r})\end{aligned} \tag{4.11}$$

where $\vec{E}_{l,t}$ and $\vec{H}_{l,t}$ are the transverse field of the lth order ideal mode for electric and magnetic waves in an ideal waveguide. β_l is the propagating constant of the lth order mode.

For the perturbed waveguide, we can expand the perturbed field with a linear summation of ideal normal modes such as:

$$\begin{aligned}\vec{E}_t(\vec{r}) &= \sum_{l=0}^{\infty} c_l e^{i\beta_l z}\vec{E}_{l,t}(x,y) = \sum_{l=0}^{\infty} a_l(z)\vec{E}_{l,t}(x,y) \\ \vec{H}_t(\vec{r}) &= \sum_{l=0}^{\infty} g_l e^{i\beta_l z}\vec{H}_{l,t}(x,y) = \sum_{l=0}^{\infty} b_l(z)\vec{H}_{l,t}(x,y)\end{aligned} \tag{4.12}$$

where the constant c_l and g_l are the slowly varying modal amplitudes that include phase term for electric and magnetic waves, respectively. $a_l(z)$ and $b_l(z)$ are the rapidly varying modal amplitudes that include phase term for electric and magnetic waves, respectively. Since $\mu = \mu_0, \varepsilon = \varepsilon_0 n^2(\vec{r})$, Eq. (4.9) can be modified by

inserting Eq. (4.12) into Eq. (4.9). From that we obtain the governing equations for the perturbed waveguide as:

$$\frac{-i}{\omega\mu_0}\sum a_l(z)\nabla_t\times\nabla_t\times\vec{E}_{l,t}(x,y)+\sum_{l=0}^{\infty}\frac{\partial b_l(z)}{\partial z}\hat{e}_z\times\vec{H}_{l,t}(x,y)=-i\omega\varepsilon_0 n^2(\vec{r})\sum_{l=0}^{\infty}a_l(z)\vec{E}_{l,t}(x,y)$$

$$\frac{i}{\omega\varepsilon_0}\sum b_l(z)\nabla_t\times\left(\frac{i}{n_0^2(\vec{r})}\nabla_t\times\vec{H}_{l,t}(x,y)\right)+\sum_{l=0}^{\infty}\frac{\partial a_l(z)}{\partial z}\hat{e}_z\times\vec{E}_t(\vec{r})=i\omega\mu_0\sum_{l=0}^{\infty}b_l(z)\vec{H}_{l,t}(x,y) \tag{4.13}$$

By substituting appropriate equations in the ideal waveguide into Eq. (4.13), we have

$$\begin{aligned}&\sum a_l(z)\left(-i\beta_l\hat{e}_z\times\vec{H}_{l,t}(x,y)-i\omega\varepsilon_0 n_0^2(x,y)\vec{E}_{l,t}(x,y)\right)\vec{E}_{l,t}(x,y)+\sum_{l=0}^{\infty}\frac{\partial b_l(z)}{\partial z}\hat{e}_z\times\vec{H}_{l,t}(x,y)\\&=-i\omega\varepsilon_0 n^2(\vec{r})\sum_{l=0}^{\infty}a_l(z)\vec{E}_{l,t}(x,y)\\&\frac{i}{\omega\varepsilon_0}\sum b_l(z)\nabla_t\times\left(\left(\frac{i}{n_0^2(x,y)}-\frac{i}{n^2(\vec{r})}\right)\nabla_t\times\vec{H}_{l,t}(x,y)\right)+\sum_{l=0}^{\infty}\left(i\beta_l b_l(z)-\frac{\partial a_l(z)}{\partial z}\right)\hat{e}_z\times\vec{E}_t(\vec{r})=0\end{aligned} \tag{4.14}$$

Rearranging Eq. (4.14) results in:

$$\begin{aligned}&\sum_{l=0}^{\infty}\left(\frac{\partial b_l(z)}{\partial z}-i\beta_l a(z)\right)\hat{e}_z\times\vec{H}_{l,t}(x,y)+i\omega\varepsilon_0\left(n^2(\vec{r})-n_0^2(x,y)a_l(z)\right)\vec{E}_{l,t}(x,y)=0\\&\sum_{l=0}^{\infty}\left(i\beta_l b_l(z)-\frac{\partial a_l(z)}{\partial z}\right)\hat{e}_z\times\vec{E}_t(\vec{r})+\tfrac{i}{\omega\varepsilon_0}\sum b_l(z)\nabla_t\times\left(\left(\frac{i}{n_0^2(x,y)}-\frac{i}{n^2(\vec{r})}\right)\nabla_t\times\vec{H}_{l,t}(x,y)\right)=0\end{aligned} \tag{4.15}$$

Using the equations for normal modes and using the normalization and orthogonalization conditions such as,

$$\int\hat{e}_z\cdot(\vec{E}_{m,t}(x,y)\times\vec{H}^*_{l,t}(x,y))dA=2\frac{\beta_l^*}{|\beta_l|}P\delta_{m,l}, \tag{4.16}$$

Where $\delta_{m,l}=1$ for $m=l$ and $\delta_{m,l}=0$ for $m\neq l$

Equation (4.15) can be modified by multiplying the terms with either $\int\vec{E}_{m,t}(x,y)\cdot$ or $\int\vec{H}_{m,t}(x,y)\cdot$

$$\begin{aligned}&\int\vec{E}_{m,t}(x,y)\cdot\left(\sum_{l=0}^{\infty}\left(\frac{\partial b_l(z)}{\partial z}-i\beta_l a(z)\right)\hat{e}_z\times\vec{H}_{l,t}(x,y)+i\omega\varepsilon_0\left(n^2(\vec{r})-n_0^2(x,y)a_l(z)\right)\vec{E}_{l,t}(x,y)=0\right)\\&\int\vec{H}_{m,t}(x,y)\cdot\left(\sum_{l=0}^{\infty}\left(\frac{\partial a_l(z)}{\partial z}-i\beta_l b_l(z)\right)\hat{e}_z\times\vec{E}_{l,t}(\vec{r})+\tfrac{i}{\omega\varepsilon_0}\sum b_l(z)\nabla_t\times\left(\left(\tfrac{i}{n^2(\vec{r})}-\tfrac{i}{n_0^2(x,y)}\right)\nabla_t\times\vec{H}_{l,t}(x,y)\right)=0\right)\end{aligned} \tag{4.17}$$

Then Eq. (4.17) can be modified by using Eq. (4.16)

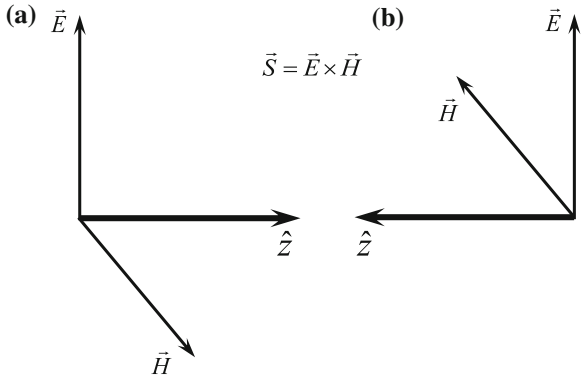

Fig. 4.2 Direction of pointing vector ($\vec{S} = \vec{E} \times \vec{H}$) propagating (**a**) forward z-direction or (**b**) backward -z-direction

$$\begin{aligned} \frac{\partial b_m(z)}{\partial z} - i\beta_m a(z) &= 2\sum_{l=0}^{\infty} a_l(z) K_{m,l}(z) \\ \frac{\partial a_m(z)}{\partial z} - i\beta_m b_m(z) &= 2\sum_{l=0}^{\infty} b_l(z) C_{m,l}(z) \end{aligned} \tag{4.18}$$

Where

$$\begin{aligned} K_{m,l}(z) &= \frac{i\omega\varepsilon_0}{4P}\frac{|\beta_m|}{\beta_m}\int \left(n^2(\vec{r}) - n_0^2(x,y)\right)\vec{E}^*_{m,t}(x,y)\cdot\vec{E}_{l,t}(x,y)dA \\ C_{m,l}(z) &= \frac{i\omega\varepsilon_0}{4P}\frac{|\beta_m|}{\beta_m^*}\int \frac{n_0^2(x,y)}{n^2(\vec{r})}\left(n^2(\vec{r}) - n_0^2(x,y)\right)\vec{E}^*_{m,z}(x,y)\cdot\vec{E}_{l,z}(x,y)dA \end{aligned} \tag{4.19}$$

The solutions for the mode coupling can be expressed as:

$$\begin{aligned} a_m(z) &= a_m^{(+)}(z) + a_m^{(-)}(z) = c_m^{(+)}(z)e^{i\beta_m z} + c_m^{(-)}(z)e^{-i\beta_m z} \\ b_m(z) &= a_m^{(+)}(z) - a_m^{(-)}(z) = c_m^{(+)}(z)e^{i\beta_m z} - c_m^{(-)}(z)e^{-i\beta_m z} \end{aligned} \tag{4.20}$$

Where $c_m^{(\pm)}$ is the amplitude for the waves propagating in forward (+) or or backward (−) direction. Here, it is evident that the relative direction of the magnetic field over the electric field of a mode is reversed when the propagation direction is reversed in an ideal waveguide as seen in Fig. 4.2.

By substituting Eq. (4.20) into Eq. (4.18), we obtain

$$\begin{aligned} \frac{\partial}{\partial z}\left(a_m^{(+)}(z) + a_m^{(-)}(z)\right) - i\beta_m\left(a_m^{(+)}(z) + a_m^{(-)}(z)\right) &= 2\sum_{l=0}^{\infty}\left(a_l^{(+)}(z) + a_l^{(-)}(z)\right)K_{m,l}(z) \\ \frac{\partial}{\partial z}\left(a_m^{(+)}(z) - a_m^{(-)}(z)\right) - i\beta_m\left(a_m^{(+)}(z) - a_m^{(-)}(z)\right) &= 2\sum_{l=0}^{\infty}\left(a_l^{(+)}(z) - a_l^{(-)}(z)\right)C_{m,l}(z) \end{aligned} \tag{4.21}$$

By taking the sum or the difference in Eq. (4.21), we obtain

$$\frac{\partial a_m^{(+)}(z)}{\partial z} - i\beta_m a_m^{(+)}(z) = \sum_{l=0}^{\infty} a_l^{(+)}(z)\big(K_{m,l}(z) + C_{m,l}(z)\big) + \sum_{l=0}^{\infty} a_l^{(-)}(z)\big(K_{m,l}(z) - C_{m,l}(z)\big)$$
$$\frac{\partial a_m^{(-)}(z)}{\partial z} + i\beta_m a_m^{(-)}(z) = \sum_{l=0}^{\infty} a_l^{(+)}(z)\big(-K_{m,l}(z) + C_{m,l}(z)\big) + \sum_{l=0}^{\infty} a_l^{(-)}(z)\big(-K_{m,l}(z) - C_{m,l}(z)\big) \tag{4.22}$$

By using Eq. (4.20), we can get the mode equations for the slowly varying factors $c_m^{(\pm)}$ as

$$\frac{\partial c_m^{(+)}(z)}{\partial z} = \sum_{l=0}^{\infty}\big(K_{m,l}(z) + C_{m,l}(z)\big)c_l^{(+)}(z)e^{i(\beta_l-\beta_m)z} + \sum_{l=0}^{\infty}\big(K_{m,l}(z) - C_{m,l}(z)\big)c_l^{(-)}(z)e^{(-i\beta_l-\beta_m)z}$$
$$\frac{\partial c_m^{(-)}(z)}{\partial z}e^{-i\beta_m z} = \sum_{l=0}^{\infty}\big(-K_{m,l}(z) + C_{m,l}(z)\big)c_l^{(+)}(z)e^{i(\beta_l+\beta_m)z} + \sum_{l=0}^{\infty}\big(-K_{m,l}(z) - C_{m,l}(z)\big)c_l^{(-)}(z)e^{(-i\beta_l+i\beta_m)z} \tag{4.23}$$

To simplify the expression, we define the coupling coefficient as:

$$\chi_{m,l}^{(p,q)}(z) = pK_{m,l}(z) + qC_{m,l}(z)$$
$$= \frac{i\omega\varepsilon_0}{4P}\int\left(\left(n^2(\vec{r}) - n_0^2(x,y)\right)\times\left(p\frac{|\beta_m|}{\beta_m}\vec{E}_{m,t}^*(x,y)\cdot\vec{E}_{l,t}(x,y) + q\frac{|\beta_m|}{\beta_m^*}\frac{n_0^2(x,y)}{n^2(\vec{r})}\vec{E}_{m,z}^*(x,y)\cdot\vec{E}_{l,z}(x,y)\right)\right)da \tag{4.24}$$

where $p, q = \pm$. Then we will finally obtain the coupled mode equation for $c_m^{(\pm)}$ as

$$\frac{\partial c_m^{(+)}(z)}{\partial z} = \sum_{l=0}^{\infty}\chi_{m,l}^{(+,+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m)z} + \sum_{l=0}^{\infty}\chi_{m,l}^{(+,-)}c_l^{(-)}(z)e^{(-i\beta_l-\beta_m)z}$$
$$\frac{\partial c_m^{(-)}(z)}{\partial z} = \sum_{l=0}^{\infty}\chi_{m,l}^{(-,+)}c_l^{(+)}(z)e^{i(\beta_l+\beta_m)z} + \sum_{l=0}^{\infty}\chi_{m,l}^{(-,-)}c_l^{(-)}(z)e^{(-i\beta_l+i\beta_m)z} \tag{4.25}$$

From Eq. (4.25), we can obtain the coupling mode equation for FBGs and LPGs by considering the refractive index profile of the perturbed waveguide $n(\vec{r})$as:

$$n(\vec{r}) = n_0(x,y) + \Delta n(\vec{r}) \tag{4.26}$$

If the perturbation is small such that $\Delta n \ll n_0$, we can approximate Eq. (4.26) as since $\Delta n^2 \to 0$:

$$n^2(\vec{r}) \approx n_0^2(x,y) + 2n_0(x,y)\Delta n(\vec{r}). \tag{4.27}$$

In conventional waveguides, most of lower guided modes have small propagating angles. Therefore the amplitude of the longitudinal field in the z-direction is much smaller than that of the transverse field in the x- and y-direction. Therefore, we can conclude:

$$\begin{aligned} \left|\vec{E}_t\right| &\gg \left|\vec{E}_z\right| \\ \left|\vec{H}_t\right| &\gg \left|\vec{H}_z\right| \\ \left|K_{m,l}(z)\right| &\gg \left|C_{m,l}(z)\right| \end{aligned} \tag{4.28}$$

Then, the coupling constant can be modified as

$$\begin{aligned} \chi_{m,l}^{(p,q)}(z) &= pK_{m,l}(z) + qC_{m,l}(z) \approx pK_{m,l}(z) \\ &= p\frac{i\omega\varepsilon_0}{4P}\iint 2n_0(x,y)\Delta n(\vec{r})\vec{E}_{m,t}^{*}(x,y)\cdot\vec{E}_{l,t}(x,y)da \end{aligned} \tag{4.29}$$

By considering that the index perturbation is uniformly induced across the cross sectional area of the waveguide in the z-direction but restricted in a localized area, Eq. (4.29) can be written as:

$$\begin{aligned} K_{m,l}(z) &= i2\Delta n(z)\Omega_{m,l} \\ \Omega_{m,l} &= \frac{i\omega\varepsilon_0}{4P}\iint 2n_0(x,\,y)\Delta n(\overrightarrow{r})\,E_{m,t}^{*}(x,\,y)\cdot\ \overrightarrow{E}_{l,t}(x,\,y)da \end{aligned} \tag{4.30}$$

where $\Omega_{m,l} = \Omega_{m,l}^{*}$. Therefore, $\Omega_{m,l}$ is real in general.

For the case of fiber gratings, the index perturbation can be written as

$$\Delta n(z) = \Delta n_{dc}(z) + \Delta n_{dc}(z)Cos(Kz + \Phi(z)) \tag{4.31}$$

where $\Delta n_{dc}(z)$ is a slowly varying DC part ($\sim$ average index variation) and $\Delta n_{ac}(z)$ is the rapidly varying AC part with a small phase variation, $\Phi(z)$.

4.1.1 LPG: Co-Directional Coupling with Sinusoidal Perturbation

From Eqs. (4.25) and (4.30), the coupled equations for the co-directional case can be written as:

$$\frac{\partial c_m^{(+)}(z)}{\partial z} = \sum_{l=0}^{\infty}\chi_{ml}^{(+,+)}(z)e^{i(\beta_l-\beta_m)z} \approx 2i\Delta n(z)\sum_{l=0}^{\infty}\Omega_{ml}^{(+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m)z} \tag{4.32}$$

By inserting the periodic variation of Eq. (4.31) and decomposing it with self coupling terms and cross coupling terms, we obtain:

$$\begin{aligned} \frac{\partial c_m^{(+)}(z)}{\partial z} =& 2i[\Delta n_{dc}(z) + \Delta n_{ac}(z)\cos(Kz + \Phi(z))]\Omega_{m,m}^{(+)}c_m^{(+)}(z) \\ &+2i[\Delta n_{dc}(z) + \Delta n_{ac}(z)\cos(Kz + \Phi(z))]\sum_{l\neq m}\Omega_{ml}^{(+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m)z} \end{aligned}, \tag{4.33}$$

The cosine function can expressed in terms of exponential functions:

$$\cos(Kz + \Phi(z)) = \frac{1}{2}\left(e^{i\phi(z)}e^{iKz} + e^{-i\phi(z)}e^{-iKz}\right) \tag{4.34}$$

Substituting Eq. (4.34) into Eq. (4.35) gives

$$\begin{aligned}\frac{\partial c_m^{(+)}(z)}{\partial z} &= 2i\Delta n_{dc}(z)\Omega_{m,m}^{(+)}c_m^{(+)}(z) \\ &+ i\Delta n_{ac}(z)e^{i\phi(z)}\Omega_{m,m}^{(+)}c_m^{(+)}(z)e^{iKz} \\ &+ i\Delta n_{ac}(z)e^{-i\phi(z)}\Omega_{m,m}^{(+)}c_m^{(+)}(z)e^{-iKz} \\ &+ 2i\Delta n_{dc}(z)\sum_{l\neq m}\Omega_{m,l}^{(+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m)z} \\ &+ i\Delta n_{ac}(z)e^{i\phi(z)}\sum_{l\neq m}\Omega_{m,l}^{(+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m+K)z} \\ &+ i\Delta n_{ac}(z)e^{-i\phi(z)}\sum_{l\neq m}\Omega_{m,l}^{(+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m-K)z}\end{aligned} \tag{4.35}$$

Since each c_m is the slowly varying modal amplitude, the effect made by the rapidly varying (or oscillating) terms in the above equation can be neglected in most cases. Also for simplicity let's only consider the mode coupling between the core and the co-propagating cladding modes in a single-mode fiber, therefore:

$$c_m^{(+)} = c_{core}, c_{cadd}^{v}(for\ v = 1, 2, 3, 4, \ldots) \tag{4.36}$$

The effective index of the core mode is greater than that of the cladding modes ($\beta_{core} > \beta_{clad}^{v}$ for any cladding of mode order, v). The grating contains many periods within the length of the grating, d. Then, we can get $Kd \gg 1$, with $K = 2\pi/\Lambda$, Λ is the grating period. The mode spacing of the cladding modes is wide enough or/and the length of the grating is long enough to satisfy:

$$\left(\beta_{clad}^{v} - \beta_{clad}^{v+1}\right)d \gg 1 \tag{4.37}$$

The coupling coefficient between cladding modes is extremely small. The terms with the following phase factors in Eq. (4.35) can be sufficiently small to be neglected;

$$e^{iKz},\ e^{-iKz},\ e^{i\left(\beta_{core}-\beta_{clad}^{v}\right)z} \tag{4.38}$$

Therefore, only three terms (one self term and two sets of cross terms) are remained;

$$\begin{aligned}\frac{\partial c_m^{(+)}(z)}{\partial z} &\approx 2i\Delta n_{dc}(z)\Omega_{mm}^{(+)}c_m^{(+)}(z) \\ &+ i\Delta n_{ac}(z)se^{i\Phi(z)}\sum_{l\neq m}\Omega_{ml}^{(+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m+K)z} \\ &+ i\Delta n_{ac}(z)e^{-i\Phi(z)}\sum_{l\neq m}\Omega_{ml}^{(+)}c_l^{(+)}(z)e^{i(\beta_l-\beta_m-K)z},\end{aligned} \tag{4.39}$$

Decomposing Eq. (4.39) (when the left side is the core mode), we can obtain the coupled core mode with a certain mode order (v) as:

$$\begin{aligned}\frac{\partial c_{core}(z)}{\partial z} &\approx 2i\Delta n_{dc}(z)\Omega_{core,core}^{(+)}c_{core}(z) \\ &+ i\Delta n_{ac}(z)e^{i\Phi(z)}\sum_{v}\Omega_{core,v}^{(+)}c_{clad}^{v}(z)e^{i\left(\beta_{clad}^{v}-\beta_{core}+K\right)z} \\ &+ i\Delta n_{ac}(z)e^{-i\Phi(z)}\sum_{v}\Omega_{core,v}^{(+)}c_{clad}^{v}(z)e^{i\left(\beta_{clad}^{v}-\beta_{core}-K\right)z}\end{aligned} \tag{4.40}$$

In the 2-nd term, the phase factor can be vanished only for one cladding mode that satisfies the phase matching condition defined as

$$\Delta\beta \equiv \beta_{core} - \beta_{clad}^{v} - K \rightarrow 0 \tag{4.41}$$

Then we can obtain the expression for the phase matching condition for the LPG as

$$\lambda_p = \Lambda(n_{core} - n_{clad}^{v}) \tag{4.42}$$

Note that for a given waveguide, the order of cladding mode (v) for the mode coupling is determined (or selected) by the amount of K. The phase of the last term of Eq. (4.40) is typically sufficiently small enough to be neglected, therefore Eq. (4.40) can be written as:

$$\frac{\partial c_{core}(z)}{\partial z} \approx 2i\Delta n_{dc}(z)\Omega_{core,core}^{(+)}c_{core}(z) + i\Delta n_{ac}(z)e^{i\Phi(z)}\Omega_{core,v}^{(+)}c_{clad}^{v}(z)e^{-i\Delta\beta z} \tag{4.43}$$

By using the same procedure, the coupling mode equation can be obtained as

$$\frac{\partial c_{clad}^{v}(z)}{\partial z} \approx 2i\Delta n_{dc}(z)\Omega_{v,v}^{(+)}c_{clad}^{v}(z) + i\Delta n_{ac}(z)e^{-i\Phi(z)}\Omega_{v,core}^{(+)}c_{core}(z)e^{i\Delta\beta z}, \tag{4.44}$$

By defining the coupling constant κ as

$$\kappa_{dc}^{co} \equiv 2\Delta n_{dc}\Omega_{core,core}^{(+)} \quad \text{and} \quad \kappa_{dc}^{cl} \equiv 2\Delta n_{dc}\Omega_{v,v}^{(+)}$$

$$\kappa = \kappa_{ac} \equiv \Delta n_{ac}\Omega_{core,v}^{(+)}, \quad \text{Then,} \quad \Delta n_{ac}\Omega_{v,core}^{(+)} = \kappa_{ac}^{*} \equiv \kappa^{*}$$

And re-defining the mode indexing as

$$A = A(z) \equiv c_{core}(z) \quad \text{Core mode amplitude (slowly varying)}$$

$$B = B(z) \equiv c^{v}_{clad}(z) \quad \text{Cladding mode amplitude (slowly varying)}$$

Then, the coupled mode equation becomes

$$A' = i\kappa^{co}_{dc}A + i\kappa_{ac}Be^{-i\Delta\beta z} \tag{4.45}$$

$$B' = i\kappa^{*}_{ac}Ae^{+i\Delta\beta z} + i\kappa^{cl}_{dc}B \tag{4.46}$$

In a matrix form, it can be written as:

$$\begin{bmatrix} A(z)' \\ B(z)' \end{bmatrix} = i\begin{bmatrix} \kappa^{co}_{dc} & \kappa_{ac}e^{-i\Delta\beta z} \\ \kappa^{*}_{ac}e^{+i\Delta\beta z} & \kappa^{cl}_{dc} \end{bmatrix}\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} \tag{4.47}$$

The solution of the coupled-mode equation of two co-propagating modes is then given by

$$\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = e^{i\frac{\beta_{co}+\beta^{v}_{cl}}{2}z}\begin{bmatrix} \left(\cos sz + i\frac{\Delta\beta}{2s}\sin sz\right)e^{i\frac{K}{2}z} & i\frac{\kappa}{s}e^{i\frac{K}{2}z}\sin sz \\ i\frac{\kappa*}{s}e^{-i\frac{K}{2}z}\sin sz & \left(\cos sz - i\frac{\Delta\beta}{2s}\sin sz\right)e^{-i\frac{K}{2}z} \end{bmatrix}\begin{bmatrix} A(0) \\ B(0) \end{bmatrix} \tag{4.48}$$

where s is $\sqrt{\left(\frac{\Delta\beta}{2}\right)^2 + \kappa\kappa*}$.

4.1.2 FPG: Co-Directional Coupling with Sinusoidal Perturbation

A core mode can be coupled to a counter propagating mode by a periodic index perturbation that exists in FBGs. Since single- mode fibers have only one core mode with a mode order l, only $c_1^{(\pm)}$ modes are possibly confined. Thus, from Eq. (4.25) we can obtain:

$$\frac{\partial c_1^{(+)}(z)}{\partial z} = \chi_{11}^{(+,+)}(z)c_1^{(+)}(z)e^{i(\beta_1-\beta_1)z} + \chi_{11}^{(+,-)}(z)c_1^{(-)}(z)e^{i(-\beta_1-\beta_1)z} \tag{4.49}$$

$$\frac{\partial c_m^{(-)}(z)}{\partial z} = \chi_{m1}^{(-,+)}(z)c_1^{(+)}(z)e^{i(\beta_1+\beta_1)z} + \chi_{m1}^{(-,-)}(z)c_1^{(-)}(z)e^{i(-\beta_1+\beta_m)z} \tag{4.50}$$

Since the mode order m becomes 1 in the FBG, Eqs. 4.48 and 4.49 become

$$\frac{\partial c_1^{(+)}(z)}{\partial z} = \chi_{11}^{(+,+)}(z)c_1^{(+)}(z) + \chi_{11}^{(+,-)}(z)c_1^{(-)}(z)e^{-2i\beta_1 z} \tag{4.51}$$

$$\frac{\partial c_1^{(-)}(z)}{\partial z} = \chi_{11}^{(-,+)}(z)c_1^{(+)}(z)e^{+2i\beta_1 z} + \chi_{11}^{(-,-)}(z)c_1^{(-)}(z) \tag{4.52}$$

Using the new coupling coefficient defined in Eq. (4.30) yields

$$\chi_{11}^{(p,q)}(z) \approx i2\Delta n(z)\Omega_{11}^{p} = pi2\Delta n(z)\Omega_{11} \tag{4.53}$$

$$\Omega_{11}^{p} \equiv p\frac{\omega\varepsilon_0}{4P}\frac{|\beta_1|}{\beta_1}\iint_{pert} n_0(x,y)\mathrm{E}_{1t}^{*}(x,y)\cdot\mathrm{E}_{1t}(x,y)dxdy \tag{4.54}$$

$$\equiv p\frac{\omega\varepsilon_0}{4P}\frac{|\beta_1|}{\beta_1}\iint_{pert} n_0(x,y)|\mathrm{E}_{1t}(x,y)|^2 dxdy, \tag{4.55}$$

with $\Omega_{11}^{(-)} = -\Omega_{11}^{(+)}$. The coupling intensity is proportional to the intensity of E field of the core mode. With the sinusoidal index variation of Eq. (4.31), we have:

$$\begin{aligned}\frac{\partial c_1^{(+)}(z)}{\partial z} &= 2i\Delta n_{dc}(z)\Omega_{11}^{(+)}c_1^{(+)}(z) \\ &\quad + i\Delta n_{ac}(z)\Omega_{11}^{(+)}c_1^{(-)}(z)\left[e^{-i\Phi(z)}e^{i(-2\beta_1-\mathrm{K})z} + e^{i\Phi(z)}e^{i(-2\beta_1+\mathrm{K})z}\right]\end{aligned} \tag{4.56}$$

$$\begin{aligned}\frac{\partial c_1^{(-)}(z)}{\partial z} &= i\Delta n_{ac}(z)\Omega_{11}^{(-)}c_1^{(+)}(z)\left[e^{-i\Phi(z)}e^{i(2\beta_1-\mathrm{K})z} + e^{i\Phi(z)}e^{i(2\beta_1+\mathrm{K})z}\right] \\ &\quad + 2i\Delta n_{dc}(z)\Omega_{11}^{(-)}c_1^{(-)}(z)\end{aligned} \tag{4.57}$$

The terms having phases of $e^{\pm i\mathrm{K}z}$ and $e^{\pm i2\beta_1}$ can be sufficiently small to be neglected. The coupling factors of the core mode to other co-directional modes were also removed for a single core fiber. Only the terms having small phase factors dominates the coupling that is $\Delta\boldsymbol{\beta} \equiv 2\boldsymbol{\beta}_{core} - \mathbf{K} \rightarrow \boldsymbol{0}$.

