

Analysis of Synthetic Diamond Wafer  
Interferograms Using a Parallelized Simulated  
Annealing Algorithm

Matthew Demas  
University of Connecticut  
Storrs, CT 06269  
demas.ma@gmail.com

May 19, 2008

## **Abstract**

Diamonds are known for both their beauty and their durability. Jefferson National Lab in Newport News, VA has found a way to utilize the diamond's strength to view the beauty of the inside of the atomic nucleus with the hopes of finding exotic forms of matter. By firing very fast electrons at a diamond sheet no thicker than a human hair, high energy particles of light known as photons are produced with a high degree of polarization that can illuminate the constituents of the nucleus known as quarks. The University of Connecticut Nuclear Physics group has responsibility for crafting these extremely thin, high quality diamond wafers. These wafers must be cut from larger stones that are about the size of a human finger, and then carefully machined down to the final thickness. The thinning of these diamonds is extremely challenging, as the diamond's greatest strength also becomes its greatest weakness. The Connecticut Nuclear Physics group has developed a novel technique to assist industrial partners in assessing the quality of the final machining steps, using a technique based on laser interferometry. The images of the diamond surface produced by the interferometer encode the thickness and shape of the diamond surface in a complex way that requires detailed analysis to extract. We have developed a novel software application to analyze these images based on the method of simulated annealing. Being able to image the surface of these diamonds without requiring costly X-ray diffraction measurements allows rapid feedback to the industrial partners as they refine their thinning techniques. Thus, by utilizing a material found to be beautiful by many, the beauty of nature can be brought more clearly into view.

# Contents

<b>1</b>	<b>Background</b>	<b>7</b>
1.1	Jefferson Lab and GlueX . . . . .	7
1.1.1	UConn and JLab . . . . .	7
1.1.2	The Beam Line . . . . .	7
1.1.3	Synthetic Diamond Wafers . . . . .	9
1.1.4	Determining Diamond Deformities . . . . .	9
1.2	Electromagnetic Theory . . . . .	10
1.2.1	Electromagnetic Radiation . . . . .	10
1.2.2	Superposition and Interference . . . . .	11
1.3	Interferometry and the Michelson Interferometer . . . . .	11
1.3.1	Interferometry . . . . .	11
1.3.2	The Michelson Interferometer . . . . .	13
1.4	Surface Resolution Approximation . . . . .	17
1.4.1	Huygen's Principle . . . . .	18
<b>2</b>	<b>Methods</b>	<b>21</b>
2.1	The Michelson Interferometer . . . . .	21
2.2	Two-Wave Interference . . . . .	23

2.3	Three-Wave Interference . . . . .	24
<b>3</b>	<b>Fringe Pattern Analysis</b>	<b>25</b>
3.1	Introduction . . . . .	25
3.2	Fringe Pattern Analysis Methods . . . . .	26
3.2.1	phase-shifting Technique . . . . .	26
3.2.2	Fourier-transform Method . . . . .	27
3.2.3	Regularization Algorithms . . . . .	28
3.2.4	Artificial Neural Network Method . . . . .	29
<b>4</b>	<b>Simulated Annealing</b>	<b>31</b>
4.1	Optimization . . . . .	31
4.2	History . . . . .	32
4.3	Method . . . . .	32
4.4	The Metropolis Algorithm . . . . .	33
4.5	Parallel Simulated Annealing . . . . .	34
4.5.1	Aarts Strategy . . . . .	35
4.5.2	MIR . . . . .	37
<b>5</b>	<b>Testing</b>	<b>40</b>
5.1	Solutions . . . . .	40
5.2	Test Problem . . . . .	41
5.3	Running ParSA . . . . .	43
<b>6</b>	<b>Results and Conclusions</b>	<b>45</b>
6.1	Visual Comparison . . . . .	45
6.1.1	Interferograms . . . . .	45

6.1.2	Surfaces . . . . .	45
6.2	Run Length Optimization . . . . .	47
6.2.1	MIR performance . . . . .	48
6.2.2	Sequential Performance . . . . .	49
6.2.3	Optimizing Run Length . . . . .	50
6.3	Conclusions . . . . .	51
<b>A</b>	<b>Theory</b>	<b>58</b>
A.1	Maxwell's Equations . . . . .	58
A.1.1	In Vacuum . . . . .	58
A.1.2	In the Presence of External Media . . . . .	58
A.1.3	In Conducting Media . . . . .	59
A.2	Wave Equations . . . . .	59
A.2.1	In Vacuum and Linear Media . . . . .	59
A.2.2	In Conducting Media . . . . .	59
A.3	Wave Equation Solutions . . . . .	60