Then we obtain the expression for the phase matching condition for the FBG as

$$\lambda_p = 2n_{core}\Lambda \tag{4.58}$$

For the case of a uniform grating ($\Phi(z) = 0$), we can write the coupled mode equation for the FBG as:

$$\frac{\partial c_1^{(+)}(z)}{\partial z} = 2i\Delta n_{dc}(z)\Omega_{11}^{(+)}c_1^{(+)}(z) + i\Delta n_{ac}(z)\Omega_{11}^{(+)}c_1^{(-)}(z)e^{-i\Delta\beta z} \tag{4.59}$$

$$\frac{\partial c_1^{(-)}(z)}{\partial z} = i\Delta n_{ac}(z)\Omega_{11}^{(-)}c_1^{(+)}(z)e^{+i\Delta\beta z} + 2i\Delta n_{dc}(z)\Omega_{11}^{(-)}c_1^{(-)}(z) \tag{4.60}$$

Here the coupling constant$\kappa$and the mode indexing can be defined as

$$\kappa_{dc}(z) \equiv 2\Delta n_{dc}(z)\Omega_{11}^{(+)}$$

$$\kappa_{ac}(z) \equiv \Delta n_{ac}(z)\Omega_{11}^{(+)}$$

$$A = A(z) \equiv c_1^{(+)}(z)$$

$$B = B(z) \equiv c_1^{(-)}(z)$$

by using the relationship of $\Omega_{11}^{(-)} = -\Omega_{11}^{(+)}$, we have the coupled mode equations for two counter-propagating modes as

$$\frac{\partial A(z)}{\partial z} = i\kappa_{dc}(z)A(z) + i\kappa_{ac}(z)B(z)e^{i\Delta\beta z} \tag{4.61}$$

$$\frac{\partial B(z)}{\partial z} = -i\kappa_{ac}^{*}(z)A(z)e^{-i\Delta\beta z} - i\kappa_{dc}(z)B(z) \tag{4.62}$$

with $\Delta\beta \equiv 2\beta_{core} - \mathrm{K}$

This can be written in a matrix form as:

$$\begin{bmatrix} A(z)' \\ B(z)' \end{bmatrix} = i\begin{bmatrix} \kappa_{dc} & \kappa_{ac}e^{-i\Delta\beta z} \\ -\kappa_{ac}^{*}e^{+i\Delta\beta z} & -\kappa_{dc} \end{bmatrix}\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} \tag{4.63}$$

The solution of the coupled-mode equation of two counter propagating modes is given by

$$\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = \begin{bmatrix} \left(\cosh sz + i\frac{\Delta\beta}{2s}\sinh sz\right)e^{i\frac{K}{2}z} & i\frac{\kappa}{s}e^{i\frac{K}{2}z}\sinh sz \\ -i\frac{\kappa*}{s}e^{-i\frac{K}{2}z}\sinh sz & \left(\cosh sz - i\frac{\Delta\beta}{2s}\sinh sz\right)e^{-i\frac{K}{2}z} \end{bmatrix}\begin{bmatrix} A(0) \\ B(0) \end{bmatrix} \tag{4.64}$$

where s is $\sqrt{\kappa\kappa* - \left(\frac{\Delta\beta}{2}\right)^2}$.

4.2 Application I: Optical Delay Control

Chirped FBGs can be used to precisely control the optical delay and this can be used for applications in telecommunication, fiber-optic sensors, microwave photonics, and bio-imaging. Chirped fiber Bragg gratings (CFBGs) have been used for dispersion compensating devices due to their fiber compatibility, polarization insensitivity, low nonlinearity, low loss and so on [6]. By compensating dispersion in optical systems, high quality optical signal transmission was realized [7, 8]. Multiple elements photonic microwave true-time delay beam-forming based on the chirped FBGs was achieved [9]. High speed and real time optical imaging technique was also obtained by using chirped FBGs [10]. The most important property of the chirped FBGs is that it can control the optical delay time without center wavelength shift. A simple and convenient method to control the optical delay time is to exploit a symmetrical bending apparatus [11]. When the linear strain like the tension or compression strain is induced along the length of a fiber grating, the resonant wavelength is shifted into the longer and shorter wavelength,

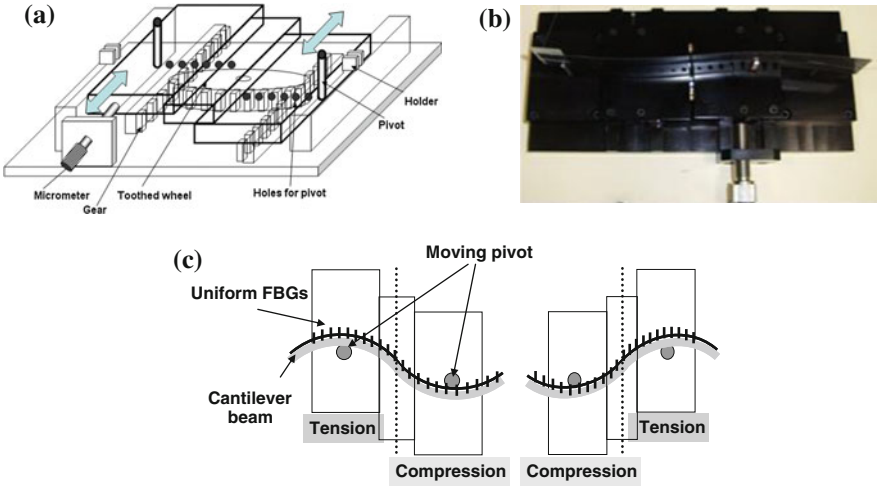

Fig. 4.3 **a** Schematic of a chromatic dispersion controller with a uniform FBG. **b** Photograph of the fabricated bending apparatus. **c** Symmetrical bending scheme based on two moving stage. Tension and compression strain depending on the bending direction can be induced [11]

respectively, Consequently, the chirp ratio of uniform FBG is changed by the tension and compression strain induced by the symmetrical bending, but the center wavelength is not changed because the effect of tension and compression strain on the center wavelength shift should be compensated mutually. Therefore, the group delay and dispersion can be controlled without the center wavelength shift.

Figures 4.3a and b show the experimental scheme and the photograph for the symmetrical bending apparatus, respectively, to control the optical delay time with a uniform FBG [11]. The sophisticated fiber bending can induce symmetrically linear strains gradient in the center of the uniform FBG. It consists of two translation stages with gears, a sawtooth wheel, two pivots, a micrometer, and two cantilever beam holder. Opposite movement of two translation stages with gears converts the linear motion of translation stage into the rotary motion of sawtooth wheel. Figure 4.3c shows the induction principle of tension and compression strain along the uniform FBG depending on the bending direction. When the left translation stage is moved forward by the micrometer, its gear rotates the sawtooth wheel and the right translation stage is moved backward by the rotary motion of the sawtooth wheel. As two translation stages are moved oppositely by the interaction between two gears and a sawtooth wheel, the position of two pivots on two translation stages is changed oppositely, which induces the symmetrical bending along the flexible cantilever beam. Consequently the tension and compression strain along the uniform FBG through the symmetrically curved cantilever beam corresponding to the bending direction can be induced. Therefore, the optical time delay of the uniform FBG can be effectively controlled by inducing the tension and compression strain at each side of the FBG without the center wavelength shift.

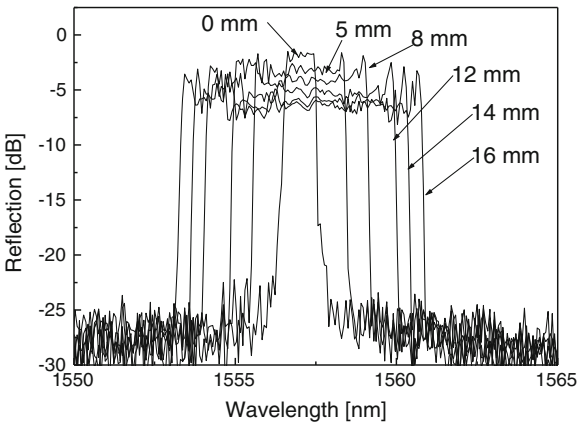

Fig. 4.4 Measured reflection spectra as a function of the variations in the translation stage moving distance [11]

The flexible cantilever beam is made of a spring steel with the high resistance against fatigue and corrosion. The uniform FBG was carefully attached to the cantilever beam using the UV curable epoxy to reduce the phase error along the fiber grating due to the microbending, which can induce additional phase error. The uniform FBG was apodized by using the Blackman profile to reduce the sidelobes and the group delay ripple [7].

Figure 4.4 shows the experimentally obtained reflection spectra of the uniform FBG as a function of the variations in the moving distance of the translation stage. As the micrometer moves, the bending curve along the cantilever beam becomes larger and this increases the amount of tension and compression strain corresponding to the bending direction. A large amount of strain gradient changes the chirp ratio along the uniform FBG and consequently makes its bandwidth be broad without the center wavelength shift.

Figures 4.5a and b show the measured group delay and the measured group velocity dispersion of the uniform FBG with the variation of the translation stage. When the left translation stage was changed, the dispersion of CFBG was controlled in the range from 312.6 ps/nm to 35.9 ps/nm. The small difference between two results may be caused by the imperfection in the fabrication of the grating or in curing process and coating material. The group delay ripple and the amplitude of group delay ripple were successfully reduced, which was measured to be less than $\sim \pm 5$ ps over the whole dispersion tuning range. Since the uniform FBG apodized by the Blackman profile was utilized, the stitching error induced by the imperfection of the phase mask could be removed and the group delay ripple could be suppressed. It is obvious that the uniform FBG is more effective for achieving tunable chromatic dispersion control compared to using the chirped FBG.

Compared to FBGs, LPGs have an advantage of being able to be mass produced due to their large periodicity and the nature of amplitude masks [1]. Therefore, LPGs have potential for various applications, for example, gain flattening of erbium doped fiber amplifier (EDFA), band-rejection filters [1]. Additionally, LPGs is attracting interest for applications in sensing strain and temperature due to

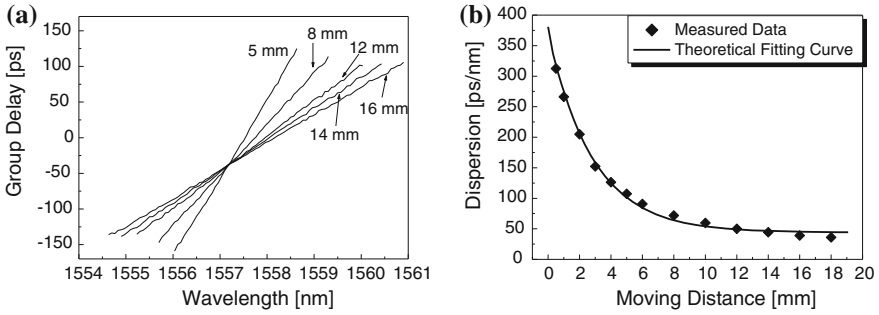

Fig. 4.5 (**a**) Measured group delay and (**b**) group velocity dispersion with variations in the moving distance of the translation stage [11]

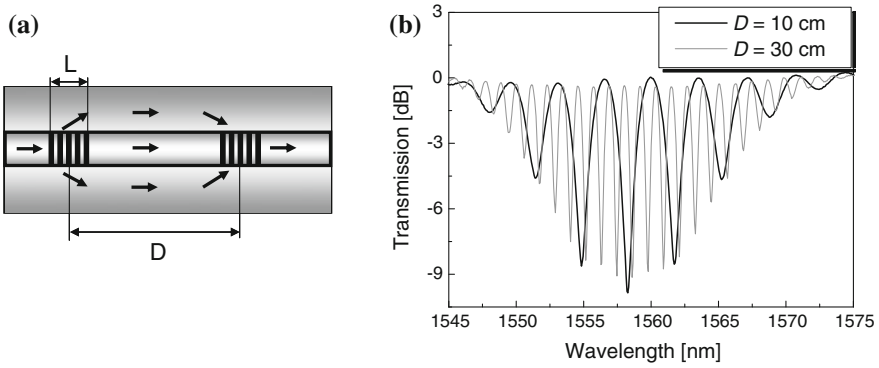

Fig. 4.6 **a** Operating principle of the cascaded LPFG resulting in the interference between the core and cladding modes. (**b**) transmission spectra of the cascaded LPGs with different separation distances [22]

its high sensitivity to these parameters [2]. LPGs are usually fabricated with the amplitude mask method or the point-by-point method. Based on these methods, several techniques have been developed for fabrication of the LPGs: (1) irradiation of UV laser beam on hydrogen-loaded GeO_2-doped fibers with the photosensitivity effect [1]; (2) periodic relaxation of residual stress with CO_2 laser [12, 13]; (3) periodic physical deformation of the core with electric arc, CO_2 laser, and flame [14–16]; (4) microbending with electric arc [17]; and (5) thermal diffusion in nitrogen-doped silica-core fiber with electric arc or CO_2 laser [18]. Several mechanisms, e.g., color center model [19] and residual stress relief model [20], have been proposed for explanation of the refractive index change.

It is possible to realize the true time delay based on the cascaded LPGs [21]. As seen in Fig. 4.6, the interference pattern resulting from the interaction between the core and the cladding modes in two identical LPGs must be generated [22]. Optical properties of the cascaded LPGs, such as peak spacing, bandwidth, and the number of peaks, are controlled by changing physical parameter, such as the length of a

LPG, the separation distance between two LPGs, the number of LPGs [22]. The differential group delay ($\Delta\tau$) in the cascaded LPGs can be derived by [21]

$$\Delta\tau = \frac{1}{c}(n_{Co} - n_{Cl}^{m} - \lambda\frac{d}{d\lambda}(n_{Co} - n_{Cl}^{m})), \tag{4.65}$$

where c is the light velocity and n_{Co} and n_{Cl}^{m} are effective indices of the core and the cladding modes, respectively. m is the cladding mode order. If a single pulse in time domain propagates along the cascaded LPGs, the first LPG divides a single pulse into several pulses, which have different delays. Since the effective index of the cladding is lower than that of the core, the optical pulse in the cladding mode propagates faster than that in the core mode. As seen in Eq. (4.65), it is clearly obvious that the amount of $\Delta\tau$ is dominantly determined by the differential effective group index between the core and the cladding modes in a single-mode fiber. However, the optical time delay induced by the cascaded LPGs is very short ($\sim$11.8 ps/m [21]) because of the small difference of the effective refractive indices.

4.3 Application II: Mechanical Sensors

Fiber gratings, in general, have high sensitivity to external perturbation such as temperature, strain and bending, which have also led to much interest for sensing applications [2]. The liquid level sensor based on phase shifting of LPGs induced by the ambient refractive index was reported [23, 24] and the application of the peak splitting of LPGs to bend sensors was reported [25, 26]. These sensors, however, have limitations, such as concurrent sensitivities to multiple perturbations; e.g., strain and temperature or bending and temperature. Versatile methods of discriminating two coexisting sensitivities have been proposed, including combination of two fiber Bragg gratings (FBGs) with different cladding diameter [27] FBG-based practical sensors including a supplementary bending cantilever beam were proposed as a promising solution for another simultaneous measurement of pressure and temperature [28] and for measuring displacement and temperature [29]. A single sampled chirped FBG (CFBG) embedded on a flexible cantilever beam is capable of discriminating bending and temperature sensitivities [30]. The sampled CFBG has multiple resonant peaks corresponding to the chirp ratio and the number of grating samples. Figure 4.7(a) and 4.7(b) show the scheme for the sensing probe and the transmission spectrum of a sampled CFBG embedded on a cantilever beam for simultaneous measurement of bending and temperature. The sampled CFBG was fabricated after exposing a photosensitive fiber to a 244 nm Ar+ laser beam through a chirped phase mask. The UV laser was periodically opened and closed by using a shutter to modulate the amplitude of the UV-induced refractive index resulting in the formation of the sampled CFBG. Then the sampled CFBG was carefully attached to the cantilever beam by using the UV curable epoxy [30].

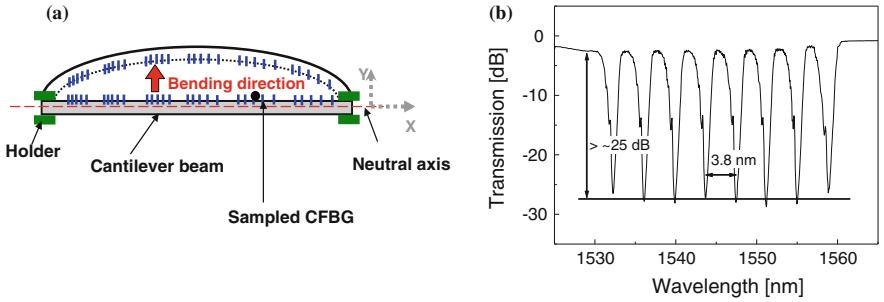

Fig. 4.7 **a** Scheme for the sensing probe based on a sampled CFBG embedded on a cantilever beam for simultaneous measurement of bending and temperature. **b** Transmission spectrum of the sampled CFBG [30]

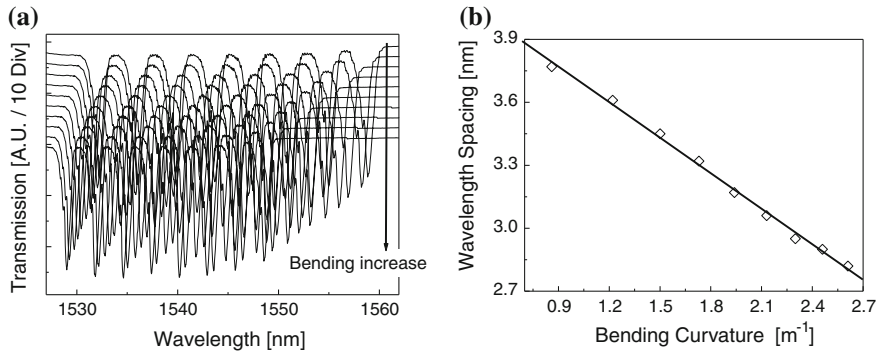

Fig. 4.8 **a** Transmission spectra of the sampled CFBG. **b** Wavelength spacing change as a function of the bending curvature change [30]

When the external bending is applied to the sampled CFBG, the wavelength spacing of the sampled CFBG is changed because the compressive strain gradient induced by the bending of the cantilever beam modifies the chirp ratio of the chirped FBG. However, the wavelength spacing of the fiber grating is not changed by the applied temperature because the chirp ratio of the sampled CFBG is not changed by variation in temperature. The multiple resonant wavelengths, however, are only shifted into the longer wavelength due to the positive thermal expansion and optic coefficients of the fiber grating [30].

Figure 4.8a shows the transmission spectra of the sampled CFBG with variations in the bending curvature. Since the compression strain gradient induced by the bending reduces the chirp ratio of the sampled CFBG, the wavelength spacing diminishes as the bending curvature increases. It should be manifest that the multiple resonant peaks also shift into the shorter wavelength due to the compression strain along the sampled CFBG [30].

Figure 4.9(a) shows the transmission spectra of the sampled CFBG as the applied temperature changes. Figuress 4.9(b) and 4.9(c) show the multiple

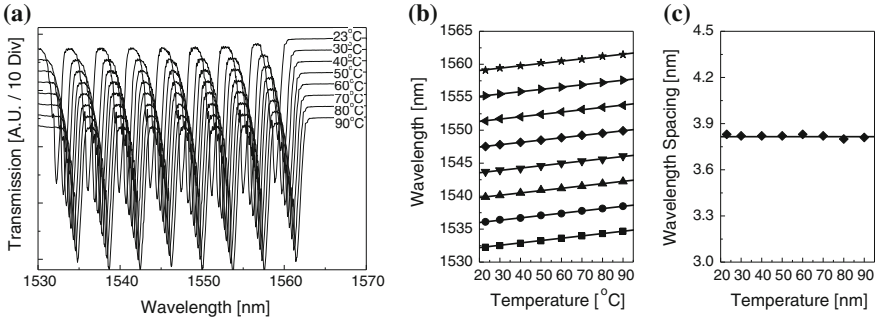

Fig. 4.9 **a** Transmission spectra of the sampled CFBG, **b** multiple resonant wavelength shifts, and (c) wavelength spacing change as a function of temperature, respectively [30]

resonant wavelength shifts and the wavelength spacing change as a function of temperature, respectively. All of multiple resonant wavelengths were shifted into the longer wavelength due to the positive thermal expansion and optic coefficient of the photosensitive fiber with high concentration of germanium. The wavelength spacing, however, was not changed by the applied temperature because of the independence of the chirp ratio on the ambient temperature variation [30].

In general, the cladding modes coupled from the core mode in the LPGs is directly interfaced with external environment. It means that the LPGs have higher sensitivity to the external perturbation than FBGs in terms of temperature sensitivity. The temperature sensitivity of LPGs is changed by controlling the doping concentrations of GeO_2 and B_2O_3 in the core region. The deviation of the waveguide property due to temperature variation ($d\Lambda/dT$) is negligible compared to that of the material property (dn/dT), and the temperature dependence of the resonance peak wavelength can be written as [31]

$$\frac{d\lambda}{dT} \approx \Lambda\left(\frac{dn_{Co}}{dT} - \frac{dn_{Cl}^{m}}{dT}\right) \tag{4.66}$$

It is obvious that the temperature sensitivity of the LPG can be controlled by adjusting dopants, such as B_2O_3 with negative thermal property. For example, the temperature sensitivity can be decreased by increasing the doping concentration of B_2O_3 in the core. On the other hand, the temperature sensitivity can be positively enhanced by doping B_2O_3 in the inner cladding [31].

Two kinds of single-mode fibers with similar properties except the doping concentrations of GeO_2 and B_2O_3 were exploited to fabricate LPGs with UV excimer laser. The physical parameters of the two fibers are: core diameter = 3.6 μm and 3.8 μm, relative index difference = 1.0 % and 0.8 %, cut off wavelength = 960 nm and 910 nm. The cladding diameter is 125 μm for both. The length and the period of the LPG were 1.17 mJ/mm^2, 2 cm, and 400 μm, respectively. The temperature and strain sensitivities of two LPGs for several cladding mode orders (m) were then measured, and the measurement results are shown in Table 4.1. LPG1 and LPG2 have positive and negative temperature sensitivities,

Table 4.1 Temperature and strain sensitivities of LPG1 and LPG2 with the cladding mode order (m) [31]

	LPG1			LPG2		
	$HE_{1,3}$	$HE_{1,4}$	$HE_{1,5}$	$HE_{1,3}$	$HE_{1,4}$	$HE_{1,5}$
Wavelength [nm]	1217.27	1303.13	1488.86	1393.52	1488.86	1681.35
$d\lambda/dT$ [nm/ °C]	0.06	0.07	0.10	−0.57	−0.59	−0.65
$d\lambda/d\varepsilon$ nm/μstrain]	0.41	0.42	0.46	0.43	0.46	0.51

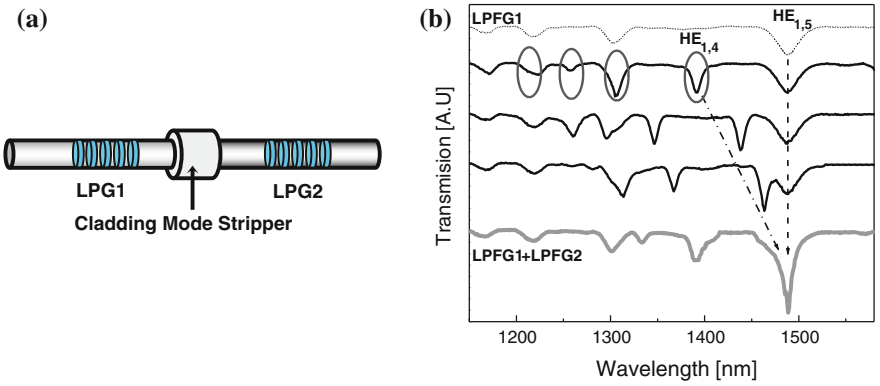

Fig. 4.10 **a** Scheme for the sensing probe based on two LPGs with similar strain sensitivities for the simultaneous measurement of strain and temperature. **b** Evolution of the transmission characteristics of two LPGs during the grating formation. [31]

respectively. The data also indicates that the temperature and strain sensitivities of LPG1 and LPG2 vary with the cladding mode order. The strain sensitivity of LPGs depends on the strain-optic coefficient, grating period Λ, and the cladding mode order [2]. In previous reports, it was shown that the strain sensitivity of LPG increases with the cladding mode order [2]. The strain sensitivity of the $HE_{1,5}$ mode in LPG1 was similar to that of the $HE_{1,4}$ mode in LPG2 and their temperature sensitivities were opposite in sign. This property makes it possible to discriminate between the temperature and strain effects simultaneously [31].

Figure 4.10a shows the simple structure of the sensing probe based on dual LPGs for simultaneous measurement of strain and temperature. A cladding mode stripper between the two gratings is required to remove the interference pattern because the cladding modes coupled from the core mode by the first LPG can interfere with the core mode again in the second LPG. After fabricating LPG1, we tried to make LPG2 such that the resonance wavelength coincides with the main resonant wavelength of LPG1. To simply the sensing signal interrogation, peak separation and shifting of two LPGs with different temperature sensitivities and the same strain sensitivites were exploited. First of all, Two LPGs with positive and negative temperature sensitivities were fabricated to induce the peak separation with variations in temperature. Two LPGs, however, had similar strain sensitivity

so that resonant peak shift could be obtained with variations in strain. Since the temperature sensitivities of two LPGs are opposite in sign while the strain sensitivities are about the same, the resonant peak separation and shift should be induced as the external temperature and strain change, respectively. This allows unambiguous and simultaneous measurement of temperature and strain. The total loss including the splicing loss and the loss due to the mode mismatch was less than 0.1 dB. The overall length of the device is about 5 cm. Figure 4.10b shows the transmission characteristics of LPGs with the overlapped resonant wavelength. The circles show the resonant peaks of LPG2. In general, the core mode ($HE_{1,1}$) can be coupled to several cladding modes ($HE_{1,m}$) of LPGs [1] and the multi-resonant peaks appears in the transmission spectrum as seen in Fig. 4.10b (the dashed line). After fabricating LPG1, we measured the variation of the transmission spectra of the two gratings during the grating formation of LPG2. The resonant wavelength of LPG2 shifted to longer wavelength due to increase of the average index during the grating formation and finally overlapped with that of LPG1 as shown in Fig. 4.10b (the gray line). The resonant wavelength of $HE_{1,4}$ of LPG2 overlapped with that of $HE_{1,5}$ of LPG1, which has similar strain sensitivity and opposite temperature sensitivity. The overlapped wavelength makes it easier to measure the strain and temperature sensitivity with a single light source since we need to detect only one wavelength. All of the resonant peaks of two LPGs cannot be made to overlap consistently due to different photo-induced refractive index changes with the cladding mode order during the grating formation

Figure 4.11 shows the peak separation of the sensing probe as a function of temperature. The transmission characteristics of LPGs with the temperature change were shown in the inset. The peak separation was induced as the temperature increased because of the opposite temperature sensitivities of the two LPGs. Since the negative temperature sensitivity of LPG2 was larger in magnitude than the positive sensitivity of LPG1, the shift of the resonant peak to the left was larger than to the right. The temperature sensitivity of the sensing probe was

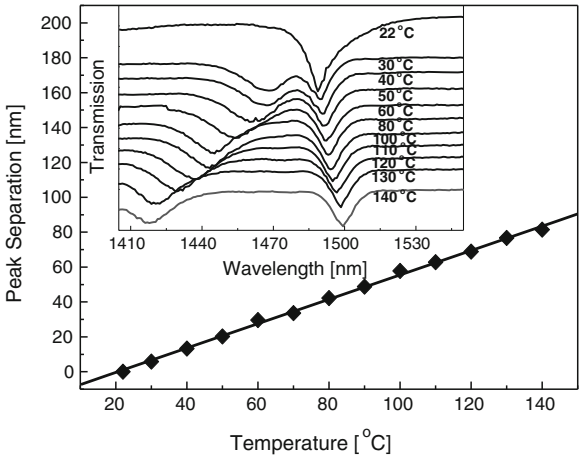

Fig. 4.11 Peak separation as a function of temperature. The transmission characteristics of the sensing probe with variations in temperature were shown in the inset [31]

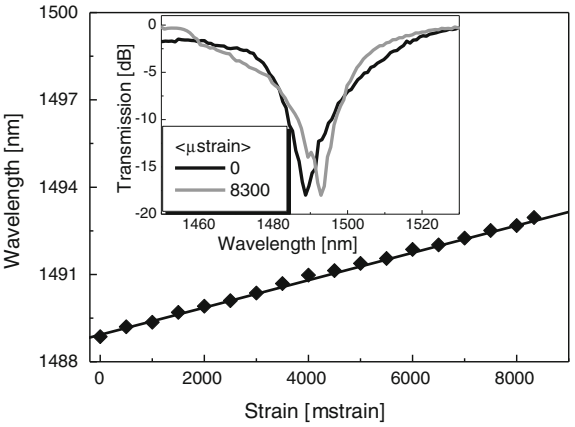

Fig. 4.12 Peak shift as a function of strain. Transmission spectra of the sensing probe with variations in strain were shown in the inset [31]

estimated as 0.69 nm/°C. Figure 4.12 shows the dependence of LPGs on the strain change and the transmission characteristics are shown in the inset. Since the two LPGs have similar strain sensitivity (∼0.46 nm/μstrain), the resonant wavelength shifted to the longer wavelength with the strain change. The measured strain sensitivity was 0.46 nm/μstrain.

4.4 Application III: Bio Sensors

Biosensors, including biomarker detection for medical diagnostics, and pathogen and toxin detection in food and water, have been attracting much attention. In general, since conventional biosensors have exploited a flouro-immunoassay method,.it is necessary to utilize the fluorescence labeling of the antigen or target DNA, which always requires additional reagents [32]. It means that the ordinary biosensors have many drawbacks, such as a high cost and the complicated and real-time detection configurations. To overcome these disadvantages, many methods have been applied to develop label-free detection biosensors [33, 34]. In particular, there has been much interest in fiber optic biosensors have been the promising techniques to realize high quality, label-free detection because of their various advantages, such as high sensitivity, fast detection speed, small size, variable and multiple detection of biosamples. A variety of biosensing techniques based on the fiber-optic surface plasmon resonance (SPR) phenomenon have been proposed [33, 35, 36]. In the case of the SPR biosensors, however, it is necessary to precisely design and fabricate the SPR biosensors because the SPR properties are highly sensitive to the metal contents, its thickness, and bio-molecules. Recently, biosensors based on fiber gratings have been widely investigated [37–41]. Most of fiber-grating-based biosensors have utilized long period gratings (LPGs) tuwith radiation mode coupling at resonance wavelengths that are very sensitive to the variation of the external medium [42]. To improve the sensitivity

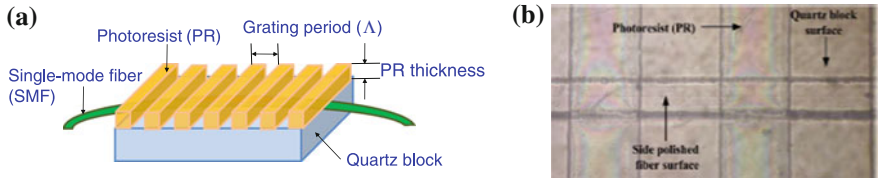

Fig. 4.13 **a** Scheme of the surface LPG. **b** Microscope image of the surface LPG [44]

of LPG-based biosensors, several methods, such as colloidal gold modified LPGs [39], LPGs with the etched cladding, [40], and LPGs with the nanostructured overlayer [41] have been proposed. However, these methods are disadvantageous because of the complicated and hazardous procedure, and the requirement of additional materials. A new, sensitive DNA biosensor based on a surface LPG inscribed on the surface of a side-polished fiber has been recently reported [43].

To fabricate a surface LPG, a single-mode fiber (SMF) was fixed by using an UV curable epoxy (NOA81) on a quartz block and ground down roughly on a brass plate by using an Al_2O_3 powder. The thickness of the cladding region was determined by measuring the transmission loss induced by the surface roughness [44]. Then, the roughly polised-SMF was polished again on a polyurethane plate by using a CeO_2 powder to reduce the surface roughness. After all polishing processes, the slurry on the flat surface should be removed by using an ultrasonic cleaning technique with de-ionized water and was dried at a temperature of $100^{O}C$ for 10 min. The photoresist (PR, Az4210) was coated on the surface of a D-shaped fiber by using a spin coater. Then, all samples were soft-baked in a thermal oven for 30 min at 90O C to remove some mixed solvent from the PR. In order to induce periodic surface gratings, the side-polished SMF with the PR coating was exposed to an UV lamp or an UV laser through a shadow long-period mask and developed by using a specific developer (Az400 k 1:3.5. Finally the periodic structure of the PR overlay on the surface of the side-polished fiber was periodically could be formed, which was a surface LPG. To stabilize the properties of the surface LPG, the post-baking process should be taken. Figures 4.13a and b show the scheme of the surface LPG and its microscope, respectively.

Since the refractive index of the PR-LPG overlay is higher than that of the core mode in the SMF, it couples the core mode to leaky modes excited by the surface PR-LPGs. When the effective index of the mth leaky mode is matched with the effective index (n_{eff}) of the core mode in the PCF, resonant coupling between the PCF and the surface PR-LPGs occurs. The resonant wavelength (λ_m) can be written as [45–47]:

$$\lambda_m = \frac{2d\sqrt{(n_g^2 - (n_{eff})^2)}}{m + \varphi}, \tag{4.67}$$

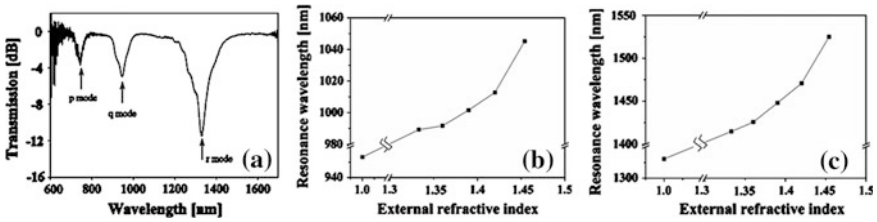

Fig. 4.14 (**a**) Transmission spectrum of the LPG, (**b**) resonance wavelengths of he q mode, and (**c**) r mode as functions of ambient indices [40]

where d and n_g are the thickness and the refractive index of the surface grating overlay, respectively. The phase change, φ, associated with external index change can be expressed by [45–47]

$$\varphi = tan^{-1}\xi \frac{(n_{eg}^2 - n_{ext}^2)^{\frac{1}{2}}}{(n_g^2 - n_{eg}^2)^{\frac{1}{2}}}, \tag{4.68}$$

where ξ is a polarization dependent constant. n_{ext} is an external refractive index. ξ can be determined by the polarization states [45–47]. It is clearly evident that external refractive index (n_{ext}) diminishes the amount of phase shift (φ), which contributes to the resonant wavelength shift to longer wavelengths [463-48].

Figure 4.14(a) shows the transmission spectrum of the surface LPG. The three different dips in the transmission spectrum indicate three different resonance wavelengths where a fundamental core mode couples to three leaky modes, p, q and r (named for convenience), in the surface LPG. Because the effective index of a high order leaky mode is smaller than a low order mode, the mode number of the r mode should be higher than that of the q mode and the resonant peak of the r mode should be appeared in a longer wavelength than other two modes. The biggest dip in Fig. 4.14a indicates that strong coupling occurred in the surface LPG between the fundamental core mode and higher mode order. Note that the resonance wavelengths, corresponding to the r and q modes, are changed with variations in the refractive indices of an external medium which are shown in Fig. 4.14b. As refractive index was increased, the resonant wavelengths shifted to a longer wavelength. When the refractive index was changed in a range from 1.333 to 1.454, the amount of resonance wavelength shifts for the q and r modes (as seen in Figs. 4.14b and c, respectively) were 84.4 nm and 110 nm towards a longer wavelength, respectively. The corresponding sensitivities were estimated to be 607.2 nm/RIU and 909.1 nm/RIU for the q and r modes, respectively. In this case, the resolutions of the biosensor were $1.6 \times 10 - 5$ and $1.1 \times 10 - 5$ RIU for the q and r modes, respectively. It is obvious that the r mode has higher sensitivity than the q mode because the higher order modes have higher sensitivity than the low order ones [2, 31].

Poly-L-lysine (PLL) has an extreme positive charge with NH_3^+ in the side chain and is frequently exploited for adsorbing biomolecules with negative charge like DNA (deoxynucleic acid) [48, 49]. A PLL solution (0.1 % W/V in water, the molecular weight = 150,000-300,000 g/mol, Sigma) was employed, which was commonly used in biology to treat glass slides. The surface LPG was initially cleaned by using a phosphate-buffered saline solution (NaH_2PO_4/Na_2HPO_4 pH 7.4, 150 mM NaCl) (PBS) before being modified with Poly-L-lysine. The PLL solution was dropped on the surface LPG to make a PLL layer for 160 min. at room temperature. The PLL layer functionalizes the surface LPG with an amino group at the free end, allowing a negatively charged DNA to be immobilized on the surface LPG. Then, the surface LPG is washed again by using a PBS buffer to remove excess PLL layers that were not immobilized on the surface LPG. The 1 μM probe of single-stranded DNA in the PBS buffer was then dropped on the surface of the PLL layer for 130 min. at room temperature. Then, the cleaning process was performed again. Finally, the 1 μM target single-stranded DNA in the PBS buffer was dropped on the surface of a probe ssDNA layer and after 65 min the sensor was washed by using a PBS buffer. About 100 μl of all biomolecules was added and removed using a micro-pipette. Figure 4.15 shows the molecular structure of the PLL, probe ssDNA and target ssDNA. The DNA sequences were (5'-CAG CGA GGT GAA AAC GAC AAA AGG GG-3') for the probe ssDNA and (5'-CCC CTT TTG TCG TTT TCA CCT CGC TG-3') for the target ssDNA.

The wavelength shift of the r mode in the surface LPG was measured, which was highly sensitive to external refractive index in a range from 1.333 to 1.454. As shown in Fig. 4.16 (a), the resonance wavelength shifted to a longer wavelength as the biomolecular layers was formed. When the PBS buffer was dropped on the surface LPG, the resonance wavelength was measured to e be 1411.69 nm. The immobilization of PLL on the surface LPG made the resonance wavelength shifted to 1413.52 nm. Then, when the probe ssDNA was immobilized on the PLL layer in the surface LPG, the shift of the resonance wavelength was measured to be

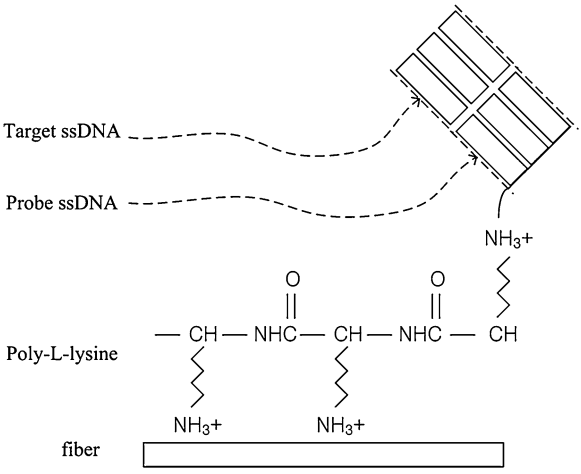

Fig. 4.15 Molecular structure of Poly-L-lysine, probe ssDNA and target ssDNA immobilized on the surface LPG [43]

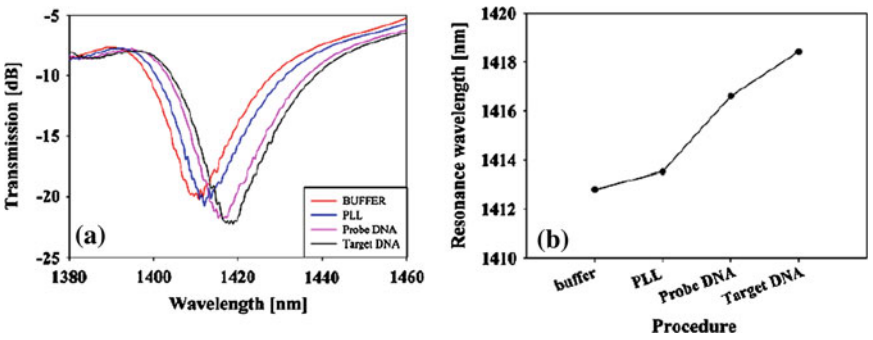

Fig. 4.16 (**a**) Transmission spectra and (**b**) resonant wavelength shifts of the surface LPG after sequential procedures of PBS buffer, PLL, probe ssDNA, and target ssDNA [43]

1416.61 nm. Finally, the hybridization with the complementary target ssDNA induced the a resonance wavelength shift of 1418.43 nm. The overall wavelength shift induced by the hybridization reaction was 1.82 nm, which is ~2.5 times higher than the previously reported biosensor based on a dual-peak LPG [49] under the same 1 μM target DNA concentration. It is evident that the surface LPG-based biosensor is highly sensitive to DNA hybridization in comparison with previously reported DNA fiber grating-based biosensors [38, 48]. The resonance wavelengths of the fiber grating-based DNA biosensor after each procedure are shown in Fig. 4.16 (b).

References

1. Vengsarkar AM, Lemaire PJ, Judkins JB, Bhatia V, Erdogan T, Sipe JE (1996) Long-period fiber gratings as band-rejection filters. J Lightwave Technol 14:58–64
2. Bhatia V, Campbell D, Claus RO, Vengsarkar AM (1996) Simultaneous strain and temperature measurement with long-period gratings. Opt Lett 21(5):336–338
3. Yariv A, Yeh P (1984) Optical Waves in Crystals. Wiley, New York, pp 177–185
4. Othonos A, Kalli K (1999) Fiber bragg gratings. Artech House, Boston, pp 189–221
5. Kogelnik H (1990) Theory of optical waveguides. In: Tamir T (ed) Guided-wave opto-eletctronics. Springer, New York
6. Eggleton BJ, Rogers JA, Westbrook PS, Strasser TA (1999) Electrically tunable power efficient dispersion compensating fiber Bragg grating. IEEE Photon Technol Lett 11(7):854–856
7. Kim J, Bae J, Han YG, Jeong JM, Kim SH, Lee SB (2004) Effectively Tunable Dispersion Compensation Based on Chirped Fiber Bragg Gratings without Central Wavelength Shift. IEEE Photon Technol Lett 16(3):849–851
8. Han YG, Lee SB (2005) Tunable dispersion compensator based on uniform fiber Bragg grating and its application to tunable pulse repetition-rate multiplication. Opt Express 13(23):9224–9229
9. Han YG, Lee JH (2007) Multiple elements photonic microwave true-time delay beamforming incorporating a tunable chirped fiber Bragg grating with symmetrical bending technique. Opt Lett 32(12):1704–1706

10. Choi ES, Na J, Ryu S, Mudhana G, Lee BH (2005) All-fiber variable optical delay line for applications in optical coherence tomography: feasibility study for a novel delay line. Opt Express 13(4):1334–1345
11. Han YG, Lee SB (2005) Tunable dispersion compensator based on uniform fiber Bragg grating and its application to tunable pulse repetition-rate multiplication. Opt Express 13(23):9224–9229
12. Akiyama M, Nishide K, Shima K, Wada A, Yamauchi R (1998) A novel long-period fiber grating using periodically released residual stress of pure-silica core fiber. Opt Fiber Conf Tech Dig 6:276–277
13. Hutsel MR, Gaylord TK (2012) Residual-stress relaxation and densification in CO2-laser-induced long-period fiber gratings. Appl Opt 51(25):6179–6187
14. Enomoto T, Shigehara M, Ishikawa S, Danzuka T, Kanamori H (1998) Long-period fiber grating in pure-silica-core fiber written by residual stress relaxation. Optical Fiber Conference Technical Digest 6:277–278
15. Lin CY, Wang LA (2001) Corrugated long-period fiber gratings as strain, torsion, and bending sensors. J Lightwave Technol 19(8):1159–1168
16. Yoon MS, Park SH, Han YG (2012) Simultaneous measurement of strain and temperature by using a micro-tapered fiber grating. J Lightwave Technol 30(8):1156–1160
17. Hwang IK, Yun SH, Kim BY (1999) Long-period fiber gratings based on periodic microbends. Opt Lett 24(18):1263–1265
18. Karpov VI, Grekov MV, Dianov EM, Golant KM, Vasiliev SA, Medvedkov OI, Khrapko RR (1998) Mode-field converters and long-period gratings fabricated by thermo-diffusion in nitrogen-doped silica-core fibers. Optical Fiber Conference Technical Digest 6:279–280
19. Dong L, Archambault JL, Reekie L, Russell PS, Payne DN (1995) Photoinduced absorption change in germanosilicate preforms: evidence for the color-center model of photosensitivity. Appl Opt 34(18):3436–3440
20. Fonjallaz PY, Limberger HG, Salathe RP, Cochet F, Leuenberger B (1995) Tension increase correlated to refractive-index change in fibers containing UV-written Bragg gratings. Opt Lett 20(11):1346–1348
21. Eom TJ, Kim SJ, Kimm TY, Park CS, Paek UC, Lee BH (2005) Realization of true-time-delay using cascaded long-period fiber gratings: theory and applications to the optical pulse multiplication and temporal encoder/decoder. J Lightwave Technol 23(2):597–608
22. Han YG, Lee BH, Han WT, Paek UC, Chung Y (2001) Controllable transmission characteristics of multi-channel long period fiber gratings. IEICE Trans Electron E84-C(3): 610–614
23. Khaliq S, James SW, Tatam RP (2001) Fiber-optic liquid-level sensor using a long-period grating. Opt Lett 26(16):1224–1226
24. Yun B, Chen N, Cui Y (2007) Highly Sensitive Liquid-Level Sensor Based on Etched Fiber Bragg Grating. IEEE Photon Technol Lett 19(21):1747–1749
25. Liu Y, Zhang L, Williams JAR, Bennion I (2000) Optical bend sensor Based on Measurement of Resonance Mode Splitting of Long-Period Fiber Grating. IEEE Photon Technol Lett 12:531–533
26. Dong X, Liu Y, Shao LY, Kang J, Zhao CL (2011) Temperature-independent fiber bending sensor based on a superimposed grating. IEEE Sensors J 11(11):3019–3022
27. James SW, Dockney ML, Tatam RP (1996) Simultaneous independent temperature and strain measurement using in-fibre Bragg Grating Sensors. Electron Lett 32:1133–1134
28. Guo T, Qiao S, Jia Z, Zhao Q, Dong X (2006) Simultaneous measurement of temperature and pressure by a single fiber Bragg grating with a broadened reflection spectrum. Appl Opt 45(13):2935–2939
29. Dong X, Liu Y, Liu Z, Dong X (2001) Simultaneous displacement and temperature measurement with cantilever-based fiber Bragg grating sensor. Opt Commun 192:213–217
30. Han YG, Dong X, Lee JH, Lee SB (2006) Simultaneous measurement of bending and temperature based on a single sampled chirped fiber Bragg grating embedded on a flexible cantilever beam. Opt Lett 31(19):2839–2841

31. Han YG, Lee SB, Kim CS, Jin U, Kang UC. Paek, and Chung Y (2003) Simultaneous measurement of temperature and strain using dual long-period fiber gratings with controlled temperature and strain sensitivity.Opt Express 11(5): 476–481
32. Charles PT, Vora GJ, Andreadis JD, Fortney AJ, Meador CE, Dulcey CS, Stenger DA (2003) Fabrication and surface characterization of DNA microarrays using amine- and thiol-terminated oligonucleotide probes. Langmuir 19:1586–1591
33. Homola J, Yee SS, Gauglitz G (1999) Surface plasmon resonance sensors: review. Sen Actuators B 54:3–15
34. Chinowskya TM, Quinn JG, Bartholomew DU, Kaiser R, Elkind JL (2003) Performance of the Spreeta 2000 integrated surface plasmon resonance affinity sensor. Sens Actuators B 91:266–274
35. Mehrvar M, Bis C, Scharer JM, Moo-Young M, Luong JH (2000) Fiber optic biosensors-trends and advances. Anal Sci 16:677–692
36. Masson JF, Barnhart M, Battaglia TM, Morris GE, Nieman RA, Young PJ, Lorson CL, Booksh KS (2004) Monitoring of recombinant survival motor neuron protein using fiber-optic surface plasmon resonance. Analyst 129:855–859
37. Maguis S, Laffont G, Ferdinand P, Carbonnier B, Kham K, Mekhalif T, Millot M (2008) Biofunctionalized tilted fiber Bragg gratings for label-free immunosensing. Opt Express 16(23):19049–19062
38. Chryssis AN, Saini SS, Lee SM, Yi H, Bentley WE, Dagenais M (2005) Detecting hybridization of DNA by highly sensitive evanescent field etched core fiber Bragg grating sensors. IEEE J Sel Top Quant Electron 11(4):864–872
39. Tang JL, Cheng SF, Hsu WT, Chiang TY, Chau LK (2006) Fiber optic biochemical sensing with a colloidal gold modified long period fiber grating. Sens Actuators B 119:105–109
40. Yang J, Sandhu P, Liang W, Xu C, Li Y (2007) Label free fiber optic biosensors with enhanced sensitivity. IEEE J Sel Top Quant Electron. 13(6):1691–1696
41. Wei X, Wei T, Xiao H, Lin YS (2008) Nano structured Pd-long period fiber gratings integrated optical sensor for hydrogen detection. Sens Actuators B 134:687–693
42. Lee KS, Erdogan T (2000) Fiber mode coupling in transmissive and reflective tilted fiber gratings. Appl Opt 39(9):1394–1404
43. Jang HS, Park KN, Kim JP, Sim SJ, Kwon OJ, Han YG, Lee KS (2009) DNA biosensor based on a long-period grating formed on the side-polished fiber surface. Opt Express 17(5): 3855–3860
44. Kwon OJ, Kim HJ, Han YG (2010) Dependence of the transmission characteristics of versatile D-shaped fibers on temperature and ambient index change for different input polarization states. J Korean Phys Soc 56(4):1274–1277
45. Kim HJ, Jun NR, Han YG (2012) Optical characteristics of a hybrid fiber grating based on surface long-period grating incorporating a fiber Bragg grating. J Korean Phys Soc 61(9):1353–1357
46. Kim HJ, Kwon OJ, Lee SB, Han YG (2012) Polarization-dependent refractometer for discrimination of ambient refractive index and temperature. Opt Lett 37(11):1802–1804
47. Kim HJ, Kwon OJ, Lee SB, Han YG (2011) Measurement of temperature and refractive index based on surface long-period gratings deposited onto a D-shaped photonic crystal fiber. Appl Phys B: Lasers Opt 102:81–85
48. Iadicicco A, Cusano A, Cutolo A, Bernini R, Giordano M (2004) Thinned fiber Bragg gratings as high sensitivity refractive index sensor. IEEE Photon Tech Lett 16(4):1149–1151
49. Sawant PD, Nicolau DV (2006) Hierarchy of DNA immobilization and hybridization on poly-L-lysine:an atomic force microscopy study. Smart Mater Struct 15:S99–S103

Chapter 5
Sagnac Loop Sensors

Li Qian

Fiber-based Sagnac loop sensors are an important subcategory of fiber sensors which utilize Sagnac interferometry. Sagnac interferometers, as opposed to all other types of common interferometers, have the unique advantage of providing the same physical path for the two counter-propagating lightwaves that create the optical interference, hence eliminating signal fading problems due to path instabilities caused by environmental disturbances. Another advantage of Sagnac interferometry as a consequence of same-path interference is the relaxed requirement on source coherence. In fact, many Sagnac loop sensors can use a broadband source (or noise source), instead of a laser source. Sagnac loop sensors work on the principle of breaking the inherent symmetry of the interferometer by the effects they are made to sense—be it the rotational speed of a gyroscope, or the Faraday effect induced by an electric current, or a traveling acoustic wave, etc. These effects introduce a measurable phase shift between the two counter-propagating lightwaves, which in turn is related to the magnitude of the effects to be sensed. Fiber-based Sagnac loop sensors have the particular advantage of being robust and low cost, for light is entirely contained in the optical fiber with no moving parts and no special alignment required. In this chapter, we will introduce the principle of Sagnac loop interferometry for optical sensing, followed by a few special examples of fiber-based Sagnac loop sensors used in practical applications.

5.1 Principle of Sagnac Loop Interferometry

Sagnac interferometer was named after the French physicist Georges Sagnac, who demonstrated in 1913 [1] that when light was split into two counter-propagating beams travelling through the same path and brought back to the same point, it

L. Qian (✉)
Department of Electrical and Computer Engineering, University of Toronto,
Toronto, ON, Canada
e-mail: l.qian@utoronto.ca

J. U. Kang (ed.), *Fiber Optic Sensing and Imaging*,
DOI: 10.1007/978-1-4614-7482-1_5,

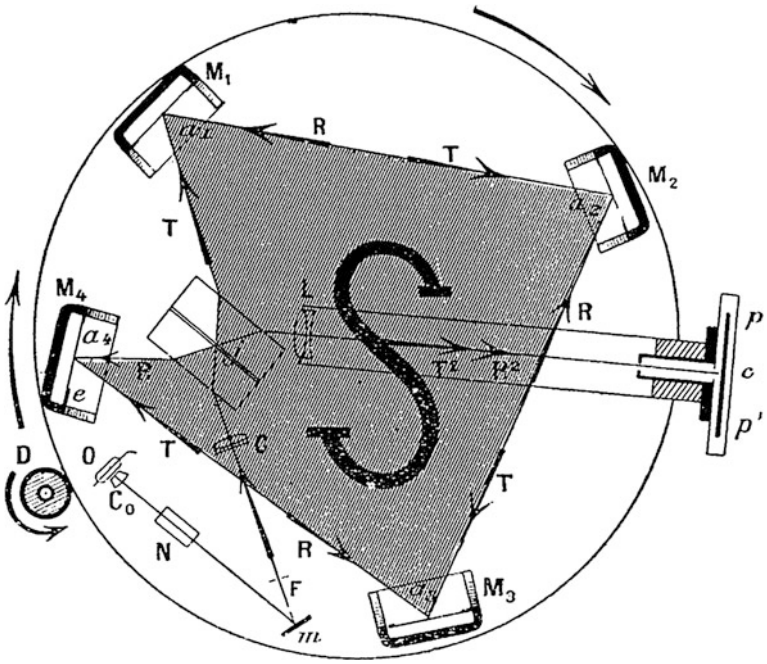

Fig. 5.1 The original experimental setup schematic published in [1] by Georges Sagnac, showing a Sagnac interferometer on a rotational platform

produced spatial interference pattern when such an experiment was set up on a rotating platform (Fig. 5.1), and the interference fringe spacing was dependent on the rotational motion of the platform. He further arrived at the conclusion that the interference pattern is consistent with the assumption that the speed of light is a constant, "independent of the collective motion of the source of light and the optical system", a concept central to the theory of Special Relativity. It is interesting to note that, more than six decades later, the first fiber optic gyroscope [2] was built on exactly the same principle.

Today, the Sagnac interferometer refers to any loop interferometer that splits the input lightwave into two that traverse exactly the same physical path in opposing directions, and recombines them at the same point. The rotating platform is not part of the Sagnac interferometer, though it is necessary for observing the Sagnac effect.

Let us first go through the principle of operation of the Sagnac interferometer, a simplified version of which is illustrated in Fig. 5.2. Here, we use an all-fiber configuration, which consists of a 50/50 fiber coupler and a single-mode fiber loop length of L. Such fiber Sagnac interferometer is also known as a fiber loop mirror, for the reason that will become clear shortly.

For simplicity, let us ignore polarization effects. We can do so for the vast majority of cases, as long as care is taken to make sure the two counter-

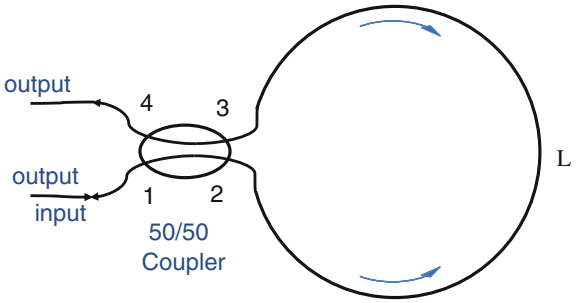

Fig. 5.2 Fiber Sagnac interferometer. The four Ports of the coupler are labelled on the diagram

propagating waves have common polarization component when they meet back at the coupler. There are a few commonly employed techniques to achieve this polarization alignment: (1) Insert a polarizer at the input and the output of the Sagnac loop. (2) Insert a fiber-based polarization controller inside the Sagnac loop. (3) Use fiber coupler and fiber loop that are made of polarization-maintaining (PM) fibers, and align input light polarization with one of the axes of the PM fiber. With polarization effects taken care of, we can use scalar electric field in the following analysis.

With the input electric field $E_0 e^{-i2\pi v_0 t}$ (where E_0 is the field amplitude, v_0 is the optical frequency and t denotes the time variable) at Port 1 of the coupler (Fig. 5.2), the electric fields at Port 2 and 3, *after* traversing the fiber loop, are given as:

$$E_2 = \frac{1}{\sqrt{2}} E_0 e^{-i2\pi v_0 t} e^{i\frac{\pi}{2}} e^{i\phi_{ccw}} \tag{5.1}$$

$$E_3 = \frac{1}{\sqrt{2}} E_0 e^{-i2\pi v_0 t} e^{i\phi_{cw}} \tag{5.2}$$

Here, the $\pi/2$ phase is added to the cross port of the directional coupler [3] and ϕ_{ccw} ϕ_{ccw} is the phase associated with light traversing the fiber loop in the counter-clockwise (clockwise) direction. These two electric fields interfere again at the coupler, resulting in output fields at Port 1 and 4, given by:

$$E_{1out} = \frac{1}{\sqrt{2}} E_2 + \frac{1}{\sqrt{2}} E_3 e^{i\frac{\pi}{2}} \tag{5.3}$$

$$E_{4out} = \frac{1}{\sqrt{2}} E_2 e^{i\frac{\pi}{2}} + \frac{1}{\sqrt{2}} E_3 \tag{5.4}$$

The output intensities at port 1 and 4 are therefore given by:

$$I_{1out} \propto \left\langle E^*_{1out} E_{1out} \right\rangle = \frac{1}{2} E_0^2 (1 + \cos \Delta\phi) \tag{5.5}$$

$$I_{4out} \propto \left\langle E^*_{4out} E_{4out} \right\rangle = \frac{1}{2} E_0^2 (1 - \cos \Delta\phi) \tag{5.6}$$

where "$\langle\,\rangle$" denotes the time average, and $\Delta\phi \overset{\text{def}}{=} \phi_{ccw} - \phi_{cw}$, is the phase difference experienced by the two counter-propagating lightwaves traversing the fiber loop.