# List of Figures

1.1	A schematic of GlueX . . . . .	8
1.2	An electromagnetic wave . . . . .	11
1.3	An interference “fringe pattern” from a Michelson interferometer	12
1.4	A Michelson interferometer . . . . .	14
1.5	Transmission probability versus thickness of the silver conducting layer. . . . .	17
1.6	Phase shifts of the transmitted (lower) and reflected (upper) waves versus thickness of the conducting layer. . . . .	18
2.1	The Michelson interferometer with the reference mirror removed (a), and a fringe pattern obtained with a real diamond crystal in this configuration (b). . . . .	22
2.2	The Michelson interferometer including the reference mirror (a), and a fringe pattern obtained with a real diamond crystal in this configuration (b). . . . .	23
3.1	The Mexican hat function . . . . .	27
3.2	Schematic of a neural network . . . . .	29
5.1	Test problem fringe pattern . . . . .	41

6.1	The test interferogram created from three random surfaces (a), and the best solution found amongst several simulated annealing runs (b). . . . .	46
6.2	The front diamond surface used to generate the interferogram in the test problem (a), and the best solution found amongst several simulated annealing runs (b). . . . .	47
6.3	The back diamond surface used to generate the interferogram in the test problem (a), and the best solution found amongst several simulated annealing runs (b). . . . .	48
6.4	Log of probability of non-convergence versus log of run length for $\alpha = 0.5$ . . . . .	49
6.5	Log of probability of non-convergence versus log of run length for $\alpha = 0.9$ . . . . .	50

# List of Tables

1.1	Boundary conditions for electric and magnetic fields . . . . .	15
5.1	Parameter values used in the submitted jobs . . . . .	42
5.2	Highest order Legendre polynomial allowed for each amplitude and phase in the test problem. . . . .	42

# Chapter 1

## Background

### 1.1 Jefferson Lab and GlueX

#### 1.1.1 UConn and JLab

The University of Connecticut's Nuclear Physics Group is currently working in conjunction with Jefferson National Laboratory on a Department of Energy sponsored experiment called GlueX (Gluonic Excitations Experiment), depicted in Fig. 1.1. The purpose of GlueX is to measure gluonic excitations in the spectrum of nuclear matter in the hopes of finding novel particles known as exotic mesons. The results of this experiment will be used to test predictions made within the Standard Model of particle physics.

#### 1.1.2 The Beam Line

The GlueX experiment employs the collision between a high-energy gamma ray and a proton to serve as an abundant source of new particles called

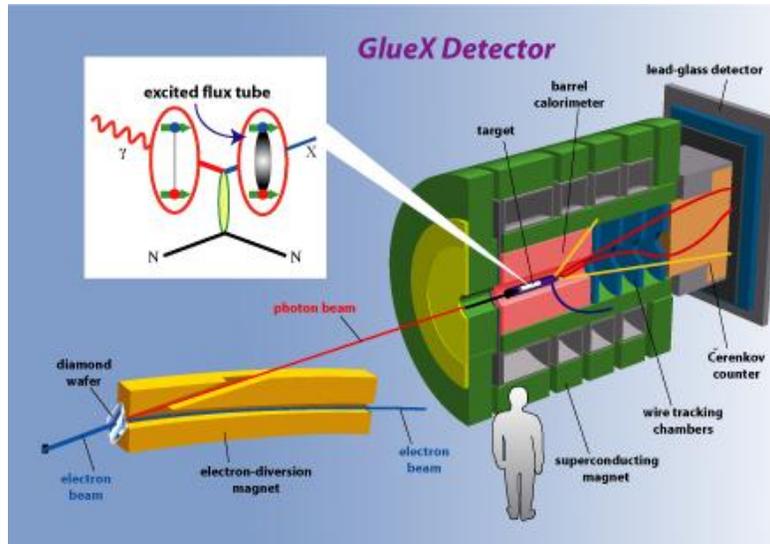


Figure 1.1: A schematic of GlueX

mesons. The way the gamma ray beam is formed is through a process known as *bremstrahlung*, which is the radiation produced by a high-energy electron upon its interaction with the Coulomb field of a nucleus in a substance. The photons of interest to GlueX need to have an extremely high energy and also a high degree of linear polarization. Conservation of energy requires that the incoming electrons must also be of high energy. The linear polarization of the *bremstrahlung* gamma rays is a consequence of the relative alignment between the electron beam direction and the normal vector to the crystal planes in the diamond radiator. To minimize the effects of multiple scattering on the collimation of the electron beam, it is important that the diamond be as thin as possible. It has been determined that the optimal thickness of the diamond *bremstrahlung* radiator to be used in GlueX is about  $20 \mu\text{m}$ .

### 1.1.3 Synthetic Diamond Wafers

Diamond is the hardest substance known to man. In the case of a *bremsstrahlung* radiator this property translates to its being radiation-hard, that is, resistant to crystal damages due to ionizing radiation. However, the diamonds must also be very thin - - on the order of 20  $\mu m$ . Manufacturing monocrystalline diamonds of this thickness is extremely difficult.

Jefferson Lab has obtained several synthetic diamond wafers from the company Element Six that specializes in the growth of large synthetic diamond monocrystals. Current manufacturing processes involve growing finger-sized ingots from a small seed, through a high-pressure, high temperature growth process. These diamonds are then cut into several thin slices, and then ground down to achieve the desired thickness. These wafers appear perfect when examined under a microscope, but they contain hidden stress due to defects that appeared during the growth process, and possibly during subsequent cutting and grinding steps. These stresses may cause the diamond to deform when its thickness is reduced to 20 microns, degrading its performance as a *bremsstrahlung* radiator.

### 1.1.4 Determining Diamond Deformities

The stresses and strains mentioned above can be measured by looking either at the surface of the diamond slices or by looking at the internal crystalline structure. Using a very large accelerator known as a synchrotron light source, x-ray diffraction experiments can be used to determine the latter, while a

simple table-top laser interferometer can be used to determine the former. The University of Connecticut Nuclear Physics Group has obtained beam time at the Cornell High Energy Synchrotron Light Source (CHESS) to perform measurements of the former type. While such measurements provide the most comprehensive information regarding the integrity of the diamond structure, they require considerable planning and human resources to carry out. In many cases, just knowing the shape of the diamond surface is sufficient to guide the manufacturing and mounting process. The goal of this research project is to develop means for extracting the shape of a diamond surface from the interference images of the crystal taken in a Michelson interferometer (see Sect. 1.3.2), with an optically flat mirror used as a reference.

## 1.2 Electromagnetic Theory

### 1.2.1 Electromagnetic Radiation

James Clerk Maxwell published his paper *On Physical Lines of Force* in 1861 which introduces the set of four equations that are known today as Maxwell's equations. Four years later he published *A Dynamical Theory of the Electromagnetic Field*. In this second paper, Maxwell derived the result that light is actually an electromagnetic wave, as depicted in Fig. 1.2 (courtesy of Ref. [1]).

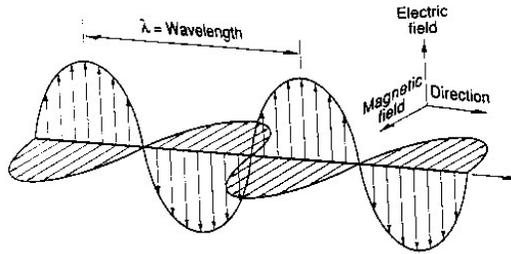


Figure 1.2: An electromagnetic wave

## 1.2.2 Superposition and Interference

Light waves (much like mechanical waves) have the property that their amplitudes add when they overlap. However, unlike mechanical waves, overlapping is not a sufficient criterion for interference. An extra condition arises from the vector nature of the oscillating electric and magnetic fields. In order to interfere, the light waves must have the same polarization.

## 1.3 Interferometry and the Michelson Interferometer

### 1.3.1 Interferometry

*Interference occurs when radiation follows more than one path from its source to the point of detection. It may be described as the local departures of the resultant intensity from the law of addition, for, as the point of detection is moved, the intensity oscillates about the sum of the separate intensities from each path. Light and dark bands are observed, called interference fringes.[2]*

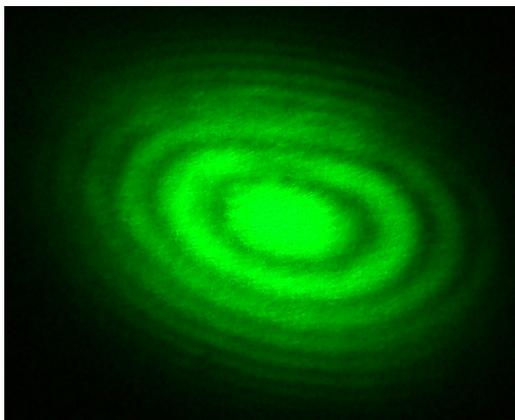


Figure 1.3: An interference “fringe pattern” from a Michelson interferometer

Interferometry is the splitting of a light beam into two or more paths and the recombining of those different beams in the manner described below, to measure the difference in optical path length between the two paths [3]. Fig. 1.3 (courtesy of Ref. [4]) shows an interference pattern from a Michelson interferometer (discussed in Sect. 1.3.2). As described by the above quote from Ref. [2], the light and dark bands result from the waves’ constructive and destructive interference. Light bands result when waves of the same polarization are in phase at the detector, whereas the darkest part of the dark bands result when the waves are  $\pi$  radians out of phase.