Note, when the Sagnac loop is stationary and in the absence of any external factors that break the reciprocity of the Sagnac loop, we have $\phi_{ccw} = \phi_{cw} = 2\pi v_0 nL/c$, where n is the effective index of the fiber mode, L is the length of the loop, and c is the speed of light in vacuum. In this case, there is zero phase difference between the two counter-propagating lightwaves, and all of the input power is reflected at Port 1 ($I_{1out} = I_{1in}$), hence it is also known as the fiber loop mirror.

However, if the fiber loop is rotating, or if other external factors exist that introduce non-reciprocity into the interferometer, then there will be a phase difference, i.e., $\Delta\phi \neq 0$, and the magnitude of the external factors may be determined by measuring $\Delta\phi$. This is the basis for fiber Sagnac loop sensing. In the following sections, we will provide a few examples of how reciprocity is broken and how $\Delta\phi$ is used in measurements and in sensing applications.

5.2 Interferometric Fiber Optic Gyroscope

Gyroscopes detect rotational movements and are commonly used in navigational systems, aeronautics, and space exploration. Historically, they were based on mechanical gyros consisting of a spinning wheel and gimbaled mounts. The optical gyroscopes based on the Sagnac effect have a major advantage over the mechanical gyros because it is "solid state" with no moving parts. It is based on the Sagnac principle discovered by Georges Sagnac in 1913.

A rigorous way of analyzing the rotating Sagnac loop is to apply relativistic electromagnetic theory to light propagation in a rotating medium [4, 5]. Below, we will adopt a simpler approach to explain the phenomenon quantitatively. Consider the fiber-optic Sagnac loop in Fig. 5.3 which has a total loop length L, wound N times around a circle of radius R, that is, $L = 2\pi RN$. If the loop is rotating counterclockwise with an angular frequency Ω, then there are two effects we need to

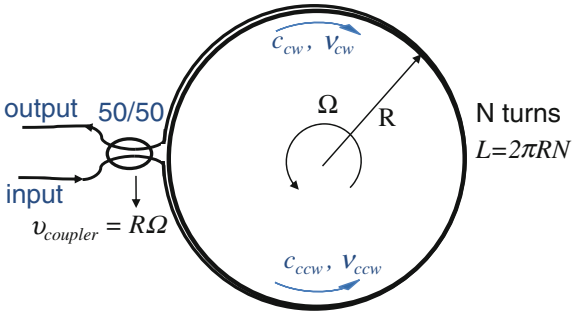

Fig. 5.3 A rotating fiber Sagnac loop seen at a fixed time instant. The coupler is moving at a speed $R\Omega$, tangential to the loop

consider in order to calculate the phase difference between the two counter propagating lightwaves.

(1) Relativistic Doppler effect due to the fact that the coupler is moving at a speed of $R\Omega$, tangential to the loop, as indicated by the arrow near the coupler in Fig. 5.3. The photons injected into the loop propagating counter-clockwise can be considered as if they come from a source that is moving in the direction of propagation, and experience a Doppler shift to a high frequency (blue shift). Conversely, light injected into the loop propagating clockwise is red shifted. When the light exits the loop, they experience the opposite Doppler shift, returning to their original frequency. The amount of Doppler shift is related to the ratio of the speed of the source (in this case, the coupler) to the speed of light propagation: $R\Omega/(c/n)$, where c is the speed of light in vacuum, and n is the effective index of the fiber mode. Here, c/n is approximately the speed of light traveling in the fiber loop (see the second point below for more details). If we denote v_{cw} (v_{ccw}) as the Doppler-shifted frequency for lightwave traveling clockwise (counter-clockwise), then its relationship with the original frequency v_0 is given by:

$$\frac{v_{cw}}{v_0} = \sqrt{\frac{1-\frac{R\Omega}{c/n}}{1+\frac{R\Omega}{c/n}}} \approx 1 - \frac{R\Omega}{c/n} \tag{5.7}$$

$$\frac{v_{ccw}}{v_0} = \sqrt{\frac{1+\frac{R\Omega}{c/n}}{1-\frac{R\Omega}{c/n}}} \approx 1 + \frac{R\Omega}{c/n} \tag{5.8}$$

(2) The Fresnel drag effect (also known as the Fizeau drag) due to the fact that the medium in which light travels, in this case, the fiber loop, is moving. When the medium is traveling in the same direction as the light, light propagates faster in the inertial frame of the observer, and the converse is also true. Let us use c_{cw} and c_{ccw} to denote the speed of light in the fiber loop traveling clockwise and counterclockwise, respectively, then we have:

$$c_{cw} = \frac{c}{n} - \left(1 - \frac{1}{n^2}\right) R\Omega \tag{5.9}$$

$$c_{ccw} = \frac{c}{n} + \left(1 - \frac{1}{n^2}\right) R\Omega \tag{5.10}$$

Now we are ready to calculate the phase shifts experienced by the two counter-propagating waves:

$$\phi_{cw} = 2\pi \frac{L}{\lambda_{cw}} = 2\pi \frac{2\pi RN}{c_{cw}/v_{cw}} = 4\pi^2 RN \frac{v_{cw}}{c_{cw}} \tag{5.11}$$

$$\phi_{ccw} = 2\pi \frac{L}{\lambda_{ccw}} = 2\pi \frac{2\pi RN}{c_{ccw}/v_{ccw}} = 4\pi^2 RN \frac{v_{ccw}}{c_{ccw}} \tag{5.12}$$

Therefore, the phase difference can be calculated as:

$$\Delta\phi = \phi_{ccw} - \phi_{cw} = 4\pi^2 RN \left(\frac{v_{ccw}}{c_{ccw}} - \frac{v_{cw}}{c_{cw}} \right) \approx \frac{4\pi^2 RN}{c^2} n^2 (v_{ccw} c_{cw} - v_{cw} c_{ccw}) \tag{5.13}$$

Note, we used an approximation $c_{cw} \approx c_{ccw} \approx c/n$ for the c_{cw} and c_{ccw} in the denominators. Substituting (5.7)–(5.10) into (5.13) and taking the first-order approximation for $R\Omega/c$, we obtain

$$\Delta\phi \approx 4\pi^2 RN v_0 \frac{2R\Omega}{c^2} = \frac{8\pi AN}{c\lambda_0} \Omega \tag{5.14}$$

where A is the area enclosed by the loop, and λ_0 is the wavelength of light in vacuum. Though (5.14) is derived with a circular geometry for the fiber loop, it is applicable to arbitrary loop shapes with the following vector modification [6]:

$$\Delta\phi = \frac{8\pi N}{c\lambda_0} \vec{A} \cdot \vec{\Omega} \tag{5.15}$$

where $\vec{A}$ has a magnitude equal to the area enclosed by the loop, and a direction normal to the loop.

Note that the "double Doppler effect" approach we took above, which was conceptually noted by Léfèvre [6], considers the spatial domain interference where the Sagnac loop is fixed at one instant of time. A more common approach is to consider temporal-domain interference [7] where the phase difference is derived by considering the time difference for the two counter-propagating waves to reach the output coupler. The result is the same, but as we are dealing with continuous-waves here, we feel it is conceptually easier to understand in the spatial domain.

There are a few important remarks we should make regarding (5.14):

1. The phase difference, $\Delta\phi$, is proportional to Ω to the first order. As this phase difference produces the interference [cf. (5.5), (5.6)] observable at the output, the fiber loop can thus be used to sense rotational speed of the system in which it is placed. The unambiguous sensing range is determined by limiting $\Delta\phi$ to $2m\pi < \Delta\phi < 2m\pi + \pi$, and the corresponding sensing range for the rotational speed is:

$$\Omega_\pi = \frac{c\lambda_0}{8AN} \tag{5.16}$$

As a numerical example, an Interferometric Fiber Optic Gyroscope (IFOG) using the telecom wavelength of 1.55 μm (where loss is minimum), with an area of 0.01 m^2, and a fiber length of 1 km, gives Ω_π around 2 rad/s.

2. The phase difference, $\Delta\phi$, is independent of n, the refractive index of the medium. This means that the rotational measurement is not affected by the dispersion of the medium.
3. The phase difference, $\Delta\phi$, is proportional to N, the number of winding turns of fiber. Therefore, the sensitivity of the IFOG can be greatly increased without noticeably increasing the dimensions of the instrument. This is one major advantage of the FOG over mechanical gyros. The sensitivity of an IFOG is ultimately limited by the detector shot noise [6], which limits the resolution of discerning a phase difference from power measurements by the detectors. If we consider 1 μrad as a good order of magnitude for sensitivity, then the sensitivity of the IFOG is given by:

$$\Omega_\mu = \frac{\Omega_\pi}{\pi} \times 10^{-6} = \frac{c\lambda_0}{8\pi AN} \times 10^{-6} \tag{5.17}$$

Obviously there is a direct trade off between sensing range and sensing sensitivity. Taking the values used in the previous numerical example, Ω_π is 0.13 deg/h, two orders of magnitude smaller than Earth's rotation rate of 15 deg/h! Honeywell has developed IFOGs with sensitivities ranging from 1–10 deg/h to 10^{-3} deg/h [8].

4. As ultimately we are detecting the optical power variation due to interference [cf. (5.5), (5.6)], it can be seen that, while $\Delta\phi$ is linear with Ω, the output power (or intensity) is not, which affects the sensitivity of the gyroscope. In fact, sensitivity is zero if $\Delta\phi$ is centered on $m\pi$, and the direction of rotation rate change cannot be discerned at this point due to the symmetry of the intensity response at this phase angle. It is therefore necessary to "bias" the phase angle to the quadrature point ($m\pi + \pi/2$) to obtain maximum sensitivity. This is done by inserting a reciprocal phase modulator near one end of the fiber loop to produce a sinusoidal phase modulation [9–11]. This technique also allows for lock-in detection, largely eliminating external phase perturbation noise. Further improvement in sensitivity over the entire sensing range (not just at one operating point) of the gyroscope is done with a closed-loop (or phase-nulling) signal processing technique [12, 13]. It uses the output of the Sagnac loop as an error signal to produce a nonreciprocal phase difference between the two counter-propagating waves, such that the total phase difference is maintained at zero by the feedback control loop. The rotation speed is now proportional to the feedback value of the nonreciprocal phase difference and is independent of the optical power measurement, and hence this scheme provides linearity and stability.

It is worth mentioning that (5.14) is derived assuming perfect reciprocity between the two counter-propagating waves (except for the nonreciprocity induced by the rotation)—this is to a large extent guaranteed by using single-mode fiber, where the two counter-propagating waves are confined to the single mode of fiber, insuring that they traverse the exact same path in opposite directions. There are, however, a host of factors [14] that can introduce non-reciprocity into the loop, resulting in noise and fluctuations. These include: (a) backscattering (b) nonlinear effects and (c) environmental temperature and stress fluctuations.

Rayleigh backscattering and other coherent backscattering due to imperfections along the fiber (e.g. splices) occur at random locations along the fiber, which introduces non-reciprocity. As a consequence, additional phase noise that is environmentally sensitive is added to the phase difference caused by the rotation of the interferometer, introducing short-term drifts in the IFOG [9, 10, 14]. This effect can be reduced to a large extent by broadening the spectral width of the source, i.e., reducing its coherence length, hence reducing the noise due to coherent backscattering. Indeed, since the Sagnac interferometer has zero path difference, one can even use incoherent light source in IFOG [15]. This is one major advantage of IFOG over ring-laser gyroscopes.

Nonlinear effects, in particular, optical Kerr effect which produces intensity-dependent refractive index, can cause non-reciprocity if there is a power imbalance between the two counter-propagating waves [16] due to an imperfect 50/50 coupler. There are two contributions to the phase due to the optical Kerr effect, one is a self-induced nonlinear phase, also known as self-phase modulation, due to the intensity of the wave itself. The other is cross-phase modulation, due to the co-existing counter-propagating wave. The latter is twice as strong as the former [17], and as a result, any power imbalance will lead to an imbalance of the nonlinear phase added to the two waves. Even though optical Kerr effect is very weak, because of the large length of fiber used in most gyroscopes, nonreciprocal phase change due to nonlinear index change can be significant. For example, silica fiber has a nonlinear refractive index n_2 of $3 \times 10^{-8}\mu m^2/W$, and the mode area of a single-mode fiber at the telecom wavelength is about $70\mu m^2$. For a loop length of 1 km, a power imbalance of 0.6 μW is sufficient to cause a 1 μrad nonreciprocal phase difference due to Kerr effect. Fortunately, the cross-phase effect can be reduced by half such that it to a large extent cancels the self-phase effect, by either amplitude modulating the waves with 50 % duty cycle, or using an incoherent source [17].

Environmental temperature and stress fluctuations tend to be nonuniform in space. As we will see in Sect. 5.3, nonuniform time-varying environmental factors can contribute to non-reciprocal phase difference. One can usually reduce the effects of such fluctuations through lock-in detection. Additionally, such environmental fluctuations can be reduced by judiciously winding the fiber coil such that symmetric points on the loop (with respect to the center of the loop) are placed in close proximity to each other [23, 24]. How this reduces the phase shift is explained in the next section.

5.3 Fiber Optic Acoustic Sensor

Sound wave is fundamentally a travelling pressure wave through the medium in which sound is propagating. The pressure that sound produces can cause a minute strain in the material, resulting in a change of the refractive index, known as the photoelastic effect [18]:

$$\Delta n = \frac{1}{2} n^3 p \varepsilon \tag{5.18}$$

where n is the nominal refractive index of the material, p is a photoelastic coefficient,[1] and ε is the mechanical strain induced by the pressure of the sound wave. If a fiber coil is placed in the path of a sound wave, this refractive index change will induce phase change in the light traveling through the fiber coil, which can be measured and related back to the property of the sound. The first fiber sensor proposed to use this principle to sense sound waves was in the form of a Mach–Zehnder interferometer (MZI) [19, 20]. Sagnac fiber-optic acoustic sensor was later proposed by Cahill and Udd [21, 22], who pointed out its advantages over MZI at higher acoustic frequencies.

At this point the reader may ask, how can a Sagnac loop, which can only be sensitive to nonreciprocal effects, be able to measure the refractive index change (which is reciprocal) induced by sound waves? It turns out that, as long as the perturbation over the fiber Sagnac loop is asymmetric with respect to the midpoint of the loop, there will be a phase difference induced between the two counter-propagating waves. Let us take a more detailed look below.

Suppose the perturbation to the refractive index is both a function of space and of time, i.e., $\Delta n(x,t)$, where x is the distance coordinate along the fiber loop in the counter-clockwise direction. Referring to Fig. 5.4, if a perturbation happening at (x,t) is imposing an additional phase shift to the counter-clockwise wave, the same perturbation will not affect the clockwise wave, because the clockwise wave has not reach it. Instead, the clockwise wave will be affected by the perturbation at a later time, $\Delta n(x, t + \frac{L-2x}{c/n})$, when it reaches the same location x. Note, we use n to denote the unperturbed effective index of the fiber mode, and the speed of light in the fiber is approximated to c/n.

Therefore, we have,

$$d\phi_{ccw} = \frac{2\pi}{\lambda_0} \Delta n(x,t) dx \tag{5.19}$$

$$d\phi_{cw} = \frac{2\pi}{\lambda_0} \Delta n\left(x, t + \frac{L-2x}{c/n}\right) dx \tag{5.20}$$

Therefore the phase difference induced by the perturbation at x is

$$d(\Delta\phi) = \frac{2\pi}{\lambda_0}\left[\Delta n(x,t) - \Delta n\left(x, t + \frac{L-2x}{c/n}\right)\right] dx \approx -\frac{2\pi}{\lambda_0} \frac{\partial \Delta n(x,t)}{\partial t} \frac{L-2x}{c/n} dx \tag{5.21}$$

[1] To be more precise, p is a tensor element of the strain-optic tensor. The strain tensor modifies the optical indicatrix (also a tensor). See [8].

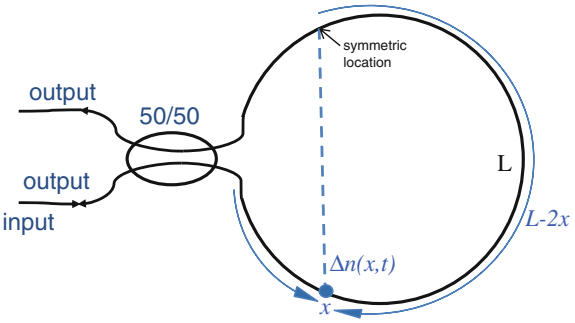

Fig. 5.4 Time-varying perturbation located asymmetrically in the Sagnac loop causes a non-reciprocal phase shift, due to the fact that the perturbation affects the clockwise and counterclockwise waves at different times

The first-order approximation used in (5.21) holds as long as the dynamics of the perturbation is slow compared to the light transit time of the loop. The total phase difference integrated over the length of the coil is

$$\Delta\phi = -\frac{2\pi n}{c\lambda_0}\int_0^L \frac{\partial \Delta n(x,t)}{\partial t}(L-2x)dx \tag{5.22}$$

It is easy to see from (5.22) that, if the perturbation is symmetric with respect to the center point of the coil, that is, if $\Delta n(x,t) = \Delta n(L-2x,t)$, then the integral will be zero, resulting in no net phase difference. Indeed, in the application of fiber gyroscope, where environmentally induced phase difference is undesirable, the fiber coil is intentionally wound to have the symmetric points on the fiber coil in close proximity to each other to minimize the non-reciprocal effect of the environmental fluctuations in temperature and pressure [23, 24]. In Sagnac acoustic sensor, however, one can use this non-reciprocal effect to measure the effect of sound waves.

From (5.22) we can also see that $\Delta\phi$ will be maximized if we expose half of the coil to the acoustic wave and shield the other half away from it. For this reason, Sagnac acoustic sensors usually have the configuration that consists of a sensing coil and a reference coil, of equal length [22, 25]. The reference coil can be shielded from the acoustic wave by a special fiber coating [26] or by other means of isolation [27].

If we now consider a simple case of harmonic acoustic wave, and assuming the acoustic wavelength is large compared to the sensing coil dimensions (which can be much smaller than the length of the fiber). Then, Δn can be treated as uniform for the sensing coil and 0 for the reference coil:

$$\Delta n(x,t) = \begin{cases} \Delta n_0 \sin(2\pi ft) & 0 < x < L/2 \\ 0 & L/2 < x < L \end{cases} \tag{5.23}$$

where f is the acoustic frequency. Substitute (5.23) into (5.22), we obtain

$$\Delta\phi = -\frac{\pi^2 L^2 n \Delta n_0}{c\lambda_0} f\cos(2\pi ft) \tag{5.24}$$

There is a linear dependence of the amplitude of $\Delta\phi$ on f, which implies that Sagnac acoustic sensors have increased sensitivities with increasing acoustic frequency,[2] making it particularly suitable for detection of underwater sound because the background acoustic noise in the ocean decreases with increasing frequency over a substantial frequency range [25].

5.4 Frequency-Shifted Sagnac Interferometer

We have seen that nonreciprocal effects can be introduced to a Sagnac loop by perturbing it *asymmetrically*. In the case of IFOG, it is largely an effect to be avoided, but it can also be taken advantage of. Non-reciprocal phase shifting elements (in the form of a phase modulator or a frequency shifter) can be introduced to the loop as a means for phase-nulling to increase the dynamic range and sensitivity of the gyroscope [12, 13]. In the case of fiber Sagnac acoustic sensing, this non-reciprocal effect is utilized to sense sound waves by isolating half of the loop from the sound wave while exposing the other half.

In this section, we specifically consider a special type of Sagnac interferometer, where a reciprocal frequency shifter is placed asymmetrically in the loop, as shown schematically in Fig. 5.5.

A frequency shifter is an optical element that produces a frequency shift to the optical wave passing through it. Such a frequency shifter can be realized using an acousto-optic modulator (AOM), in which a traveling acoustic wave is driven through an acousto-optic crystal, creating a moving refractive index grating that diffracts the light wave and Doppler shifts its frequency by an amount equal to the driving frequency. Depending on the diffraction order, both positive and negative frequency shifts can be realized. AOMs are also reciprocal in the sense that it works equally well when the lightwave reverses direction. Commercial AOMs with fiber pigtails are available, with an operation frequency range typically about ± 10 % of the intended driving frequency. (A larger operating range will cause excessive loss due to the larger diffraction angle affecting light coupling to the fiber). Currently, most common AOMs are designed to operate in the range of tens to hundreds of megahertz, independent of polarization. Gigahertz AOMs are also commercially available at a high cost, though their performance is typically polarization dependent. Techniques other than AOMs exist that can create a frequency shift to the lightwave. For example, phase or amplitude modulation of a lightwave will create frequency sidebands that are frequency shifted from the baseband signal, which can be isolated through filtering or signal processing. Single-side-band modulation scheme can also be used to create a frequency shift.

[2] The acoustic range under discussion here is quite low as we made simplifying assumptions to arrive at (5.24): (a) the acoustic wavelength is larger than the coil dimension, and (b) the acoustic frequency is lower than the characteristic frequency of the Sagnac loop.

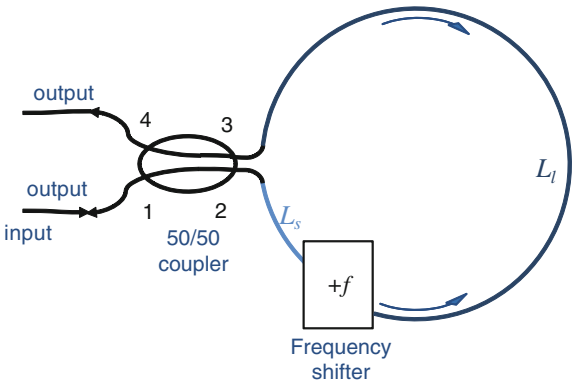

Fig. 5.5 A fiber Sagnac interferometer with a frequency shifter placed asymmetrically in the loop

These techniques, however, may not be reciprocal, and are also typically polarization dependent. Therefore, in many cases AOMs are preferred for performing the function of the frequency shift in frequency-shifted Sagnac interferometers (FSSI) due to their polarization independence and reciprocity.

The effect of the frequency shifter asymmetrically located in the Sagnac loop can be easily understood by examining the phase difference it introduces to the counter-propagating lightwaves. Consider the configuration illustrated in Fig. 5.5. The clockwise wave travels the long section of the loop (L_l) at this original frequency v_0 and completes the short section of the loop (L_s) at the up-shifted frequency $v_0 + f$. The counterclockwise wave does the reverse, traversing the short section at v_0 and the long section at $v_0 + f$. Therefore, the resulting phase difference is:

$$\Delta\phi = \frac{2\pi n}{c} f(L_l - L_s) \tag{5.25}$$

Note the phase difference is linearly dependent on the amount of frequency *shift* f, as well as the length imbalance $(L_l - L_s)$ marked by the location of the frequency shifter. This simple configuration turns out to be extremely versatile for a range of sensing applications. Below, we will review a few of them.

5.4.1 FSSI for Strain Sensing

One of the first applications of such a configuration is in long-range fiber strain sensing [28, 29]: A portion of the fiber or the entire loop can act as a strain gauge—when the strain changes, it creates a change in $\Delta\phi$. The method was implemented in combination with the phase-nulling technique developed for fiber gyros, with a close-loop feedback to adjust f such that the phase difference produced by the strain is canceled by the phase difference produced by varying f (Fig. 5.6). By differentiating (5.25) while keeping $\Delta\phi$ constant, we obtain:

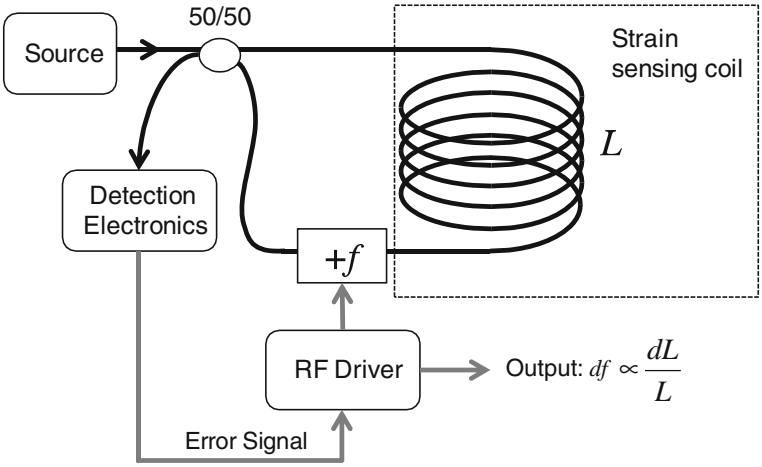

Fig. 5.6 Schematic arrangement of strain sensing based on FSSI, combined with the phase-nulling technique. *Black lines* indicate optical fibers while *grey lines* are electrical wires

$$\left|\frac{df}{f}\right| = \left|\frac{dL}{L}\right| \tag{5.26}$$

where dL/L is the strain experienced by the sensing coil of length L. The strain resolution is then determined by the frequency resolution, which can easily reach 10^{-8}, e.g., 1 Hz resolution for 100 MHz of driving frequency. It has been proposed that such high-sensitivity can be used for geological monitoring, such as monitoring Earth's slippage in earthquake prone regions, slippage of large structures built near cliffs, settling of dams, etc. [29].

5.4.2 FSSI for Fiber Length and Dispersion Measurements

Another usage of (5.25) is to measure fiber length (if n, the effective mode index of a single-mode fiber is accurately known), or to measure fiber dispersion $n(\lambda)$ (if the fiber length is known). Both aspects were explored by Abedin et al. [30] using a Faraday-mirror enabled Sagnac interferometer with an RF phase modulator (Fig. 5.7a), and later, by Qi et al. [31] using a simpler configuration with an AOM as frequency shifter (Fig. 5.7b). The latter demonstrated a large dynamic range for fiber length measurement and a relatively accuracy of 10^{-6} for long fibers. The chromatic dispersion (CD) measurement based on FSSI [31] yields results in close agreement with commercial CD instrument based on the modulation phase shift method, yet the measurement configuration of FSSI is much simpler, requiring only low-cost components and a cw laser source. (In the case of fiber length measurement, one can even use incoherent broadband source, due to the zero-path-length difference of the Sagnac interferometer).

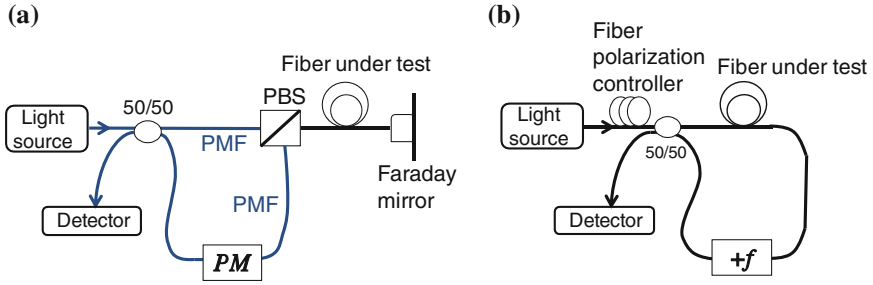

Fig. 5.7 Schematics for measuring fiber length and chromatic dispersion using FSSI used by **a** Abedin et al. [30], and **b** used by Qi et al. [31]. Schematics are simplified to show only the Sagnac loop. *PM* Phase modulator, *PMF* Polarization Maintaining Fiber

5.4.3 FSSI for Ultrasonic Sensing

The usefulness of the FSSI can be enhanced with a slight modification of its configuration to enable single-end access of the test fiber. Consider the configuration illustrated in Fig. 5.8a, one realizes that, even though it looks like a Mach–Zehnder interferometer, it is in fact a Sagnac interferometer. Because of the second 50/50 coupler, there are in fact four waves propagating in the loop (Fig. 5.8b). Two of them (taking Path 1 and Path 2) are the usual clockwise and counter clockwise waves, completing the loop after reflection from the end of the "dangling" fiber probe. Each of these two waves passes the frequency shifter once, and therefore both have frequency $v_0 + f$, and the resulting interference of the two waves is centered on DC. The other two waves (taking Path 3 and Path 4), trace back their respective original path after reflection. As a result, one wave (taking path 4) is frequency shifted twice to become $v_0 + 2f$, while the other (taking path 3) is not shifted at all. The interferences involving these two waves, with each other and with the other waves, are centered on f and $2f$, which can be easily filtered out.

Such single-ended FSSI configuration was first used for ultrasonic sensing by Fomitchov et al. [32] for both extrinsic and intrinsic sensing of ultrasonic waves. The "dangling" fiber in Fig 5.8a was used as a sensing probe. In the case of extrinsic sensing (Fig. 5.9a), the probe was placed close to the surface of a structure emitting ultrasonic signal, and light was collimated out of the fiber probe, reflected off the target surface, and collected back into the probe. In the intrinsic case (Fig. 5.9b), the fiber probe was embedded into the structure in the path of an ultrasonic wave.