The simplest solutions to Maxwell’s equations in free space (see App. A.3) are plane waves. When two plane waves combine, the total electric field amplitude has the form:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{A}_1 e^{i(\mathbf{k}_1 \cdot \mathbf{r} - \omega t)} + \mathbf{A}_2 e^{i(\mathbf{k}_2 \cdot \mathbf{r} - \omega t)} \quad (1.1)$$

However, the human eye does not see amplitudes, rather it sees intensities (i.e. the absolute value squared of the amplitude):

$$I(\mathbf{r}, t) = |(\mathbf{E}(\mathbf{r}, t))|^2 = \left| \mathbf{A}_1 e^{i(\mathbf{k}_1 \cdot \mathbf{r} - \omega t)} + \mathbf{A}_2 e^{i(\mathbf{k}_2 \cdot \mathbf{r} - \omega t)} \right|^2 \quad (1.2)$$

The above expression is a mathematical representation of an interference “fringe pattern” such as the one illustrated in Fig. 1.3. The amplitudes  $\mathbf{A}_1$  and  $\mathbf{A}_2$  are generally complex and can be written in polar form:

$$\mathbf{A}_i = A_i e^{i\phi_i} \hat{\epsilon} \quad (1.3)$$

This allows one to write the complex amplitude in terms of a real amplitude ( $A_i$ ) and a real phase  $\phi_i$ . This phase term is important when extracting information from the fringe pattern, as path length differences between beams 1 and 2 result in the light and dark patterns described above.

If one views the interference pattern on a plane perpendicular to the direction of propagation of the wave, the viewing plane can be defined as  $z = 0$  and thus the intensity pattern can be thought of as an amplitude and a phase varying with  $x$  and  $y$  ( $A_i = A_i(x, y)$ ,  $\phi_i = \phi_i(x, y)$ ).

$$I(x, y) = \left| A_1 e^{i\phi_1} + A_2 e^{i\phi_2} \right|^2 \quad (1.4)$$

### 1.3.2 The Michelson Interferometer

The Michelson Interferometer was invented in 1882 by Albert A. Michelson. Its original purpose was to measure the difference in path length seen by

light beams travelling along perpendicular paths through the device in order to determine the motion of the experiment through the æther. The failure of Michelson’s experiment to detect absolute motion later helped to support Albert Einstein’s theory of special relativity. Michelson has been called the “father of visible light interferometry” and “was awarded in 1907 the Nobel prize in physics for ‘his optical instruments of precision and the spectroscopic and metrological investigations he has executed with them’ ” [2].

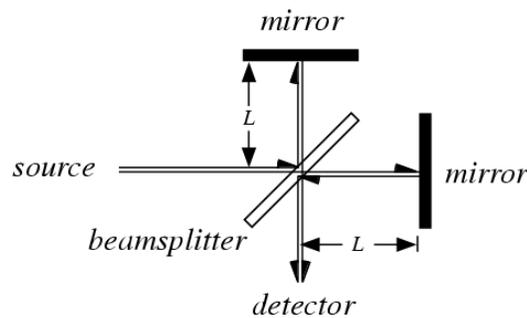


Figure 1.4: A Michelson interferometer

The Michelson interferometer is considered to be the epitome of a class of two-beam interferometers that feature a “division of amplitude” [2]. Fig. 1.4 (courtesy of Ref. [5]) depicts a diagram of a Michelson interferometer. In the diagram, a beam of light is emitted from a source at the left and directed toward a semi-transparent mirror (beam splitter) that is at an incidence angle of  $45^\circ$ . The front surface of the beam splitter is coated with a thin layer of some conducting material (e.g. silver) whose thickness is chosen to split the incident beam intensity equally into reflected and transmitted beams. Each of these beams travels to its respective mirror and is reflected back toward

the beam splitter. Back at the beam splitter each of the returning beams is split again, with one ray returning back toward the source and the other toward the detector. The two waves that are directed toward the detector interfere and produce a fringe pattern, while the waves that are directed back toward the source are absorbed.

$$\begin{aligned} D_1^\perp - D_2^\perp &= \sigma_{free} & \mathbf{E}_1^\parallel - \mathbf{E}_2^\parallel &= 0 \\ B_1^\perp - B_2^\perp &= 0 & \mathbf{H}_1^\parallel - \mathbf{H}_2^\parallel &= \mathbf{K}_{free} \times \hat{n} \end{aligned}$$