There are two components to the phase difference between the counterpropagating waves in the Sagnac loop: one is the *static* phase difference $(\Delta\phi_s)$ due to the asymmetric location of the frequency shifter, the other is the *dynamic* phase difference $(\Delta\phi_d(t))$ due to the modulation of the ultrasonic wave on the phases of the two counter propagating waves:

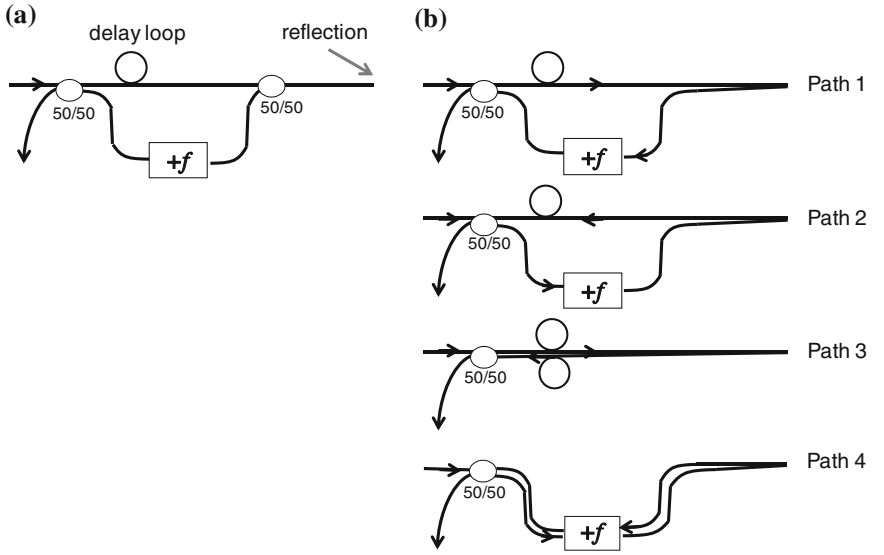

Fig. 5.8 **a** A frequency-shifted Sagnac interferometer with a dangling fiber probe and a delay loop used for ultrasonic sensing. **b** The four possible paths lightwave circulating in the loop can take

$$\Delta\phi(t) = \Delta\phi_s + \Delta\phi_d(t) \tag{5.27}$$

The static phase difference ($\Delta\phi_s$) above is given by (5.25), and is dependent on both the amount of frequency shift and the fiber length imbalance, which is controlled by the length of the delay loop. In this particular application, the static phase difference is tuned to the quadrature point (i.e., $\Delta\phi_s = m\pi + \pi/2$) to maximize the sensitivity to the ultrasound-induced phase change. The dynamic phase difference is dependent on the arrival time difference ($\Delta\tau$) between the clockwise (CW) and the counter clockwise (CCW) waves at the fiber probing tip, as shown in Fig. 5.9. For simplicity, suppose in Fig 5.8a the optical path of the upper half of the loop is exactly balanced out by that of the lower half of the loop *except* for the length of the delay loop, then $\Delta\tau = L_{delay}n/c$, and the dynamic phase difference can be calculated by:

$$\Delta\phi_d(t) = \phi_d(t) - \phi_d(t - \Delta\tau) \tag{5.28}$$

where ϕ_d is the phase modulation of the lightwave due to the ultrasound. In the case of extrinsic sensing, ϕ_d is a direct result of optical-path-length modulation due to surface displacement $\Delta s(t)$, i.e., $\phi_d(t) = 2\frac{2\pi\nu}{c}\Delta s(t)$. In the case of intrinsic sensing, the structural density change due to ultrasound can induce both a physical length change of the embedded fiber $\Delta l(t)$, as well as a refractive index change $\Delta n(t)$ due to the resulting strain. Therefore, $\phi_d(t) = \frac{2\pi\nu}{c}(n\Delta l(t) + l\Delta n(t))$, where n, l are the refractive index and the length of the embedded fiber, respectively.

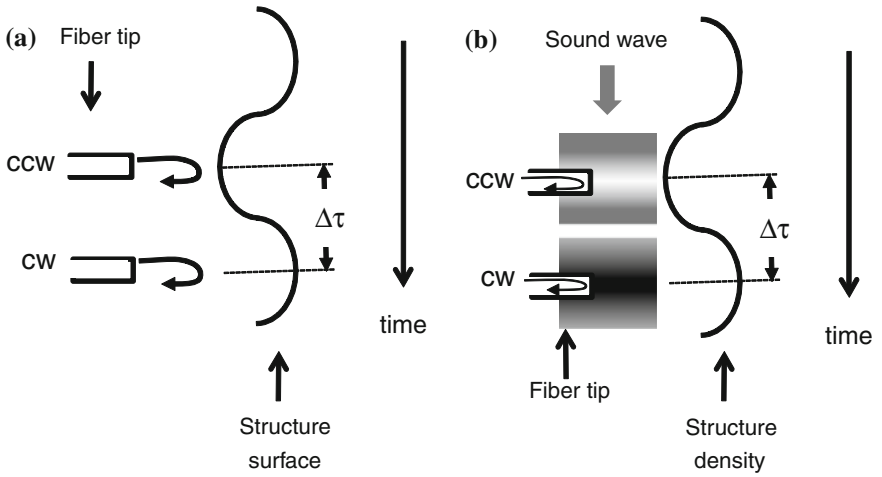

Fig. 5.9 Illustration of extrinsic and intrinsic ultrasonic sensing with a fiber probing tip from the single-ended FSSI, adapted from [32]. **a** In the case of extrinsic sensing, structural surface modified by ultrasonic wave modulates the optical path length for the CW and CCW wave at different instants of time. **b** In the case of intrinsic sensing, structural density modified by ultrasonic wave induces phase changes to the light traveling down the fiber tip

(Note, since the frequency shift is orders of magnitude smaller than the optical frequency, its effect on ϕ_d is negligible.) In either case, for a harmonic ultrasonic perturbation, ϕ_d will take the form of $A\sin(2\pi v_a t)$, where v_a is the acoustic frequency and A is the amplitude of the induced phase shift. Substituting this form into (5.28), and then using (5.27) and (5.6), we obtain the ac part of the output intensity as

$$I_{ac} \propto \cos\left\{2A\sin(\pi v_a \Delta\tau)\cos\left[2\pi v_a\left(t-\frac{\Delta\tau}{2}\right)\right]\right\} \tag{5.29}$$

Therefore, it can be seen that, in order to minimize signal distortion, $\Delta\tau$ (that is the delay length) should be tailored such that $v_a\Delta\tau = \frac{1}{2}$, v_a being the acoustic frequency it is designed to measure. It is therefore a narrow-band acoustic sensor that is insensitive to low-frequency vibrational and thermal noise. Fomitchov et al. [32] demonstrated that this ultrasonic sensing system can be used to detect structural flaws (such as flaws in an aircraft wheel) by monitoring the scattered ultrasonic waves.

5.4.4 FSSI for Sensor Array Interrogation

The frequency shifter used in the previous example (5.4.3) serves the function of biasing the measured phase difference at quadrature point. However, if we re-

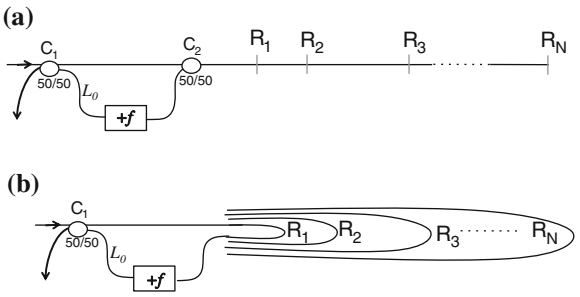

Fig. 5.10 **a** FSSI used for sensor array interrogation. R1, R2, …, Rn are weak reflective sensor elements along the fiber. **b** The equivalent "circuit" for FSSI in (**a**), with many Sagnac loops of different lengths

examine (5.25), we notice that because the phase shift is proportional to the length imbalance, FSSI can be used for location-resolved sensing. In particular, using the single-end access configuration, one can put a number of weakly reflecting optical sensors, such as FBGs, along the fiber, and use FSSI to resolve the reflections (which contain sensor information) sent back by different sensors according to their location (Fig. 5.10a). In other words, FSSI can be an effective sensor array interrogation technique. Qi et al. [33] used such a system to locate multiple weak reflections along a fiber link with high sensitivity (to −67 dB), while Ye et al. used it to interrogate an FBG sensor array [34] as well as a gas sensor array [35].

The system as illustrated in Fig. 5.10a in fact is equivalent to multiple Sagnac interferometers of different loop lengths connected in parallel (Fig. 5.10b). If one linearly sweeps the frequency of the frequency shifter, for the n^{th} loop, the output intensity I_n will vary sinusoidally with frequency:

$$I_n \propto 1 - \cos\left[\frac{2\pi n}{c}(L_n - L_0)f\right] \tag{5.30}$$

where L_n corresponds to the fiber length from the first coupler (C_1) to the n^{th} reflective sensor and back to the frequency shifter, and L_0 is the fiber length from the frequency shifter to the first coupler, which is common to all loops. One can immediately see that, with all reflective sensors present, the resulting output intensity is in the form of:

$$I_o \propto \sum_{n=1}^{N} A_n\left[1 - \cos\left[\frac{2\pi n}{c}(L_n - L_0)f\right]\right] \tag{5.31}$$

where A_n's are proportionality constants which may be different for different reflective sensors due to the difference in their reflectivities, as well as due to loss encountered along the fiber. Equation (5.31) is nothing other than the summation of sinusoids with different frequencies. As such, it can be easily "demodulated" by taking a Fourier transform. Figure 5.11 gives a sample of I_o (with Dc removed) and its Fourier transform.

It can be seen from Fig. 5.11 that each reflective sensor produces a peak in the Fourier domain, corresponding to its location. From (5.31), one can also utilize the

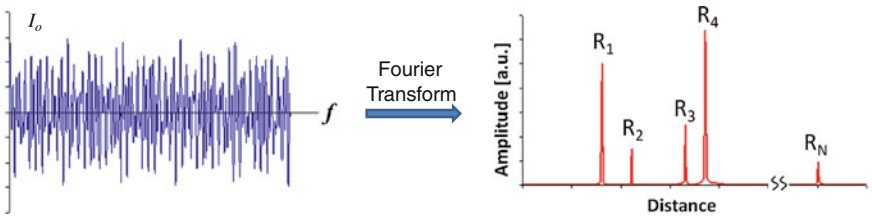

Fig. 5.11 An example of output intensity (**a**) and its Fourier transform (*horizontal axis converted to distance*) (**b**) obtained from a single-end-access FSSI system with multiple reflective sensors in an array

fact that the location information is solely contained in the phase (or frequency) of the signal, and the amplitude of the output intensity is independent of the location (ignoring fiber loss, which is mostly spectrally flat). The implication is twofold: (1) One can use the spectral information of the sensors (contained in the amplitude) for sensing, and (2) The spectral information sent back from different sensors will not interfere with each other, because the technique is capable of distinguishing the sensors (an the information sent back by the sensors) by their location. In other words, the technique allows different sensors to overlap spectrally. However, the problem of spectral shadowing has to be considered. Spectral shadowing refers to the fact that a sensor upstream in the array can modify the spectrum of the light passing through it, and so light reflected by the sensors downstream also contains the modification made by the sensor upstream. Hence, the sensor upstream can cast "spectral shadows" on the measured spectra of sensors downstream. Fortunately, this spectral shadowing effect can be easily corrected by a simple division of the amplitude [35]:

$$\frac{A_n}{A_{n-1}} \propto R_n(\lambda) \tag{5.32}$$

where $R_n(\lambda)$ is the spectral response of the n^{th} sensor.

Figure 5.12 shows an example of using FSSI for interrogating a senor array of 10 FBGs [34, 35] before (Fig.5.12a) and after (Fig. 5.12b) strain is applied to some of the FBGs. Spectral overlap of some of the FBGs is evident from the overall optical spectrum seen on the Optical Spectrum Analyzer. Nevertheless, the location-resolved results obtained from the Fourier transform of the output intensity as a function of frequency sweep, f, reveal the reflection spectra for individual gratings.

This sensor array interrogation technique based on FSSI, unlike the more commonly used wavelength-division-multiplexing (WDM) techniques, does not impose any spectral requirement on the sensing element. Because sensors are allowed to have spectral overlap, FSSI has two main advantages over the WDM technique: (1) The sensors (such as FBGs for strain or temperature sensing) can have a much larger dynamic range, as each sensor can occupy the entire bandwidth available to the system. (2) Fabrication tolerances on sensors are much more

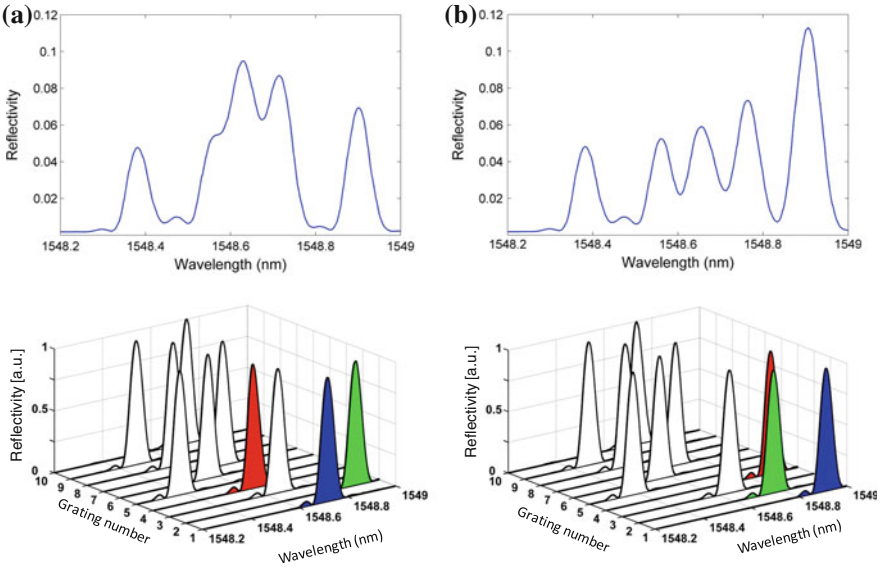

Fig. 5.12 Location-resolved FBG interrogation system showing individual FBG reflection spectrum **a** before, and **b** after applying strain to FBG 1 (*blue*), 2(*Green*), and 4 (*Red*). The spectra above were obtained on an Optical Spectrum Analyzer, indicating spectral overlap of some of the FBGs. Adapted from [36]

relaxed, and the sensors no longer need to be sorted or grouped according to their spectral characteristics, and they no longer have to avoid spectral overlapping. The FSSI technique also has an advantage over time-division-multiplexing (TDM) techniques for sensor array interrogation, because FSSI uses a continuous-wave light source and slow detectors, as opposed to the requirement for pulsed source and fast detection in TDM.

5.4.5 FSSI for Cavity Ring Down Sensing

The possibility of using frequency-shifted Sagnac interferometer for sensor multiplexing and for location-resolved sensing has opened up other opportunities for FSSI-based sensing. One novel extension is the use of FSSI for continuous-wave cavity ring down sensing.

Cavity ring down (CRD) technique was developed to measure minute quantities of target molecules using light, by measuring the minute optical loss due to absorption or scattering by the molecules. It conventionally employs a laser pulse circulating in a high-finesse optical cavity (the ring-down cavity) containing the target material (usually gas or liquid). The pulse peak intensity coupled out of the cavity is slightly reduced each time the pulse completes a cavity round trip due to

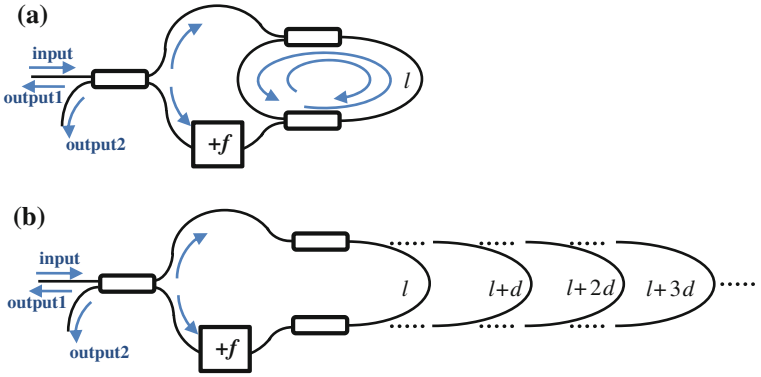

Fig. 5.13 **a** A Frequency-shifted Sagnac interferometer with a nested ring-down cavity, which is equivalent to (**b**) **b** an FSSI system with Sagnac loops of different lengths connected in parallel

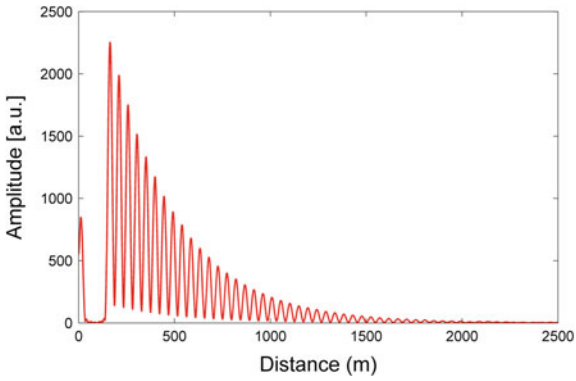

Fig. 5.14 An example of output intensity decay as a function of distance (or the number of cavity round trips), obtained through Fourier transform of the measured intensity at the FSSI output as a function of frequency sweep. The decay length is used to calculate the cavity loss

absorption or scattering of the target material. If the decay of pulse intensity is recorded as a function of time, the cavity loss can be calculated from the characteristic decay time of the pulse train [37]. More recently, CRD techniques that use fiber cavity for sensing has also been developed [38–40].

Using the FSSI scheme, however, cavity ring down measurements can be done with continuous-wave sources [41], without the need for pulses or temporal modulation. This is because FSSI can distinguish light going through different path lengths in the Sagnac interferometer. Consider the schematic illustrated in Fig. 5.13a, with its equivalent "circuit" shown in Fig. 5.13b. Note Fig. 5.13b is essentially the same as Fig. 5.10b, except that, in the case of nested ring-down cavity, the number of Sagnac loops in parallel depends on the cavity loss. Therefore, we can apply the same analysis we used for analyzing FSSI with multiple reflections to the case of FSSI with a nested ring-down cavity. The output intensity will have the same form as (5.31), and its Fourier transform will present a series of equally spaced intensity peaks with a characteristic decay distance (Fig. 5.14), from which the cavity loss can be determined:

$$Cavity\ Loss\,(\mathrm{dB}) = 10\log e\,\frac{cavity\ length}{decay\ length} \tag{5.33}$$

where the decay length is defined as the distance required for the peak intensity to drop to 1/e of its initial value.

This method has been applied to measure fiber bend loss [41], the concentration of 1-octyne in a decane solution, and the refractive index of sodium chloride solutions [42]. Preliminary results demonstrated that a refractive index sensitivity of 1×10^{-4} RIU can be obtained with this method.

5.5 Fiber Optic Sagnac Current Sensor

The magnetooptic property of optical fibers [43] can be utilized for fiber-optic current sensing, based on the well-known Faraday effect, or Faraday rotation. The effect is usually described as a rotation of the azimuth angle of the plane of polarization for a linearly polarized light propagating in a magnetic field **H** for a path length L:

$$\Delta\theta = V\int_0^L \mathbf{H}\cdot\mathbf{dl} \tag{5.34}$$

where $\Delta\theta$ is the amount of azimuthal rotation, and V is the Verdet constant which is material dependent. For standard silica fibers, the magneto-optic effect is very weak, with V on the order of a few μrad/A. Fibers doped with certain rare-earth element can have significantly enhanced Verdet constants [44].

An alternate way to describe the Faraday effect, which is more pertinent to the Sagnac current senor, is to regard the effect as a circular birefringence produced by the magnetic field. A circular birefringence results in different refractive indices for the two orthogonal circular polarizations: Right-handed circular polarization (RHCP) and left-handed circular polarization (LHCP). A linear polarization can be considered as the superposition of co-propagating RHCP and LHCP with a relative phase difference that equals to twice the azimuth of the linear polarization. Therefore, the phase difference between RHCP and LHCP generated by the magnetic field is

$$\Delta\phi = 2V\int_0^L \mathbf{H}\cdot\mathbf{dl} \tag{5.35}$$

For a fiber coil of N turns wrapped around a wire carrying current I, the phase difference can be further enhanced by a factor of N:

$$\Delta\phi = 2NV\oint \mathbf{H}\cdot\mathbf{dl} = 2NVI \tag{5.36}$$

Equation (5.36) is the basis for interferometric fiber-optic current sensing.

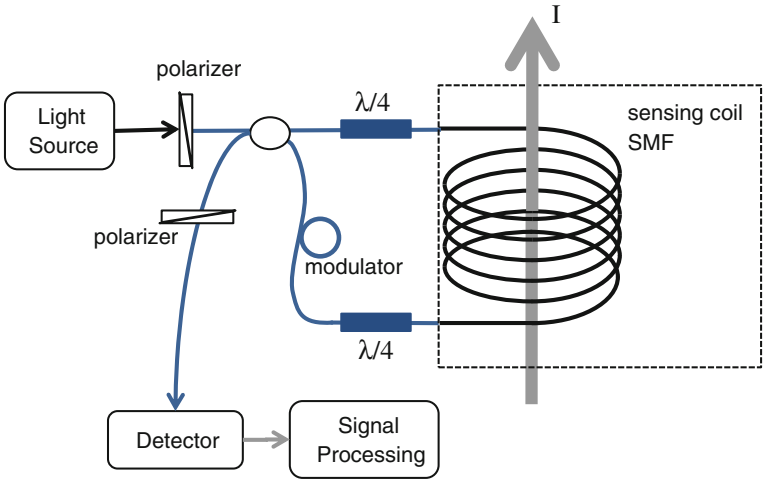

Fig. 5.15 A simple schematic for implementing current sensing using a Sagnac interferometer (*Blue lines* indicate polarization-maintaining fibers). $\lambda/4$ indicates fiber-based Quarter-wave plates

The Faraday effect is non-reciprocal, and therefore Sagnac interferometers are ideally suited for measuring such an effect. In addition, signal-processing techniques developed for the IFOG to enhance sensitivity and improve noise immunity can be readily "borrowed" to produce high-sensitivity fiber-based Sagnac current sensors. In a Sagnac interferometer, the Faraday effect introduces a phase difference (5.36) between *counter-propagating* lightwaves of the *same* handedness.[3] Thus, measuring this phase difference involves converting a linearly polarized lightwave into circular polarization (RHCP or LHCP) using a quarter-wave plate before launching it into the sensing coil, and converting the circularly polarized light back to the original linear polarization at the output of the sensing coil, and then interfering them to measure the phase difference [45]. As in the case for the IFOG, a phase modulator can be added to the Sagnac loop to "bias" the measured phase angle to the quadrature point for better sensitivity (see Sect. 5.1). A simple schematic is illustrated in Fig. 5.15.

Sagnac interferometry is only one of the many different techniques for fiber-optic current sensing, i.e., measuring $\Delta\theta$ (5.34) or $\Delta\phi$ (5.36), but it has the advantage of being sensitive only to non-reciprocal effects (of which Faraday effect is one), while being immune to reciprocal effects, as long as the reciprocal effects are not fast time-varying (see Sect. 5.3) and the Sagnac interferometer is truly single-mode. In practice, however, it turns out that meeting these two

[3] Note that the handedness of the circularly polarized light is in reference to its propagation direction. Using a common convention, the E field of a right-hand circularly polarized lightwave rotates clockwise (in time) as seen by an observer facing the incoming lightwave.

conditions can be quite a challenge, especially when a measurement accuracy of better than 0.1 % is required, for example, for revenue metering applications.

First, linear birefringence in a standard optical fiber (caused by imperfection in the fiber fabrication process, and exasperated by bending and winding in the coiling process) makes it impossible for the launched circularly polarized light to propagate in a *single* circular polarization mode throughout the sensing coil. Somewhere along the way, some power will couple back and forth between the two circular polarization modes, depending on local birefringence, which is in general not uniform (varies with location) and not constant (varies with time.) The net effect is that different amounts of polarization noise are added to the two counter-propagating waves, resulting in noise and instability in the phase measurement. It is worth noting that this effect cannot be completely prevented even if one polarization mode is completely filtered out at the output, because mode coupling back and forth destroys reciprocity [45].

To alleviate this problem created by the fiber's linear birefringence, one can either significantly reduce the linear birefringence of the fiber coil, or significantly overwhelm the linear birefringence of the fiber, by using fiber with a very high circular birefringence. In the former case, fiber coils can be annealed [46] to release stress and therefore stress-induced birefringence, or one can use special fibers that produce less birefringence under stress [47], and/or use special fiber packaging to reduce stress [48]. Alternatively, one can use fibers with a very high intrinsic circular birefringence such that the fiber is essentially a "circular polarization maintaining" fiber, in which weak linear birefringence cannot cause noticeable mode coupling (in the same way that regular polarization-maintaining fiber reduces mode coupling of linearly polarized modes). Quantitative analysis showed that adding circular birefringence to the sensing coil stabilizes the scale factor [49]. A high circular birefringence can be created by drawing an elliptical-core fiber while spinning the preform at a high rate during the drawing process (spun fiber) [50]. It should be noted that a high circular birefringence of the fiber does not affect the sensitivity of the Sagnac interferometer, since the circular birefringence induced by the Faraday effect is *additive* to the existing circular birefringence of the fiber.

Second, time-varying perturbations (such as vibration and thermal fluctuations) induce non-reciprocal effects when the perturbation site is asymmetrically located with respect to the center of the Sagnac fiber loop (See Sect. 5.3). An ingenious solution was developed (independently by Blake et al. [51] and by Frosio and Dandliker [52]) to mitigate this problem by using a reflection topology and sending both RHCP and LHCP down the sensing fiber, as illustrated in Fig. 5.16.

In this reflection Sagnac topology, linearly polarized light is split into two equal components of two orthogonal linear polarizations (via a 45° splice) and launched into a polarization-maintaining fiber. Both polarizations are converted into circular polarizations (which are orthogonal to each other) through a short piece of birefringe fiber that acts as a quarter-wave plate, and travel down the sensing coil. Upon reflection, RHCP becomes LHCP and vice versa. The polarizations then travel exactly the same path back up the sensing coil. They are then converted

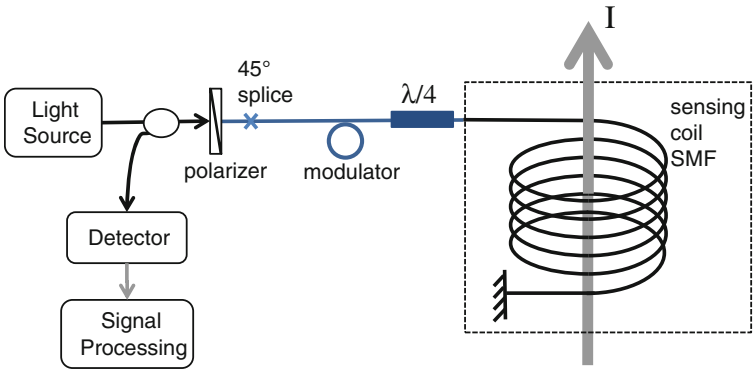

Fig. 5.16 A simplified reflection Sagnac interferometer used for current sensing (*Blue lines* indicate polarization-maintaining fibers). *λ*/4 indicates fiber-based Quarter-wave plates

back into linear polarizations and brought to interfere with each other through the polarizer that is oriented at 45° with respect to the two linear polarizations (because of the 45° splice).

There are two distinct advantages to this reflection topology. First, the total accumulated phase difference between the two circular polarizations due to the Faraday effect is doubled compared to (5.36), i.e., $\Delta\phi = 4NVI$. Second, since the two interfering waves are now *co-propagating* along the fiber, any time-varying perturbation (as long as it is not polarization dependent) affects both waves equally, and at the same time, so it would not add to any phase difference between the two. This topology has in fact been adopted by commercial fiber-optic current sensors [48].

A third challenge, which is common to all optical current sensors, is to overcome or compensate the temperature dependence of the Verdet constant, which has a coefficient of $(1/V)dV/dT = 0.7 \times 10^{-4}/^{\circ}\mathrm{C}$ [53]. It may seem small, but in order to qualify for field applications, current sensors are required to be accurate over a temperature range of −40–8 °C [54]. Without temperature compensation, the accuracy cannot reach <0.1 %, which is usually required by the power industry. Though one can add a temperature sensor co-located with the current sensor and calibrate the scale factor of the current sensor according to its temperature, such an arrangement would add considerable cost and complexity to the sensing system. Passive compensation is much more desirable if it can be done to specification. It turns out that the fiber-based quarter-wave plate(s) (shown in Figs. 5.15 and 5.16) also has a temperature dependence which can be used to compensate for the effect due to the variation of Verdet constant with temperature. An ideal zeroth-order quarter-wave plate produces exactly $\pi/2$ phase shift (called retardation) between the two orthogonal linear polarizations that are aligned to its axes. However, the amount of retardation will be a function of temperature as the refractive index of the quarter-wave plate is temperature dependent, causing the

actual retardation to deviate from $\pi/2$. If we use ε_2 and ε_2 to denote the deviations of retardation from $\pi/2$ for the two fiber-base quarter-wave plates shown in Fig. 5.15, then the modified phase difference ($\Delta\phi_\varepsilon$) between the two counter-propagating waves is [55]:

$$\Delta\phi_\varepsilon = \tan^{-1}\left(\frac{(\cos\varepsilon_1 + \cos\varepsilon_2)\sin\ \Delta\phi}{(1+\cos\varepsilon_1\cos\varepsilon_2)\cos\ \Delta\phi - \sin\varepsilon_1\sin\varepsilon_2\cos 2\chi}\right) \tag{5.37}$$

where $\Delta\phi$ is given by (5.36), and χ is the angle between the fast axes of the two quarter-wave plates. For the reflection topology shown in Fig. 5.16, where only one quarter-wave plate is used, the modified phase difference ($\Delta\phi_\varepsilon$) between the two co-propagating waves when they interfere at the detector is [48]:

$$\Delta\phi_\varepsilon = \tan^{-1}\left(\frac{2\cos\varepsilon\sin(2\Delta\phi)}{(1+\cos^2\varepsilon)\cos(2\Delta\phi) - \sin^2\varepsilon}\right) \tag{5.38}$$

For small ε, (5.38) is approximated by $\Delta\phi_\varepsilon \approx 2\Delta\phi/\cos\varepsilon$, or

$$\Delta\phi_\varepsilon - 2\Delta\phi \approx 2\Delta\phi\varepsilon^2 \tag{5.39}$$

Therefore, the phase deviation from the expected $2\Delta\phi$ is a quadratic function of ε, which in turn is approximately a linear function of temperature. Pre-biasing ε can cause $\Delta\phi_\varepsilon$ to either increase or decrease with temperature (depending on which branch of the quadratic curve ε is on). Furthermore, the rate of increase or decrease with temperature can also be determined by judiciously choosing the range of ε. To compensate for a positive temperature coefficient of the Verdet constant, one needs to put a positive bias on ε (i.e., larger than $\pi/2$ retardation). More detailed analysis reveals that a retardation of 0.558π would compensate the temperature variation of the Verdet constant almost exactly [48].

Combining all of the abovementioned engineering solutions—reducing unwanted non-reciprocal effects (due to linear birefringence and time-varying perturbations) and compensating for temperature dependence of the Verdet constant, as well as applying signal processing techniques developed for IFOG (such as quadrature bias or phase nulling), today's state-of-the-art fiber-optic current sensors based on Sagnac interferometers can achieve accuracies well within 0.1 % [55]. Compared to traditional electric current transformers, fiber-optic current sensors can have a much higher measurement bandwidth (for AC currents), limited only by the transit time of light in the sensing coil. Furthermore, they have the attractive advantages of being immune to electromagnetic interference, requiring no electrical insulation, low cost, and light weight, particularly for high voltage systems. (A conventional electrical current transformer sensor requires costly and bulky oil-filled insulation tower, and for a 500 kV system, the sensor can weigh as much as 7,000 kg! [56]).

5.6 Summary

We have shown in this chapter that the fiber Sagnac interferometer is a versatile tool for a range of sensing applications. The key feature of the Sagnac loop is that it measures non-reciprocal effects (e.g. Sagnac effect, Faraday effect, time-varying effects such as those caused by acoustic signals), while being insensitive to reciprocal effects. Topological modifications of the Sagnac interferometer leads to novel applications such as sensor multiplexing, as well as spatial-domain continous-wave cavity-ring-down sensing.

Acknowledgment The author of this Chapter wishes to thank Fei Ye and Dr. Bing Qi for providing results and diagrams used in Sect. 5.4.4.

References

1. Sagnac G (1913) Sur la preuve de la réalité de l'éther lumineux par l'expérience de l'interférographe tournant. Comptes Rendus 157:1410–1413 (translated as "On the proof of the reality of the luminiferous aether by the experiment with a rotating interferometer")
2. Vali V, Shorthill RW (1976) Fiber ring interferometer. Appl Opt 15:1099–1100
3. Regarding the origin of the $\pi/2$ phase difference between the two output ports of a direction coupler, see, for example, Chap. 13.3. In: Yariv A, Yeh P (2007) Photonics, 6th edn. Oxford University Press
4. Arditty HJ, Lefèvre HC (1982) Theoretical basis of Sagnac effect in fiber gyroscope, vol 32. Series in Optical Sciences, Springer, Berlin, pp 44–51
5. Lefèvre HC, Arditty HJ (1982) Electromagnétisme des milieux diélectriques linéaires en rotation et application à la propagation d'ondes guidées. Appl Opt 21:1400–1409 in French
6. Chapter 2 of Lefèvre H (1993) The fiber-optic gyroscope. Artech House, Inc., Boston
7. see, for example, Blake J (1998) Fiber optic gyroscopes. In: Grattan KTV, Meggitt BT (eds) Optical fiber sensor technology, vol 2. Chapman and Hall, New York
8. Sanders GA, Szafraniec B, Liu R-Y, Bielas MS, Strandjord LK (1995) Fiber-optic gyro development for a broad range of applications. In: DePaula RP, Berthold III JW (eds) Proceedings of SPIE, vol 2510. Fiber Optic and Laser Sensors XIII, Ramon, pp 3–11
9. Böhm K, Russer P, Weidel E, Ulrich R (1981) Low-noise fiber-optic rotation sensing. Opt Lett 6:64–66
10. Bergh RA, Lefevre HC, Shaw HJ (1981) All-single-mode fiber-optic gyroscope with long-term stability. Opt Lett 6:502–504
11. Chapter 3 of Lefèvre H (1993) The fiber-optic gyroscope, Artech House, Inc., Boston
12. Davis JL, Ezekiel S (1982) Closed-loop low-noise fiber-optic rotation sensor. Opt Lett 6:505–507
13. Kim BY, Shaw HJ (1984) Gated phase-modulation feedback approach to fiber-optic gyroscopes. Opt Lett 9:263–265
14. Cutler CC, Newton SA, Shaw HJ (1980) Limitation of rotation sensing by scattering. Opt Lett 5:488–490
15. Burns WK, Moeller RP (1983) Rayleigh backscattering in a fiber gyroscope with limited coherence sources. IEEE J Lightwave Technol 1:381
16. Ezekiel S, Davis JL, Hellwarth RW Intensity dependent nonreciprocal phase shift in a fiberoptic gyroscope. In: Springer Series in Optical Sciences, vol 32. pp 332–336 (SPIE MS 8 pp 308–312)

17. Chapter 7 of Lefèvre H (1993) The fiber-optic gyroscope. Artech House, Inc., Boston
18. Saleh BEA, Teich MC (2007) Fundamentals of photonics, 2nd edn. Wiley-Interscience, Chap. 19
19. Bucaro J, Dardy HD, Carome EF (1977) Optical-fiber acoustic sensor. Appl Opt 16:1761–1762
20. Cole JH, Johnson RL, Bhuta RG (1977) Fiber-optic detection of sound. J Acoust Soc Am 62:1136–1138
21. Cahill RF, Udd E (1983) Optical acoustic sensor. US Patent 4,375,680, 1 Mar 1983
22. Udd E (1983) Fiber-optic acoustic sensor based on the Sagnac interferometer. Proc SPIE 425:90–95
23. Frigo NJ (1983) Compensation of linear sources of nonreciprocity in Sagnac interferometers. Proc SPIE 412:268–271
24. Schupe DM (1980) Thermally induced nonreciprocity in the fiber-optic interferometer. Appl Opt 19:654
25. Kråkenes K, Bløtekjaer K (1989) Sagnac interferometer for underwater sound detection: noise properties. Opt Lett 14:1152–1154
26. N. Lagakos, E.U. Schnaus, H.H. Cole, J. Jarznski, and J.A. Bucaro, (1982) Optimizing fiber coatings for interferometric acoustic sensors. IEEE Trans Microw Theory Tech MTT-30:529
27. Lagakos N, Bucaro JA (1981) Pressure desensitization of optical fibers. Appl Opt 20:2715
28. Michal RJ, Udd E, Theriault JP (1986) Derivative fiber-optic sensors based on phase-nulling optical gyro development. Proc SPIE 719:150–154
29. Udd E, Blom RG, Tralli DM, Saaski E, Dokka R (1994) Application of the Sagnac Interferometer based strain sensor to an earth movement detection system. Proc SPIE 2191:127–136
30. Abedin KS, Hyodo M, Onodera N (2000) A novel Sagnac Interferometer for characterizing the chromatic dispersion and length of installed optical fiber. Proc SPIE 4087:169–176
31. Qi B, Tausz A, Qian L, Lo HK (2005) High-resolution, large dynamic range fiber length measurement based on a frequency-shifted asymmetric Sagnac interferometer. Opt Lett 30:3287–3289
32. Fomitchov PA, Krishnaswamy S, Achenbach JD (2000) Extrinsic and intrinsic fiber optic Sagnac ultrasound sensors. Opt Eng 39:1972–1984
33. Qi Bing, Qian Li, Tausz Andrew, Lo Hoi-Kwong (2006) Frequency-shifted Mach-Zehnder Interferometer for Locating Multiple Weak Reflections along a Fiber Link. IEEE Photonics Technol Lett 18(1):p295
34. Ye F, Qian L, Liu Y, Qi B (2008) Using frequency-shifted interferometry for multiplexing a fiber Bragg grating array. IEEE Photonics Technol Lett 20:1488–1490
35. Ye F, Qian L, Qi B (2009) Multipoint chemical gas sensing using frequency-shifted interferometry. J Lightwave Technol 27:5356–5364
36. Ye F (2008) Frequency-shifted interferometry for multipoint chemical gas sensing and other sensing applications. Thesis, University of Toronto
37. Berden G, Peeters R, Meijer G (2000) Cavity ring-down spectroscopy: experimental schemes and applications. Int Rev Chem 19:565–607
38. Waechter H, Litman J, Cheung AH, Barnes JA, Loock H-P (2010) Chemical sensing using fiber cavity ring-down spectroscopy. Sensors 10:1716–1742
39. Tarsa PB, Rabinowitz P, Lehmann KK (2004) Evanescent field absorption in a passive optical fiber resonator using continuous-wave cavity ring-down spectroscopy. Chem Phys Lett 383:297–303
40. Wang C, Herath C (2010) High-sensitivity fiber-loop ringdown evanescent-field index sensors using single-mode fiber. Opt Lett 35:1639–1641
41. Ye F, Qi B, Qian L (2011) Continuous-wave fiber cavity ringdown measurements using frequency-shifted interferometry. Opt Lett 36:2080–2082
42. Ye F, Zhou C, Qi B, Qian L (2013) Continuous-wave fiber loop cavity ring-down evanescent-field sensing based on frequency-shifted interferometry. Sensor Actuat B Chem 184:150–155

43. Smith AM (1978) Polarization and magnetooptic properties of single mode optical fiber. Appl Opt 17:52–56
44. Ballato J, Snitzer E (1995) Fabrication of fibers with high rare-earth concentrations for Faraday isolator applications. Appl Opt 34:6848–6854
45. Nicati PA, Robert Ph (1988) Stabilized current sensor using Sagnac interferometer. J Phys E: Sci Instrum 21:791–796
46. Tang D, Rose AH, Day GW, Etzel SM (1991) Annealing of linear birefringence in single-made fiber coils: application to optical fiber current sensor. IEEE J Lightwave Technol 9:1031–1037
47. Kurosawa K, Yoshida S, Sakamoto K, Masuda I, Yamashita T (1994) An optical fiber-type current sensor utilizing the Faraday effect of flint glass fiber. Proc SPIE 2360:24–27
48. Bohnert K, Gabus P, Kostovic J, Brändle H (2005) Optical fiber sensors for the electric power industry. Opt Lasers Eng 43:511–526
49. Short SX, de Arruda JU, Tselikov A, Blake JN (1998) Elimination of birefringence induced scale factor errors in the in-line Sagnac interferometer current sensor. IEEE J Lightwave Technol 16:1844–1850
50. Laming RI, Payne DN (1989) Electric current sensors employing spun highly birefringent optical fibers. IEEE J Lightwave Technol 7:2084–2094
51. Blake J, Tantaswadi P, de Carvalho RT (1996) In-line Sagnac interferometer current sensor. IEEE Trans Power Delivery 11:116–121
52. Frosio G, Dändliker R (1994) Reciprocal reflection interferometer for a fiber-optic Faraday current sensor. Appl Opt 33:6111–6122
53. Williams PA, Rose AH, Day GW, Milner TE, Deeter MN (1991) Temperature dependence of the Verdet constant in several diamagnetic glasses. Appl Opt 30:1176–1178
54. International Standard for Current Transformers (1987) IEC 185. International Electrotechical Commission (IEC), Geneva, Switzerland
55. Bohnert K, Brändle H, Brunzel MG, Gabus P, Guggenbach P (2007) Highly accurate fiber-optic DC current sensor for the electrowinning industry. IEEE Trans Ind Appl 43:180–187
56. Day GW, Rochford KB, Rose AH (1996) Fundamentals and problems of fiber current sensors. In: Proceedings of OFS-11: 11th international conference on optical fiber sensors

Chapter 6
Principles of Optical Coherence Tomography

Kang Zhang and Jin U. Kang

6.1 Optical Coherence and Interference

Optical coherence is a property describing a degree of correlation between phases of optical waves. An optical wave with arbitrary amplitude and phase can be described by a complex wave function $a(\mathbf{r}, t)$. The time-averaged intensity $I(\mathbf{r}, t)$ can be defined as,

$$I(\mathbf{r}, t) = \langle a(\mathbf{r}, t) a^*(\mathbf{r}, t) \rangle \tag{6.1}$$

where $\langle\ \rangle$ is the averaging operation. When $I(\mathbf{r}, t)$ is time-invariant, the optical wave is viewed as statistically stationary.

The spatial and temporal coherence of a stationary optical wave $a(\mathbf{r}, t)$ can be mutually described by a correlation function $G(\mathbf{r}_1, \mathbf{r}_2, \tau)$ as:

$$G(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle a(\mathbf{r}_1, t) a^*(\mathbf{r}_2, t + \tau) \rangle \tag{6.2}$$

The complex degree of coherence function $g(\mathbf{r}_1, \mathbf{r}_2, \tau)$ can be then defined as,

$$g(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{G(\mathbf{r}_1, \mathbf{r}_2, \tau)}{\sqrt{I(\mathbf{r}_1) I(\mathbf{r}_2)}} \tag{6.3}$$

In situations where two waves' spatial positions coincide, e.g. $\mathbf{r} = \mathbf{r}_1 = \mathbf{r}_2$, Eq. (6.2) becomes a temporal coherence function $G(\tau)$.

$$G(\tau) = \langle a(t) a^*(t + \tau) \rangle \tag{6.4}$$

K. Zhang (✉)
GE Global Research, 1 Research Circle, Niskayuna, New York 12309, USA
e-mail: kangzhang2011@gmail.com

J. U. Kang
Electrical and Computer Engineering Department, Johns Hopkins University,
N. Charles Street 3400, Baltimore 21218, USA
e-mail: jkang@jhu.edu

J. U. Kang (ed.), *Fiber Optic Sensing and Imaging*,
DOI: 10.1007/978-1-4614-7482-1_6,

Then the temporal complex degree of coherence $g(\tau)$ can be written as:

$$g(\tau) = \frac{\langle a(t)a^*(t+\tau)\rangle}{\langle a(t)a^*(t)\rangle} \tag{6.5}$$

The degree of coherence of a light source can be described qualitatively using $g(\tau)$ as:

$$|g(\tau)| = 1 \quad \text{Perfectly coherent} \tag{6.6}$$

$$0 < |g(\tau)| < 1 \quad \text{Partially coherence} \tag{6.7}$$

$$|g(\tau)| = 0 \quad \text{Totally incoherent} \tag{6.8}$$

The coherence time τ_c and corresponding coherence length are defined as,

$$\tau_c = \int_{-\infty}^{\infty} |g(\tau)|^2 d\tau \tag{6.9}$$

$$l_c = c\tau_c \tag{6.10}$$

The spectral function of a light source can be described by the power spectral density (PSD) as,

$$S(v) = \int_{-\infty}^{\infty} G(\tau)e^{j2\pi v\tau} d\tau \tag{6.11}$$

Optical coherence applications use a wide range of interferometers that introduce an optical delay time, τ by inducing an optical path-length (OPL) difference, $\Delta L = c\tau$.

6.2 Time Domain Optical Coherence Tomography

The first time domain optical coherence tomography (TD-OCT) system was based on a low-coherence Michelson interferometer [1]. Figure 6.1 illustrates the principle of a time domain system, where a low-coherent light source with a Gaussian PSD, $S(v)$ is used as the light source. The image sample can be modeled as a multi-layered structure with interfaces R_i. A reference mirror moves along the axial direction and interferes with each interface over the image range ΔD in depth.

For a low-coherent light source, the coherence function can be modeled as,

$$G(\tau) = G_a(\tau)e^{-j2\pi v_0\tau} = |G_a(\tau)|e^{-j2\pi v_0\tau - \varphi_a(\tau)} \tag{6.12}$$

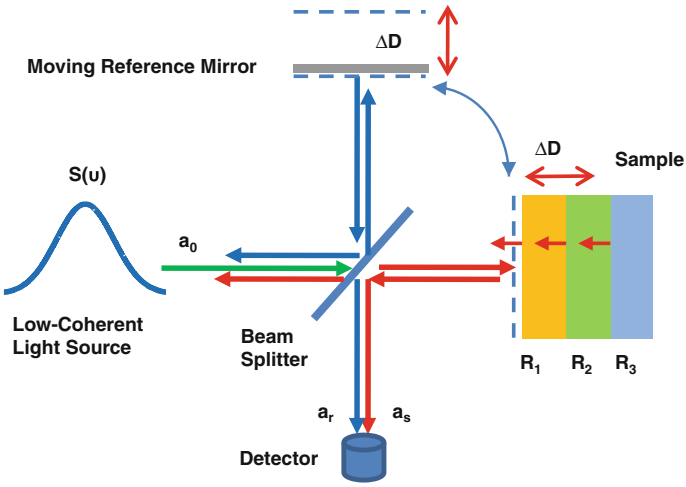

Fig. 6.1 Principle of a TD-OCT system

where v_0 is the central frequency of $S(v)$ [2]. Here the time delay τ relates to the spatial displacement of an interface relevant to the fixed reference plane by $\tau = \Delta D/c$.

The detector measures the inference between the two optical waves a_r and a_s which come from the reference mirror and the sample respectively. Considering Eq. (6.4), the detected average intensity can be written as,

$$\begin{aligned} I &= \left\langle |a_r + a_s|^2 \right\rangle = \left\langle |a_r|^2 \right\rangle + \left\langle |a_s|^2 \right\rangle + \left\langle a_r^* a_s \right\rangle + \left\langle a_r a_s^* \right\rangle \\ &= I_r + I_s + 2\mathrm{Re}\{G(\tau)\} = I_r + I_s + 2|G(\tau)|\cos\varphi(\tau) \end{aligned} \tag{6.13}$$

Assuming that the spectral width satisfies $\Delta v_c = 1/\tau_c \ll v_0$, $|G_a(\tau)|$ and $\phi_a(\tau)$ are slow varying respect to τ_c, and $|G_a(\tau)| = |G(\tau)|$, for a single interface with reflectivity R_i, Eq. (6.13) can be further expressed as,

$$\begin{aligned} I &= I_r + I_s + 2R_i|G(\tau - \tau_i)|\cos[2\pi v_0(\tau - \tau_i) + \varphi_a(\tau - \tau_i)] \\ &= I_r + I_s + \frac{2R_i}{c}|G(\Delta D - \Delta D_i)|\cos[2\pi v_0(\Delta D - \Delta D_i) + \varphi_a(\Delta D - \Delta D_i)] \end{aligned} \tag{6.14}$$

Here gives the response of a single interface, which presents the point spread function (PSF) of the OCT imaging system. For multiple interfaces, Eq. (6.14) turns to,

$$\begin{aligned} I &= I_r + \sum_i I_s \\ &+ \frac{2}{c}\sum_i R_i|G(\Delta D - \Delta D_i)|\cos[2\pi v_0(\Delta D - \Delta D_i) + \varphi_a(\Delta D - \Delta D_i)] \\ &+ \frac{2}{c}\sum_{ij} R_i R_j\left|G(\Delta D_i - \Delta D_j)\right|\cos[2\pi v_0(\Delta D_i - \Delta D_j) + \varphi_a(\Delta D_i - \Delta D_j)] \end{aligned} \tag{6.15}$$

where the last term results from autocorrelation between interfaces.

Equation (6.15) presents the basic 1-D depth information inside the sample, which is referred to as an "A-scan". Correspondingly, a 2-D cross-sectional image can be achieved by moving the light spot across a line on the sample, which is referred to as a "B-scan"; a 3-D volumetric image can be achieved by moving the light spot across a region on the sample, which is referred to as a "C-scan". For some special cases, the light spot stays in the same position on the sample, and the series of A-scan over time is referred to as an "M-scan".

For a Gaussian spectral shape of $S(v)$, according to Eq. (6.11), $G(\tau)$ can be expressed as,

$$G(\tau) = \int_0^\infty S(v) e^{-j2\pi v\tau} dv \tag{6.16}$$

With a normalized $S(v)$ [3], we have,

$$S(v) = \frac{2\sqrt{\ln 2/\pi}}{\Delta v_c} e^{-4\ln 2\left(\frac{v-v_0}{v}\right)^2} \tag{6.17}$$

Then by Eqs. (6.16) and (6.17), $G(\tau)$ can be expressed as,

$$G(\tau) = e^{-\left(\frac{\pi \Delta v \tau}{2\sqrt{\ln 2}}\right)^2} * e^{-j2\pi v_0 \tau} \tag{6.18}$$

6.3 Fourier Domain Optical Coherence Tomography

Fourier domain optical coherence tomography (FD-OCT) has been advancing rapidly since the first demonstration [4], and now dominates the OCT research and market with superior A-scan rates in the order of tens of kHz up to MHz, as well as a much higher sensitivity—usually two order of magnitude higher than the TD-OCT [5, 6]. FD-OCT can be categorized into two different types: spectrometer based OCT (Fig. 6.2a) and swept-source based OCT (Fig. 6.2b), which follow the same interference principle. However, during the rest of this chapter spectrometer-based SD-OCT will be simply referred to as "FD-OCT".

Figure 6.3 presents the layout and signal processing steps of FD-OCT. Each interface of the multi-layered sample produces a spectral interferogram, $S_i(v)$. The detected spectrum is a superposition of the interferograms from all the layers. The A-scan signal can be then reconstructed by taking the Fourier transform of the combined interferogram, e.g. measured spectrum from the spectrometer.

The measured spectrum can be expressed as [7],

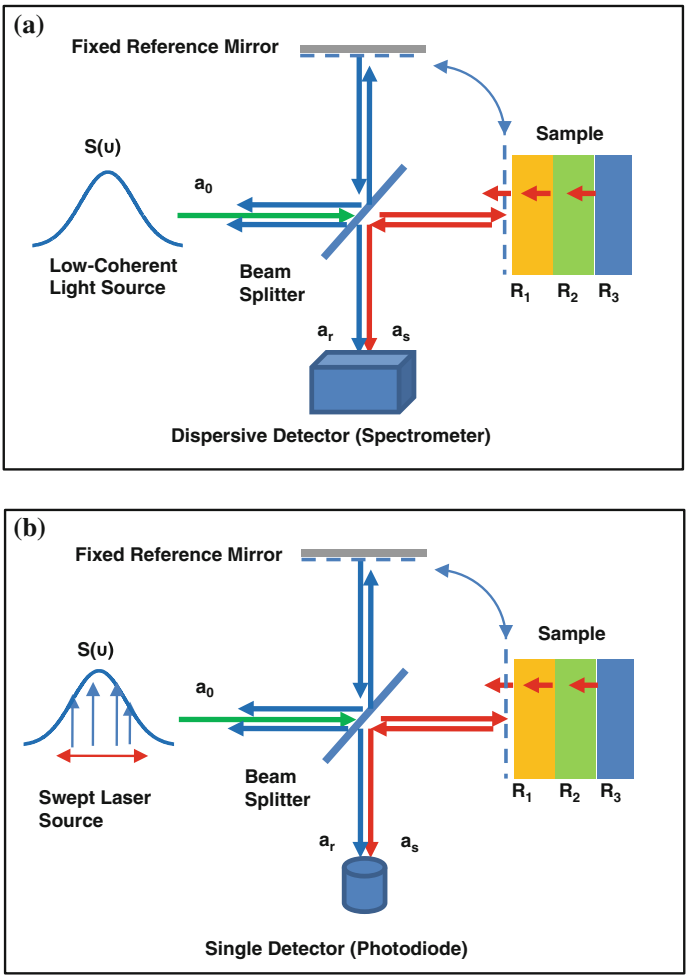

Fig. 6.2 Two types of FD-OCT: **a** Spectrometer based. **b** Swept source based

$$S_{combined}(v) = S(v) * \rho(v) * \left[R_r + \sum_i R_i + 2\sum_i \sqrt{R_r R_i}\cos(2\pi v \tau_i) + 2\sum_{i,j} \sqrt{R_i R_j}\cos(2\pi v(\tau_i - \tau_j)) \right] \tag{6.19}$$

where $\rho(v)$ is the spectral response of the spectrometer, which is usually non-linear with respect to v. To reconstruct A-scan signal $D(\tau)$, a Fourier transform on $S_{combined}(v)$ is performed, assuming a uniform $\rho(v)$ for simplicity, as,

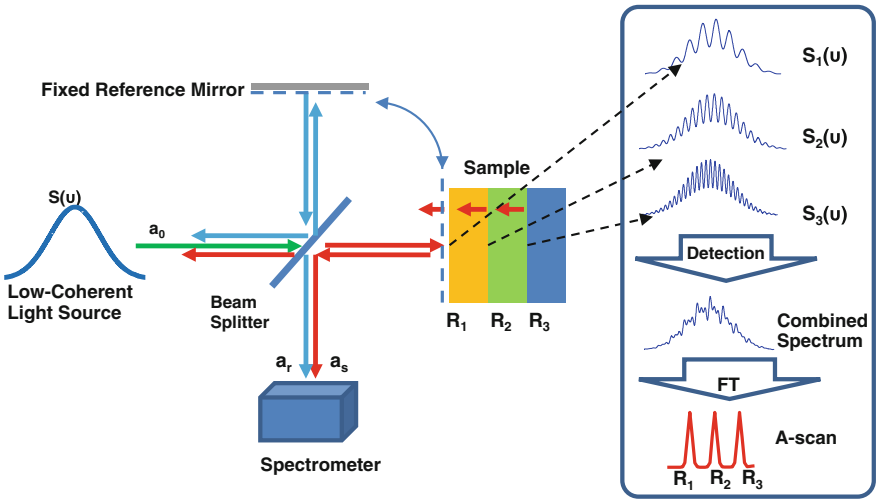

Fig. 6.3 Layout and signal processing of FD-OCT

$$
\begin{aligned}
D(\tau) &= FT_{v\to\tau}\{S_{combined}(v)\}\\
&= \left(R_r+\sum_i R_i\right)G(\tau)+\sum_i\sqrt{R_rR_i}(G(\tau)\otimes\delta(\tau\pm\tau_i))+\sum_{i,j}\sqrt{R_iR_j}\big(G(\tau)\otimes\delta(\tau\pm(\tau_i-\tau_j))\big)\\
&= \frac{1}{c}\left(R_r+\sum_i R_i\right)G(\Delta D)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(DC Component)}\\
&\quad+\frac{1}{c}\sum_i\sqrt{R_rR_i}(G(\Delta D)\otimes\delta(\Delta D\pm\Delta D_i))\ldots\ldots\ldots\ldots\text{(Cross-Correlation Component)}\\
&\quad+\frac{1}{c}\sum_{i,j}\sqrt{R_iR_j}\big(G(\Delta D)\otimes\delta(\Delta D\pm(\Delta D_i-\Delta D_j)\big)\ldots\text{(Auto-Correlation Component)}
\end{aligned}
\tag{6.20}
$$

Here the delta functions $\delta(\Delta D\pm\Delta D_i)$ and reflectivities R_i represent the multi-layered sample. The cross-correlation component is the desired signal for Michelson type FD-OCT.

Obvious from Eq. (6.20), a direct Fourier transform generates a pair of conjugate signals at each side of the zero-delay plane, which is the major artifact in FD-OCT imaging. Several phase-shift interferometery methods have been developed to tackle such artifacts [8–12]. Later in this chapter, a simple and effective method called the “Simultaneous B-M-mode scanning method” is adopted by applying a linear phase modulation $\varphi(\mathrm{t}|\mathrm{x})=\beta*(\mathrm{t}|\mathrm{x})$ to each M-scan/B-scan’s 2D interferogram frame $I(t|x,\lambda)$ [12], as,

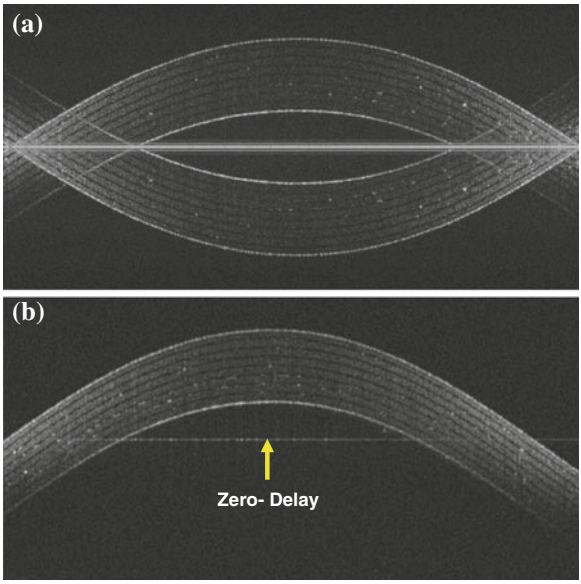

Fig. 6.4 Image comparison between **a** Standard (half-range) FD-OCT, **b** Full-range FD-OCT

$$
\begin{aligned}
F_{t|x\to f|u}[I(t|x,\lambda)] &= |a_r|^2\delta(f|u) \\
&+ \Gamma_{f|u}\{F_{t|x\to f|u}[a_s(t|x,\lambda)]\} \\
&+ F_{t|x\to f|u}[a_r^*(t|x,\lambda)a_r(\lambda)] \otimes \delta(f|u+\beta) \\
&+ F_{t|x\to f|u}[a_s(t|x,\lambda)a_r^*(\lambda)] \otimes \delta(f|u-\beta)\,,
\end{aligned}
\tag{6.21}
$$

where $\Gamma_{f|u}\{\}$ is the correlation operator. The first three terms on the right hand of Eq. (6.21) present the DC noise, autocorrelation noise, and complex-conjugate noise, respectively. The last term can be filtered out by a proper band-pass filter in the f|u domain and then convert back to the t|x domain by applying an inverse Fourier transform along the f|u direction which is a modified Hilbert transform. Figure 6.4 compares the images of a multi-layered phantom using standard (half-range) FD-OCT and full-range FD-OCT.

6.4 Imaging Resolution of OCT

From Eqs. (6.15) and (6.21), the A-scan signal describes the spatial convolution of PSF $G(\Delta D)$ and the sample's structure function $\delta(\Delta D \pm \Delta D_i)$, therefore the FWHM axial resolution of TD-OCT and FD-OCT are the same according to Eq. (6.18)

$$
\Delta d_{OCT} = \frac{2\ln 2}{\pi}\frac{\lambda_0}{\Delta\lambda} \tag{6.22}
$$

The lateral resolution of both TD-OCT and FD-OCT systems are determined by the numerical aperture (N. A.) of the imaging lens, as

$$\Delta x_{OCT} = \frac{2\sqrt{\ln 2}}{\pi}\frac{\lambda_0}{N.A.} \tag{6.23}$$

6.5 Optical Dispersion Mismatch and Compensation in OCT

Optical dispersion mismatch is a common problem for all Michelson type OCT systems, especially more so for an ultra-high resolution FD-OCT system using an extremely broadband light source. The optical dispersion arises from wavelength-dependent phase velocity described by the optical material-related propagation constant $\beta(\omega)$, which can be expanded as a Taylor series as,

$$\begin{aligned}\beta(\omega) =& \beta(\omega_0) + \left.\frac{d\beta}{d\omega}\right|_{\omega_0}(\omega-\omega_0) + \frac{1}{2}\left.\frac{d^2\beta}{d\omega^2}\right|_{\omega_0}(\omega-\omega_0)^2 \\ &+ \frac{1}{6}\left.\frac{d^3\beta}{d\omega^3}\right|_{\omega_0}(\omega-\omega_0)^3 + \ldots\end{aligned} \tag{6.24}$$

Here $\frac{d\beta}{d\omega}$ is the inverse group velocity, and $\frac{d^2\beta}{d\omega^2}$ describes the group velocity dispersion, which majorly contributes to the broadening of system PSF. The third-order dispersion $\frac{d^3\beta}{d\omega^3}$ produces asymmetric distortion of the PSF.

The dispersion mismatch in an OCT system usually results from unbalance between optical components in the sample arm and reference arm, and for some cases like retinal OCT, also comes from the imaging sample itself (the vitreous humor). Dispersion compensation is necessary for most ultra-high resolution FD-OCT systems and usually two approaches are utilized,

(1) Hardware Approach: The dispersion of the sample arm is physically matched by putting balancing optical components on the reference arm. One simplest way may be using identical optics, and an alternative way is using dispersion matching prism pairs. Therefore hardware approach usually need additional cost for optical components.
(2) Software Approach: A phase correction via Hilbert transformation is applied to the original spectrum, as [13],

$$\overline{\Phi} = -a_2(\omega-\omega_0)^2 - a_3(\omega-\omega_0)^3 \tag{6.25}$$

Here only the second and third order dispersions are compensated), where a_2 and a_3 were pre-optimized parameters depending on the system. Software approach is cost-effective but brings large computing burden due to the Hilbert transformation.

6.6 Frontiers in OCT: GPU Accelerated Real-Time Ultra-High Speed FD-OCT

6.6.1 Background and Overview

In recent years, the raw data acquisition speed of FD-OCT systems has been advancing rapidly so that 100 kHz level FD-OCT system are common today [14–23]. For a spectrometer-based OCT, an ultrahigh speed CMOS line scan camera based system achieved up to 312,500 line/s in the year 2008 [14]; while for a swept-source based OCT, Multi-MHz rate was achieved in the year 2010 [22]. Such ultrahigh acquisition speed enables time-resolved volumetric (4D) recording and reconstruction of dynamic physical processes [17–21].

While the acquisition speed of FD-OCT has been satisfying the interventional imaging requirements, the image (A-scan) reconstruction and visualization speeds have not kept up with the ever increasing rate of data acquisition, becoming one of the limiting factors for employing FD-OCT systems in practical therapeutic intervention applications. For most FD-OCT systems, the raw data is acquired in real-time but only saved to hard drive for later reconstruction and visualization. For surgeries, such imaging protocol provides valuable "pre-operative/post-operative" images, but is incapable of providing real-time, "inter-operative" imaging for surgical guidance and intervention.

The image reconstruction and visualization schematic for a general FD-OCT system is show as Fig. 6.5, for both standard and full-range FD-OCT. The A-scans are acquired and processed independently; therefore FD-OCT is inherently ideal for a massive parallel processing architecture. Several parallel processing methods have been implemented to improve A-scan data of FD-OCT images including FPGA [24, 25] and multi-core CPU parallel processing [26, 27]. Recent progresses in general-purpose GPU makes it possible to implement ultrahigh throughput OCT data processing and visualization on a variety of low-cost, many-core graphics cards [28–36]. As shown in Fig. 6.6, one or multiple GPUs can be directly integrated into the FD-OCT system without needing any optical modification. Compared to FPGA and multi-core processing methods, GPU acceleration is more cost-effective in terms of price/performance ratio and convenience of system integration.

For the standard half-range FD-OCT, GPU-based image reconstruction has been implemented on both linear-k and non-linear-k systems [28–32]. As will be shown in this section, >3,000,000 A-scan/second processing rate for 1024-pixel standard FD-OCT is achieved.

For the full-range FD-OCT, the image reconstruction requires signal processing workload more than 3 times that of the standard OCT, since each A-scan requires a Modified Hilbert Transform implemented by three fast Fourier transforms (FFT) in different dimensions of the frame, a band-pass filtering, and necessary matrix transpose [37]. Several publications have demonstrated the GPU accelerated processing of full-range FD-OCT [33–35]. As will be shown in this section,

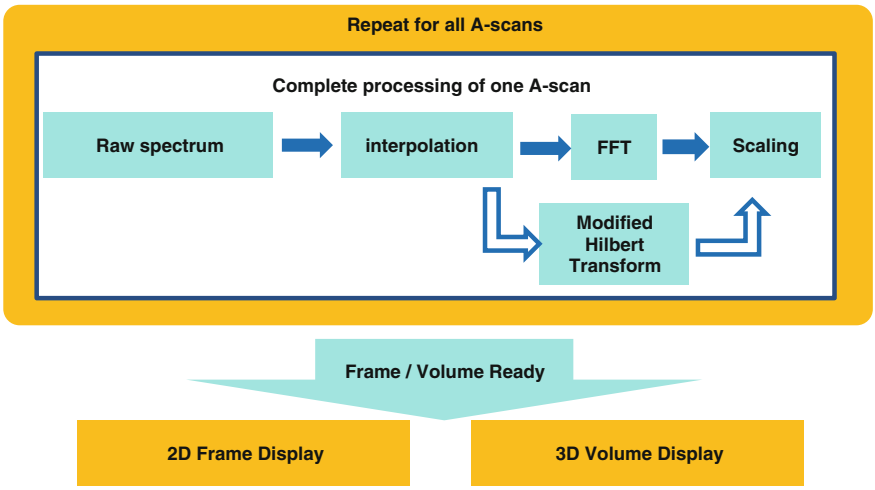

Fig. 6.5 Basic data processing and visualization flowchart for a genetic FD-OCT system. Hollow arrows indicate the data path for full-range FD-OCT processing

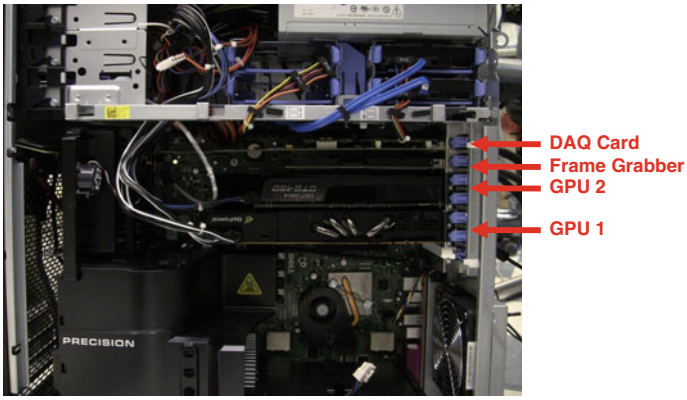

Fig. 6.6 GPU integration into the FD-OCT system

>1,300,000 A-scan/second processing rate for 1024-pixel full-range FD-OCT is achieved.

For volumetric visualization, multiple 2D slice extraction and co-registration is the simplest approach, while volume rendering offers more comprehensive spatial view of the whole 3D data set, which is not immediately available from 2D slices. However, volume rendering such as ray-casting is usually very time-consuming for CPU. So real-time rendering for a large data volume is only available through massive parallel processors such as GPU. Moreover, a complete 3D data set must be ready prior to any volumetric visualization due to the feature of FD-OCT signal processing method, which still would require a solution. As will be presented later

in this section, real-time 4D OCT imaging has also been achieved through GPU-based volume rendering [29, 36].

6.6.2 System Layout and Data Flow

The GPU acceleration technologies can be incorporated with a regular FD-OCT system through a hardware-software heterogeneous computing platform, as shown in Fig. 6.7. A 12-bit dual-line CMOS line-scan camera (Sprint spL2048-140 k, Basler AG, Germany) works as the detector of the OCT spectrometer. A super-luminescence diode (SLED) ($\lambda_0 = 825$ nm, $\Delta\lambda = 70$ nm, Superlum, Ireland) was used as the low-coherent light source, which provided a measured axial resolution of approximately 5.5 μm in air, 4.1 μm in water. Since the SLED's spectrum only covers less than half of the CMOS camera's sensor array, the camera is set to work at 1024-pixel mode by selecting the area-of-interest (AOI). The camera works at the "dual-line averaging mode" to get 3 dB higher SNR of the raw spectrum [14], and the minimum line period is camera-limited to 7.8 μs, which corresponds to a

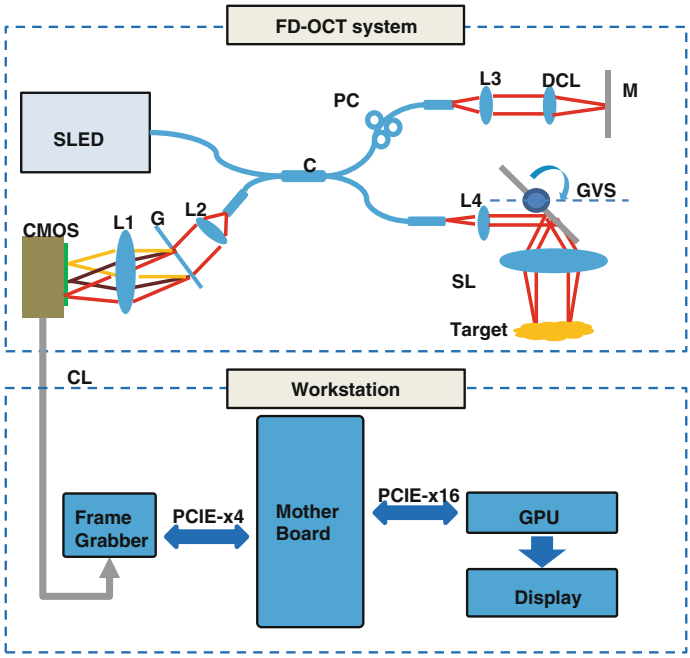

Fig. 6.7 System configuration: CMOS, CMOS line scan camera; L, spectrometer lens; G, grating; C1, C2, C3, achromatic collimators; C, 50:50 broadband fiber coupler; CL, camera link cable; GPU, graphics processing unit; GVS, galvanometer (only the first galvanometer is illustrated for simplicity); SL, scanning lens; DCL, dispersion compensation lens; M, reference mirror; PC, polarization controller; SP, Sample

maximum line rate of 128 k A-scan/s. The beam scanning was implemented by a pair of high speed galvanometer mirrors driven by a dual channel function generator and synchronized with a high speed frame grabber (PCIE-1429, National Instruments, USA). The transversal resolution was approximately 20 μm, assuming Gaussian beam profile. A quad-core Dell T7500 workstation was used to host the frame grabber (PCIE x4 interface) and GPU (PCIE x16 interface), and in our experiments, certain models of GPUs manufactured by NVIDIA was used to perform the FD-OCT image reconstruction and visualization. The GPU is programmed through NVIDIA's CUDA technology. The FFT operation is implemented by the CUFFT library [38]. The software is developed under Microsoft Visual C++ environment with the NI-IMAQ Win32 API (National Instrument).

For the full-range FD-OCT mode, a phase modulation is applied to each B-scan's 2D interferogram frame by slightly displacing the probe beam off the first galvanometer's pivoting point (here only the first galvanometer is illustrated in the figure) [37, 39]. In this experimental setup, the scanning angle is 8° and the beam offset is 1.4 mm, therefore a carrier frequency of 30 kHz is obtained according to Ref. [39].

The Data Flow on a CPU-GPU heterogeneous computing architecture is shown in Fig. 6.8, where three major threads are used for the data acquisition (Thread 1), the GPU processing (Thread 2), and the image display (Thread 3), respectively.

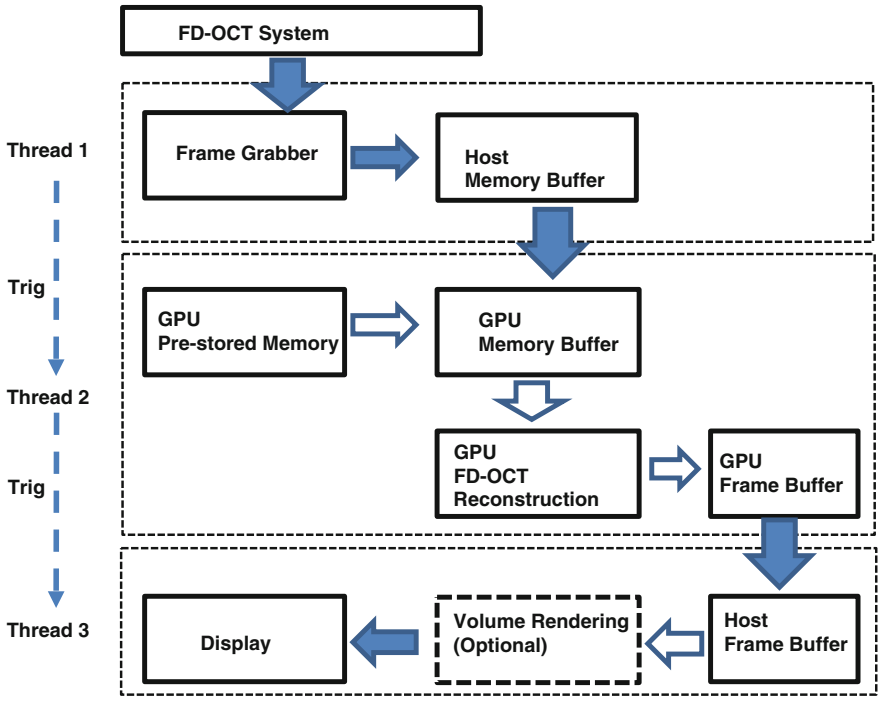

Fig. 6.8 Data flow on a CPU-GPU heterogeneous computing architecture

The three threads are triggered unidirectionally and work in the synchronized pipeline mode, as indicated by the blue dashed arrows.

Thread 1 is implemented through National Instruments IMAQ library, which controls the interface between frame grabber and CMOS camera.

Thread 2 is a GPU-CPU hybrid thread which consists of tens of thousands of GPU threads. The solid arrows describe the main data stream and the hollow arrows indicate the internal data flow of the GPU. Some parameters for the interpolation and the reference frame are stored in the "GPU Pre-stored Memory" section. In our previous work, we have implemented a series of algorithms into the "GPU FD-OCT Reconstruction" box [29, 33, 35, 36].

Thread 3 is responsible for image display. The GPU based volume rendering is also available if the OCT system is doing C-scan to get 3D data set. The volume rendering can be done on the same GPU for reconstruction [29], or on a secondary GPU dedicated to visualization [36], as will be shown later.

6.6.3 Benchmark Tests

(1) **GPU processing line rate for different FD-OCT methods**

First we performed benchmark line rate test of different FD-OCT processing methods on the GTX 480 GPU as follows:

LIFFT	Standard FD-OCT with Linear Spline Interpolation (LSI);
LIFFT-C	Full-range FD-OCT with LSI;
CIFFT	Standard FD-OCT with Cubic Spline Interpolation (CSI);
CIFFT-C	Full-range FD-OCT with CSI;
NUDFT	Standard FD-OCT with Non-Uniform Discrete Fourier Transform (NUDFT);
NUDFT-C	Full-range FD-OCT with NUDFT;
NUFFT	Standard FD-OCT with Non-Uniform Fast Fourier Transform (NUFFT);
NUFFT-C	Full-range FD-OCT with NUFFT;

The details of the different algorithms are presented in Ref. [35]. All algorithms are tested with 4,096 lines of both 1024-pixel spectrum and 2048-pixel spectrum. For each case, both the peak internal processing line rate and the reduced line rate considering the data transfer bandwidth of PCIE x16 interface are listed in Fig. 6.9. The processing time is measured with CUDA API functions and all CUDA threads are synchronized before the measurement. In the system, the PCIE x16 interface has a data transfer bandwidth of about 4.6 GByte/s for both transfer data into and out of the GPU from the Host side.

As in Fig. 6.9a, the processing speed of LIFFT goes up to >3,000,000 line/s (effectively >1,000,000 line/s under data transfer limit), achieving the fastest

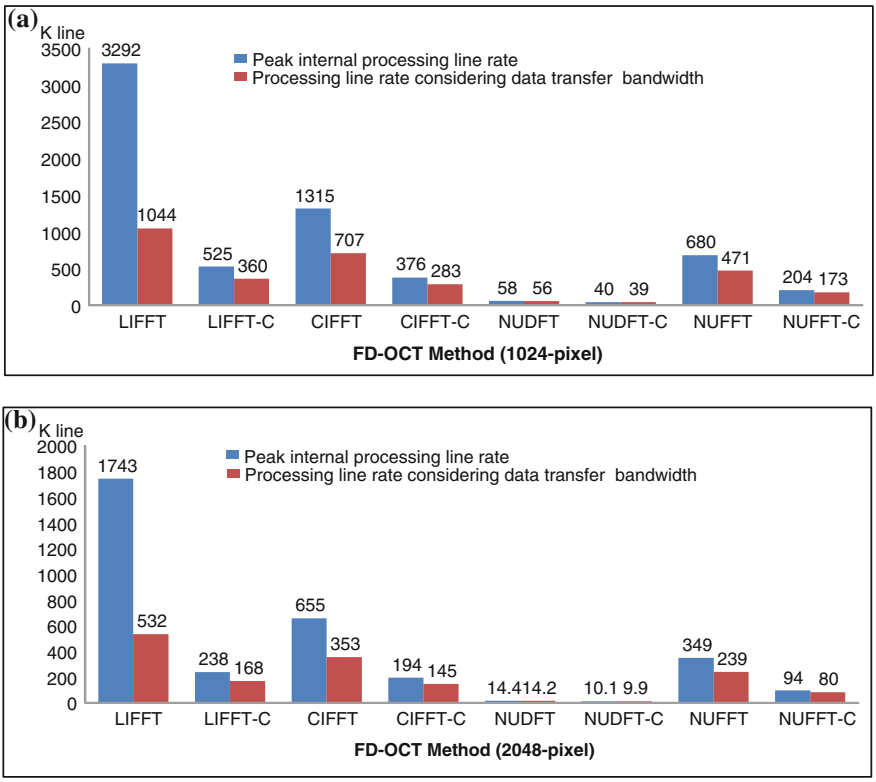

Fig. 6.9 Benchmark line rate test of different FD-OCT processing method. **a** 1024-pixel FD-OCT; **b** 2048-pxiel FD-OCT

processing line rate of FD-OCT by the time of publication [35]. The final processing line rate for 1024-pixel full-range FD-OCT with GPU-NUFFT is 173 k line/s, which is still higher than the maximum camera acquisition rate of 128 k line/s, while the GPU-NUDFT speed is relatively lower in both standard and complex FD-OCT.

(2) **Comparison of point spread function**

Then we compared the depth resolved point spread function (PSF) of standard FD-OCT by different processing methods, as shown in Fig. 6.10, using GPU based LIFFT, CIFFT, NUDFT and NUFFT, respectively. A mirror is used as an image sample for evaluating the PSF. Here 1,024 of GPU-processed 1024-pixel A-scans are averaged to present the mean noise level for the PSF in each axial position [40]. From Fig. 6.10a, it is evident that using LIFFT method introduces a significant background noise and side-lobes around the signal peak. The side-lobes tend to be broad and extend to their neighboring peaks, which can results in significant OCT image degradation. When CIFFT is used instead, as shown in Fig. 6.10b, the side-lobes are suppressed in the shallow image depth, but still considerably high in

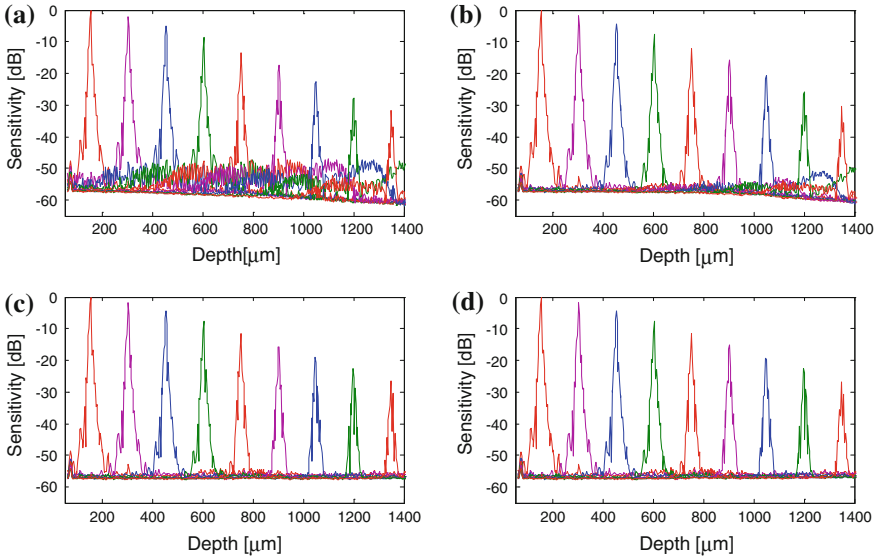

Fig. 6.10 Depth resolved point spread function for standard FD-OCT by different processing methods: **a** LIFFT; **b** CIFFT; **c** NUDFT; **d** NUFFT

deeper depth. Figure 6.10c and d shows significant suppression of side-lobes over the whole image depth which utilized NUDFT and NUFFT, respectively.

The PSFs and sensitivity roll-off of the full-range FD-OCT mode are shown in Fig. 6.11, where a different scanning lens was used and the spot size was approximately 40 μm [36]. The data was processed by a NVIDIA GTX 580 GPU (512 cores, 1.59 GHz, processor clock and 1.5 GBytes graphics memory). At both sides of the zero-delay, PSFs at different positions are processed as A-scans using LSI with FFT (Fig. 6.11a and b) and NUFFT (Fig. 6.11c and d), respectively.

As one can see, using NUFFT processing, the system obtained a sensitivity fall-off of 19.6 dB from position near zero-delay to the negative edge, and 18.8 dB to the positive edge, which is lower than using linear interpolation with FFT (24.7 dB to the negative edge and 24.5 dB to the positive edge). Moreover, compared to the linear interpolation method, NUFFT obtained a constant background noise level over the whole A-scan range. The blue arrows in Fig. 6.11a and b indicates side lobes in the PSFs near both positive and negative edges as a result of interpolation error. By applying a proper super-Gaussian filter to the modified Hilbert transform [34], the conjugate suppression ratios of 37.0 and 40.9 dB are obtained respectively at the positive and negative sides near zero-delay.

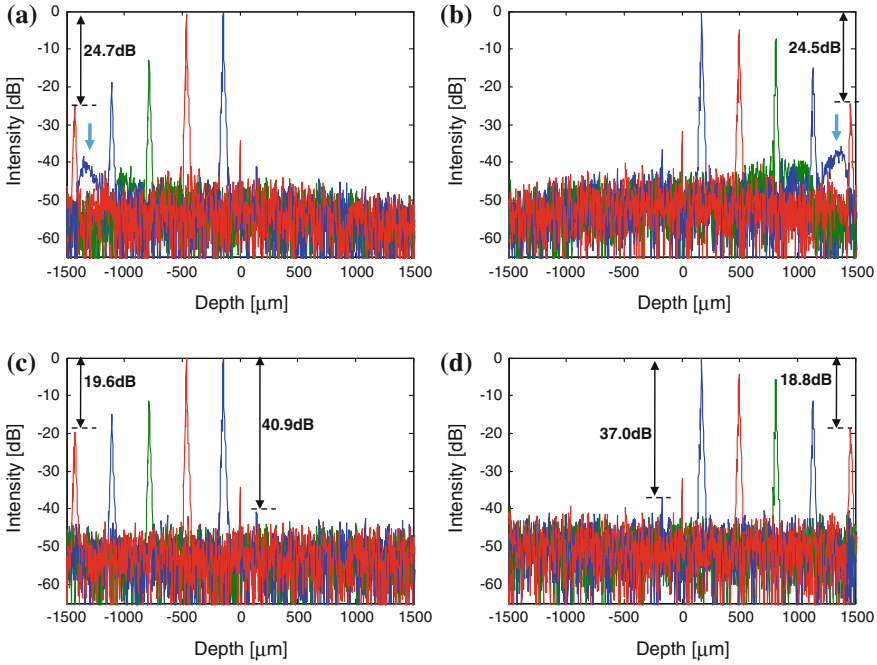

Fig. 6.11 Point spread function and sensitivity roll-off for full-range FD-OCT: (**a**) and (**b**), PSFs processed by linear interpolation with FFT, *blue arrows* indicate the side lobes of PSFs near positive and negative edges due to interpolation error. (**c**) and (**d**), PSFs processed by NUFFT

6.6.4 Real-Time 2D Imaging Tests

In vivo human finger imaging using GPU-NUFFT-based full-range FD-OCT is shown in Fig. 6.12. Images were displayed at 29.8fps with original frame size of 4,096 pixel (lateral) × 1,024 pixel (axial). Figure 6.12a and b present the coronal scans of the fingertip and palm, where the epithelial structures such as sweat duct (SD), stratum corneum (SC) and stratum spinosum (SS) are clearly distinguishable. Figure 6.12c and d present the coronal scans of the finger nail fold region, showing the major dermatologic structures such as epidermis (E), dermis (D), nail bed (NB), and nail root (NR), as well as in the sagittal scans in Fig. 6.12e and f.

Compared to standard FD-OCT, the GPU-NUFFT full-range FD-OCT image is free of conjugate artifact, DC noise, and autocorrelation noise. These noises are problematical to remove in standard FD-OCT. Moreover, due to the implementation of the complex OCT processing, the image depth is effectively doubled, with the highest SNR region in the zero delay point.

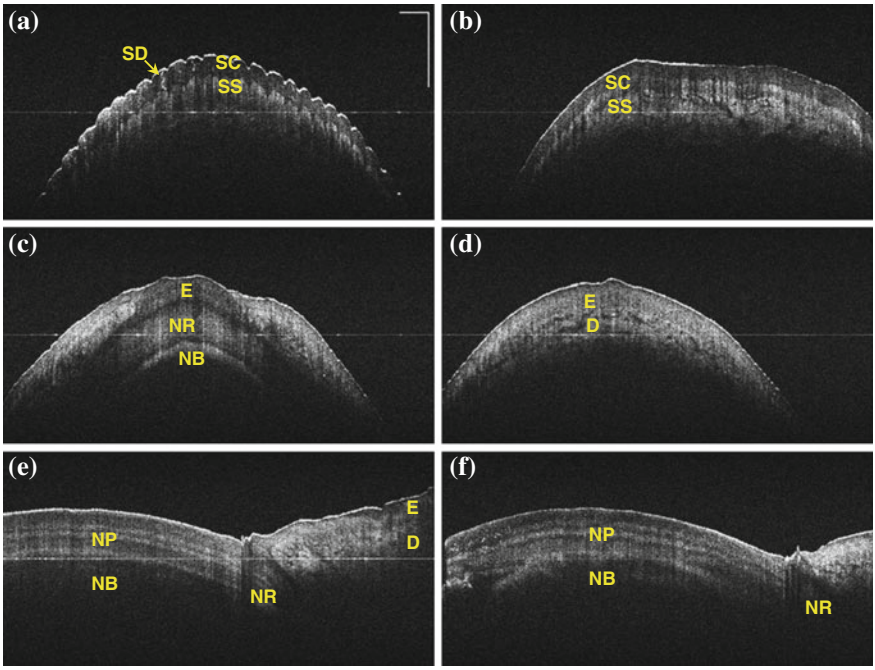

Fig. 6.12 Real-time full-range FD-OCT images using GPU-NUFFT, where the bars represent 1 mm in both dimensions for all images: **a** Finger tip, (coronal). **b** Finger palm (coronal). **c–d** Finger nail fold (coronal); **e–f** Finger nail (sagittal). SD, sweat duct; SC, stratum corneum; SS, stratum spinosum; NP, nail plate; NB, nail bed; NR, nail root; E, epidermis; D, dermis

6.6.5 Real-Time 3D (4D) Imaging Tests

(1) **4D Visualization by *en face* slice extraction**

First, the simplest 4D visualization was tested by en face slice extraction, implemented by a single GPU (NVIDIA Quadro FX5800, with 240 stream processors, 1.3 GHz clock rate and 4 GBytes graphics memory). The FD-OCT worked at 1024-pixel mode. The line scan rate was set to 100,000 line/second. A Naval orange juice sac was used as the sample. Three different volume sizes are tested: $250 \times 160 \times 512$ voxels (40,000 A-scans/volume); $250 \times 80 \times 512$ voxels (20,000 A-scans/volume); $125 \times 80 \times 512$ voxels (10,000 A-scans/volume); corresponding to a volume rate of 2.5, 5 and 10 volume/s, respectively. Figure 6.13 shows the *en face* slices of approximately 1×1 mm region in two different depths extracted from the same volumetric data and the depth difference of about 25 μm. All the A-scans of one volume were acquired and processed as one batch and remapped for en face slice extraction. More than one *en face* images at different depth can be quickly reconstructed and displayed simultaneously since the complete 3D data is available. As one can see, with decreasing volume size and

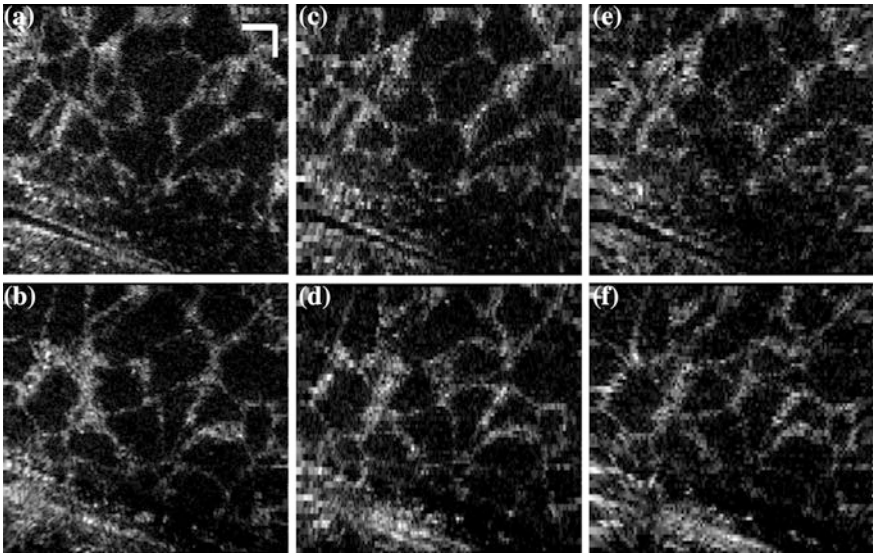

Fig. 6.13 4D visualization by *en face* slice extraction, the scale bar represents 100 μm for all images: **a** 250 × 160 × 512 voxels; **b** from the same volume as (**a**) but 25 μm deeper; **c** 250 × 80 × 512 voxels; **d** from the same volume as (**c**) but 25 μm deeper; **e** 125 × 80 × 512 voxels; **f** from the same volume as (**e**) but 25 μm deeper

increasing volume rate, the image quality degenerate but the major details such as cell wall are still clear enough to be visible compared with the largest volume size slices as in Fig. 6.13a and b.

(2) **4D Visualization by volume rendering (Standard FD-OCT, Single GPU)**

This technology implemented using a NVIDIA Quadro FX 5800 GPU can perform real-time volume rendering of continuous acquired data volume at 10 volume per second 4D FD-OCT "live" image using. The system's acquisition line rate is set to be 125,000 line/s at 1024-OCT mode. The acquisition volume size is set to be 12,500 A-scans, providing 125(X) × 100(Y) × 512(Z) voxels after the signal processing stage, which takes less than 10 ms and leaves more than 90 ms for each volume interval at the volume rate of 10 volume/s. The image plane is set to 512 × 512 pixels, which means a total number of 5,122 = 262,144 eye rays are used to accumulate though the whole rendering volume for the ray-casting process [41]. The actual rendering time is recorded during the imaging processing to be ∼3 ms for half volume and ∼6 ms for full volume, which is much shorter than the volume interval residual (>90 ms).

First we tested the real-time visualization ability by imaging non-biological samples. Here the volume rendering is applied and the real volume size is approximately 4 × 4 × 0.66 mm. Figure 6.14a presents the top surface of a piece of sugar-shell coated chocolate. Here the perspective projection is used for the

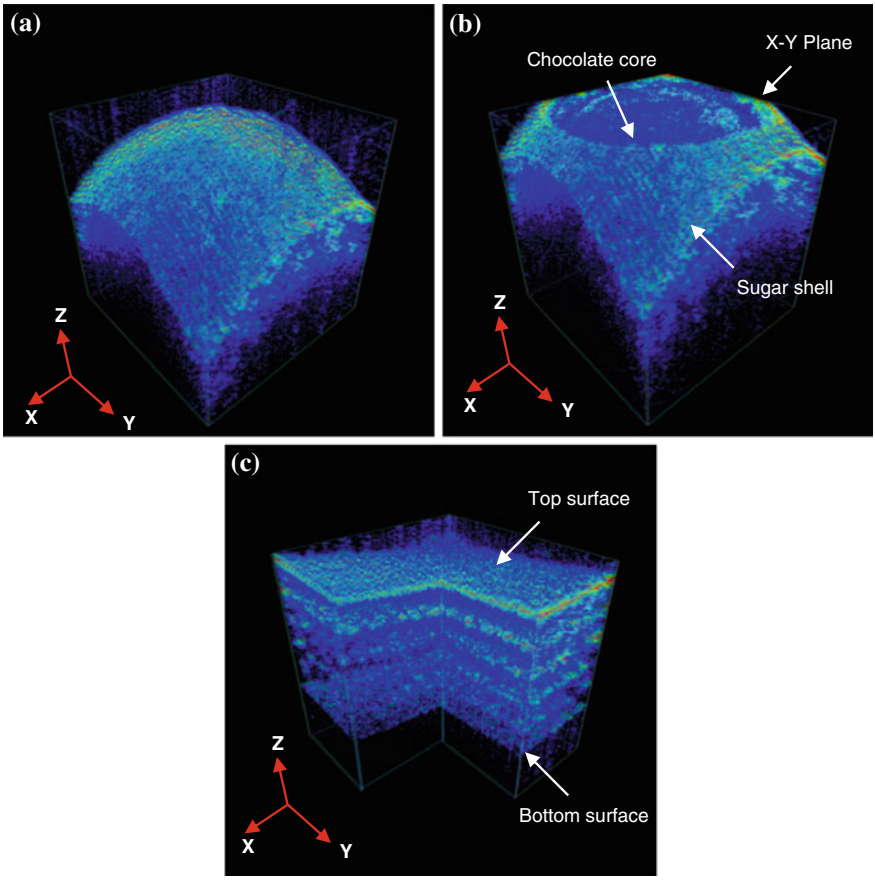

Fig. 6.14 **a** The 3D OCT image of a piece of sugar-shell coated chocolate; **b** sugar-shell top truncated by the X–Y plane, inner structure visible; **c** A five-layer phantom

eye's viewpoint [42], and the rendering volume frame is indicated by the white lines. Figure 6.14b shows the situation when the target surface is truncated by the rendering volume's boundary, the X–Y plane, where the sugar shell is virtually "peeled" and the inner structures of the chocolate core is clearly recognizable. Figure 6.14c illustrates a five-layer polymer phantom, where the layers are easily distinguishable. The volume rendering frame in Fig. 6.14c is configured as "L" shape so the tapes are virtually "cut" to reveal the inside layer structures.

In vivo real-time 3D imaging of a human fingertip is shown in Fig. 6.15. Figure 6.15a shows the skin and fingernail connection, the full volume rendering is applied here giving a real size of $4 \times 4 \times 1.32$ mm considering the large topology range of the nail connection region. The major dermatologic structures such as epidermis (E), dermis (D), nail fold (NF), nail root (NF) and nail body (N) are clearly distinguishable from Fig. 6.15a. The fingerprint is imaged and shown in

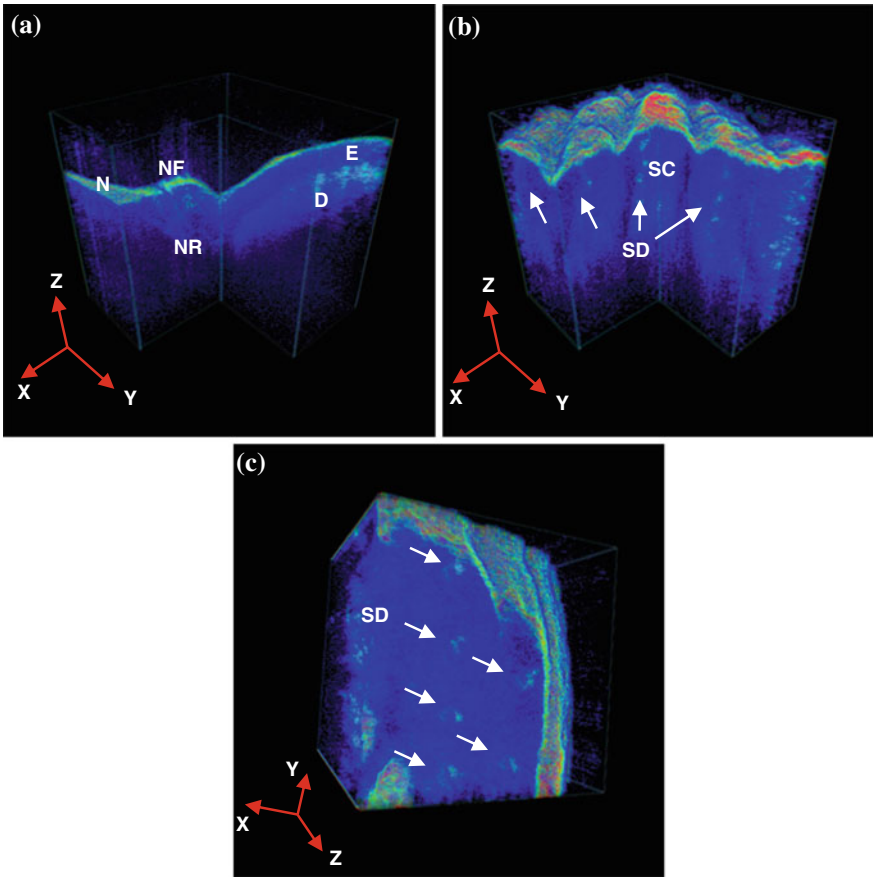

Fig. 6.15 *In vivo* real-time 3D imaging of a human fingertip: **a** Skin and fingernail connection; **b** Fingerprint, *side-view* with "L" volume rendering frame; **c** Fingerprint, *top-view*

Fig. 6.15b, where the epithelium structures such as sweat duct (SD), stratum corneum (SC) can be clearly identified. Figure 6.30c offers a top-view of the fingerprint region, where the surface is virtually peeled by the X-Y plane and the inner sweat duct are clearly visible. The volume size for Fig. 6.15b and c is $2 \times 2 \times 0.66$ mm.

To make full use of the computing power and the whole 3D data, multiple 2D frames processed from the same 3D data set with different model view matrix can be rendered. This is shown as a side-view (Fig. 6.16a, b, d, e), top-view (Fig. 6.16c) and bottom-view (Fig. 6.16f), where Fig. 6.16a and d are using the same model view matrix but displayed with the "L" volume rendering frame. All frames are rendered from the same volume period and displayed simultaneously, thus gives more comprehensive information of the target. The two vertexes with the big red and green dot indicate the same edge for each rendering volume frame.

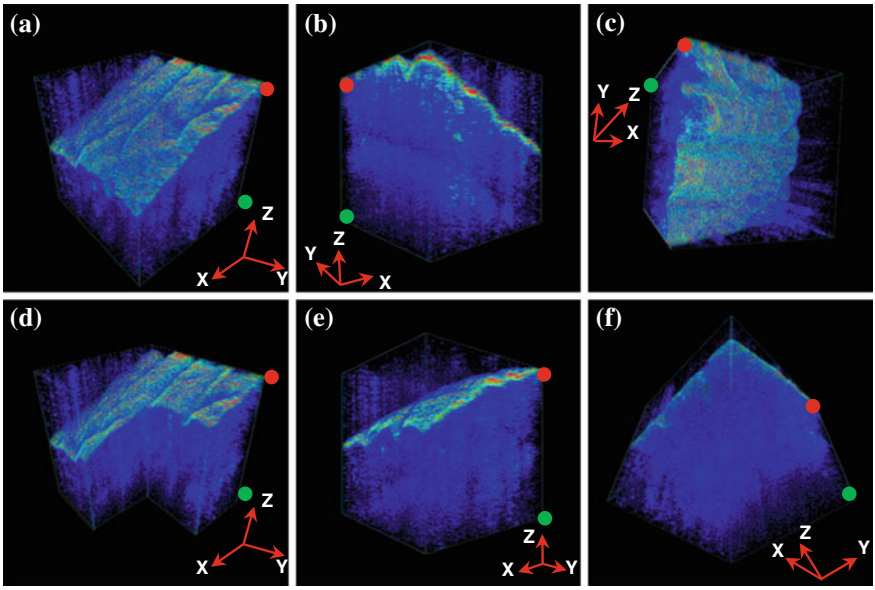

Fig. 6.16 Real-time volume imaging of the finger print by 4D standard FD-OCT

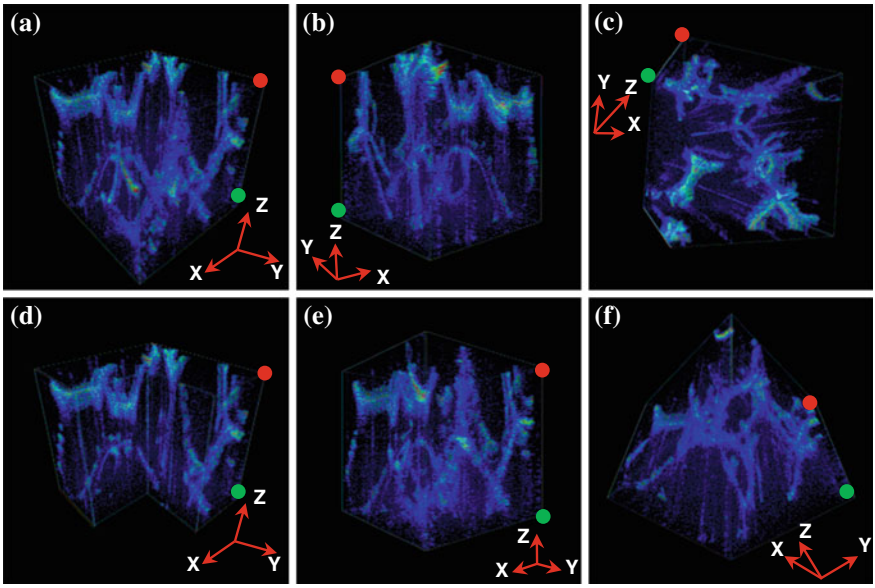

Fig. 6.17 Real-time volume imaging of the polymer foam by 4D standard FD-OCT

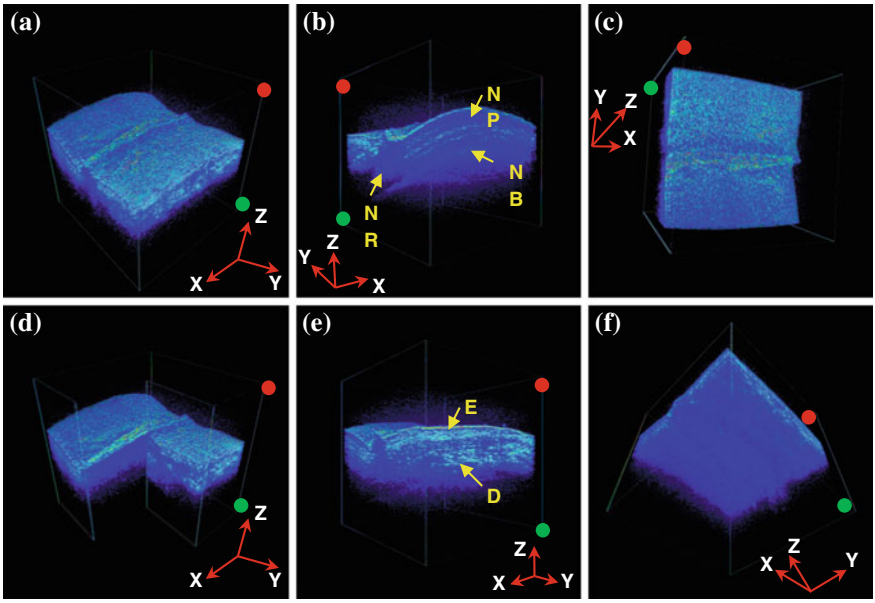

Fig. 6.18 Real-time volume imaging of the nail fold by 4D full-range FD-OCT

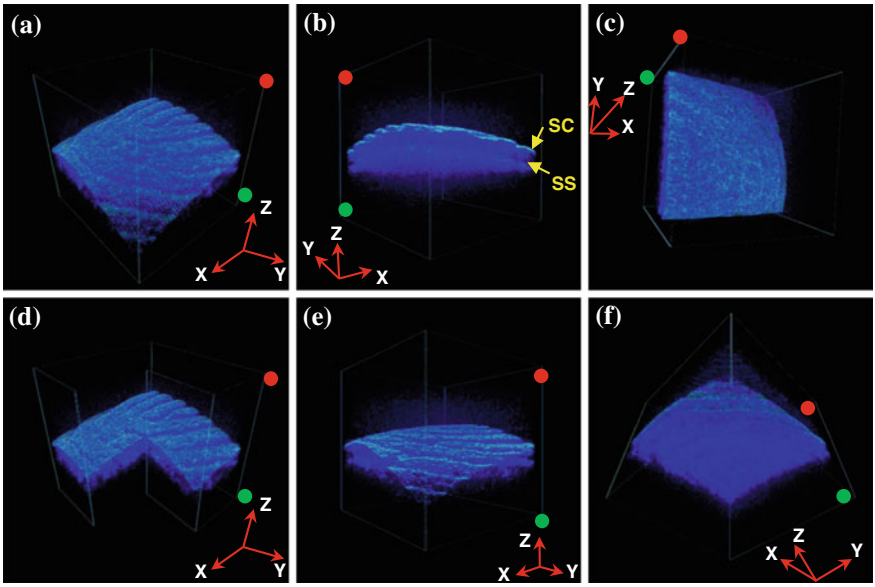

Fig. 6.19 Real-time volume imaging of the finger print by 4D full-range FD-OCT

Figure 6.17 shows the imaging of a polymer foam sample following the same rendering protocol. The volume size for both Figs. 6.16 and 6.17 is 2 × 2 × 0.66 mm.

(3) **4D Visualization by volume rendering (Full-range FD-OCT, Dual-GPU)**

Real-time 4D full-range FD-OCT using the dual-GPU platform can improve the system performance and stability [43]. The system consists of GPU-1 (NVIDIA GeForce GTX 480) with 480 stream processors, 1.45 GHz processor clock and 1.5 GBytes graphics memory is dedicated for raw data processing of B-scan frames. GPU-2 (NVIDIA GeForce GTS 450) with 192 stream processors, 1.76 GHz processor clock and 1.0 GBytes graphics memory is dedicated for the volume rendering and display of the complete C-scan image processed by GPU-1. The A-scans were processed by CIFFT-C method.

In this demonstration, the B-scan size is set to 256 A-scans with 1024 pixel each. The C-scan size is set to 100 B-scans, resulting in 256(Y) × 100(X) × 1024(Z) voxels (effectively 250(Y) × 98(X) × 1024(Z) voxels after removing of edge pixels due to fly-back time of galvanometers), and 5 volumes/second. It takes GPU-2 about 8 ms to render one 2D image with 512 × 512 pixel from this 3D data set using the ray-casting algorithm.

Figures 6.18 and 6.19 presents the multiple real-time volume imaging of the nail fold and finger print by 4D full-range FD-OCT, respectively. The scanning range is 3.5 mm (X) × 3.5 mm (Y) lateral and 3 mm (Z) for the axial full-range. The major dermatologic structures such as epidermis (E), dermis (D), nail plate (NP), nail root (NR) and nail bed (NB), stratum corneum (SC) and stratum spinosum (SS) are clearly distinguishable.

References

1. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T et al (1991) Optical coherence tomography. Science 254(5035):1178–1181
2. Saleh BEA, Teich MC (2007) Fundamentals of photonics, Chapter 11, 2nd Edn, Wiley-Interscience (March 9)
3. Schmitt JM (1999) Optical coherence tomography (OCT): a review. IEEE J Sel Top Quantum Electron 5:1205
4. Fercher A (1995) Measurement of intraocular distances by backscattering spectral interferometry. Optics Commun 117:43
5. Drexler W, Fujimoto JG (2008) Optical coherence tomography: technology and applications. Springer, New York
6. Leitgeb R, Hitzenberger C, Fercher A (2003) Performance of fourier domain vs. time domain optical coherence tomography. Opt Express 11(8):889–894
7. Wojtkowski M (2010) High-speed optical coherence tomography: basics and applications. Appl Opt 49:D30–D61
8. Leitgeb RA, Hitzenberger CK, Fercher AF, Bajraszewski T (2003) Phase-shifting algorithm to achieve high-speed long-depth-range probing by frequency-domain optical coherence tomography. Opt Lett 28:2201–2203

9. Choma MA, Yang CH, Izatt JA (2003) Instantaneous quadrature low-coherence interferometry with 3 × 3 fiber-optic couplers. Opt Lett 28:2162
10. Bhushan B, Wyant JC, Koliopoulos CL (1985) Measurement of surface topography of magnetic tapes by Mirau interferometry. Appl Opt 24:1489–1497
11. Dubois A (2001) Phase-map measurements by interferometry with sinusoidal phase modulation and four integrating buckets. J Opt Soc Am A 18:1972–1979
12. Yasuno Y, Makita S, Endo T, Aoki G, Itoh M, Yatagai T (2006) Simultaneous B-M-mode scanning method for real-time full-range Fourier domain optical coherence tomography. Appl Opt 45:1861–1865
13. Wojtkowski M, Srinivasan V, Ko T, Fujimoto J, Kowalczyk A, Duker J (2004) Ultrahigh-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation. Opt Express 12:2404–2422
14. Potsaid B, Gorczynska I, Srinivasan VJ, Chen Y, Jiang J, Cable A, Fujimoto JG (2008) Ultrahigh speed Spectral/Fourier domain OCT ophthalmic imaging at 70,000 to 312,500 axial scans per second. Opt Express 16:15149–15169
15. Huber R, Adler DC, Fujimoto JG (2006) Buffered Fourier domain mode locking: unidirectional swept laser sources for optical coherence tomography imaging at 370,000 lines/s. Opt Lett 31:2975–2977
16. Oh W-Y, Vakoc BJ, Shishkov M, Tearney GJ, Bouma BE (2010) >400 kHz repetition rate wavelength-swept laser and application to high-speed optical frequency domain imaging. Opt Lett 35:2919–2921
17. Grulkowski I, Gora M, Szkulmowski M, Gorczynska I, Szlag D, Marcos S, Kowalczyk A, Wojtkowski M (2009) Anterior segment imaging with Spectral OCT system using a high-speed CMOS camera. Opt Express 17:4842–4858
18. Gora M, Karnowski K, Szkulmowski M, Kaluzny BJ, Huber R, Kowalczyk A, Wojtkowski M (2009) Ultra high-speed swept source OCT imaging of the anterior segment of human eye at 200 kHz with adjustable imaging range. Opt Express 17:14880–14894
19. Gargesha M, Jenkins MW, Rollins AM, Wilson DL (2008) Denoising and 4D visualization of OCT images. Opt Express 16:12313–12333
20. Gargesha M, Jenkins MW, Wilson DL, Rollins AM (2009) High temporal resolution OCT using image-based retrospective gating. Opt Express 17:10786–10799
21. Jenkins MW, Rothenberg F, Roy D, Nikolski VP, Hu Z, Watanabe M, Wilson DL, Efimov IR, Rollins AM (2006) 4D embryonic cardiography using gated optical coherence tomography. Opt Express 14:736–748
22. Wieser W, Biedermann BR, Klein T, Eigenwillig CM, Huber R (2010) Multi-megahertz OCT: high quality 3D imaging at 20 million A-scans and 4.5 GVoxels per second. Opt Express 18:14685–14704
23. Klein T, Wieser W, Eigenwillig CM, Biedermann BR, Huber R (2011) Megahertz OCT for ultrawide-field retinal imaging with a 1050 nm Fourier domain mode-locked laser. Opt Express 19:3044–3062
24. Ustun TE, Iftimia NV, Ferguson RD, Hammer DX (2008) Real-time processing for Fourier domain optical coherence tomography using a field programmable gate array. Rev Sci Instrum 79:114301
25. Desjardins AE, Vakoc BJ, Suter MJ, Yun SH, Tearney GJ, Bouma BE (2009) Real-time FPGA processing for high-speed optical frequency domain imaging. IEEE Trans Med Imaging 28:1468–1472
26. Liu G, Zhang J, Yu L, Xie T, Chen Z (2009) Real-time polarization-sensitive optical coherence tomography data processing with parallel computing. Appl Opt 48:6365–6370
27. Probst J, Hillmann D, Lankenau E, Winter C, Oelckers S, Koch P, Hüttmann G (2010) Optical coherence tomography with online visualization of more than seven rendered volumes per second. J Biomed Opt 15:026014
28. Watanabe Y, Itagaki T (2009) Real-time display on Fourier domain optical coherence tomography system using a graphics processing unit. J Biomed Opt 14:060506

29. Zhang K, Kang JU (2010) Real-time 4D signal processing and visualization using graphics processing unit on a regular nonlinear-k Fourier-domain OCT system. Opt Express 18:11772–11784
30. Jeught SV, Bradu A, Podoleanu AG (2010) Real-time resampling in Fourier domain optical coherence tomography using a graphics processing unit. J Biomed Opt 15:030511
31. Sylwestrzak M, Szkulmowski M, Szlag D, Targowski P (2010) Real-time imaging for Spectral optical coherence tomography with massively parallel data processing. Photonics Lett Poland 2:137–139
32. Rasakanthan J, Sugden K, Tomlins PH (2011) Processing and rendering of Fourier domain optical coherence tomography images at a line rate over 524 kHz using a graphics processing unit. J Biomed Opt 16:020505
33. Kang JU, Zhang K (2010) Real-time complex optical coherence tomography using graphics processing unit for surgical intervention. Proceedings of the 23rd Annual Meeting of the IEEE Photonics Society, p 38-39
34. Watanabe Y, Maeno S, Aoshima K, Hasegawa H, Koseki H (2010) Real-time processing for full-range Fourier-domain optical-coherence tomography with zero-filling interpolation using multiple graphic processing units. Appl Opt 49:4756–4762
35. Zhang K, Kang JU (2010) Graphics processing unit accelerated non-uniform fast Fourier transform for ultrahigh-speed, real-time Fourier-domain OCT. Opt Express 18:23472–23487
36. Zhang K, Kang JU (2011) Real-time intraoperative 4D full-range FD-OCT based on the dual graphics processing units architecture for microsurgery guidance. Biomed Opt Express 2:764–770
37. Yasuno Y, Makita S, Endo T, Aoki G, Itoh M, Yatagai T (2006) Simultaneous B-M-mode scanning method for real-time full-range Fourier domain optical coherence tomography. Appl Opt 45:1861–1865
38. NVIDIA (2012) NVIDIA CUDA CUFFT library version 5.0
39. Baumann B, Pircher M, Götzinger E, Hitzenberger CK (2007) Full range complex spectral domain optical coherence tomography without additional phase shifters. Opt Express 15:13375–13387
40. Vergnole S, Lévesque D, Lamouche G (2010) Experimental validation of an optimized signal processing method to handle non-linearity in swept-source optical coherence tomography. Opt Express 18:10446–10461
41. Levoy M (1988) Display of surfaces from volume data. IEEE Comp Graph Appl 8:29–37
42. Kaufman A, Mueller K (2005) Overview of volume rendering. In: Johnson C, Hansen C (eds) The visualization handbook. Academic, New York
43. Zhang K, Kang JU (2011) Real-time dual-mode standard/complex Fourier-domain OCT system using graphics processing unit accelerated 4D signal processing and visualization. Proc SPIE 7904:79040J–79041J