Table 1.1: Boundary conditions for electric and magnetic fields

It is interesting to know the phase shifts that occur when the beam of light is both reflected and transmitted through the conducting layer of the beam splitter. In order to determine these quantities, one must solve Maxwell's equations using the boundary conditions (Table 1.1) for an interface between two media. For the case considered here, there are no free charges or currents at the interface, so  $\sigma_{free} = 0$  and  $\mathbf{K}_{free} = \mathbf{0}$ . For the sake of simplicity one can take the angle of incidence to be normal to the beam splitter. Despite the fact that the actual Michelson interferometer utilizes a beam that is at oblique incidence to the beam splitter, all of the essential features of the problem are retained with fewer complications if the incidence angle is taken to be 0 degrees.

Using the boundary conditions and the solutions to the wave equations in homogeneous media, one obtains a system of linear equations that can be solved as follows.

$$Mv = b \quad (1.5)$$

where

$$M = \begin{bmatrix} -1 & 1 & 1 & 0 \\ Z_1^{-1} & Z_2^{-1} & -Z_2^{-1} & 0 \\ 0 & e^{ik_2a} & e^{-ik_2a} & -e^{-ik_1a} \\ 0 & Z_2^{-1}e^{ik_2a} & Z_2^{-1}e^{-ik_2a} & -Z_1^{-1}e^{-ik_1a} \end{bmatrix} \quad (1.6)$$

$$v = \begin{bmatrix} E_r \\ E_f \\ E_b \\ E_t \end{bmatrix}; \quad b = \begin{bmatrix} 1 \\ Z_1^{-1} \\ 0 \\ 0 \end{bmatrix} \quad (1.7)$$

with the solution  $v = M^{-1}b$ . It should be noted that  $a$  is the thickness of the conducting layer,  $Z_1 = \sqrt{\frac{\mu_0}{\epsilon_0}}$  and  $Z_2 = \sqrt{\frac{\mu_0}{\epsilon_{Ag}}}$ . The components of the vector  $v$  are as follows:

- $E_r$  is the amplitude of the reflected wave,
- $E_f$  is the amplitude of the forward moving wave,
- $E_b$  is the amplitude of the backward moving wave,
- $E_t$  is the amplitude of the transmitted wave.

The incident wave number  $k_1 = 2\pi/\lambda$  where  $\lambda$  is the wavelength emitted by the monochromatic source. Inside the conductor the wave is described by a complex wave number  $k_2 = \sqrt{n_2^2k_1^2 + ik_1\sigma Z_1}$  where  $n_2$  is the refractive index

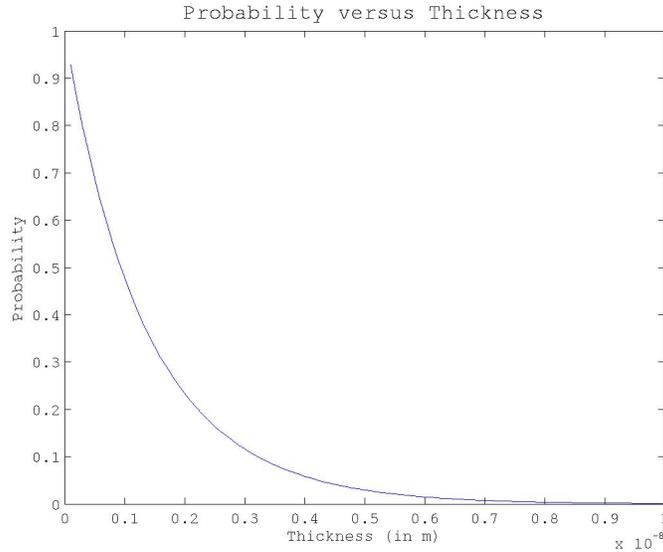


Figure 1.5: Transmission probability versus thickness of the silver conducting layer.

of the conducting material at frequency  $\omega = k_1 c$  and  $\sigma$  is the corresponding conductivity. Fig. 1.5 shows the probability of transmission versus thickness. Fig. 1.6 shows the phase shift of both the transmitted and reflected waves versus thickness. The thickness that corresponds to 50% transmission occurs around  $1 \text{ nm}$ . The corresponding phase shifts for this thickness can then be determined from Fig. 1.6.

## 1.4 Surface Resolution Approximation

An important part of topological interferometry is that the surface profile is imprinted on the phase of the wave reflected from the sample surface. Plane wave solutions can be used when the height of surface features is much

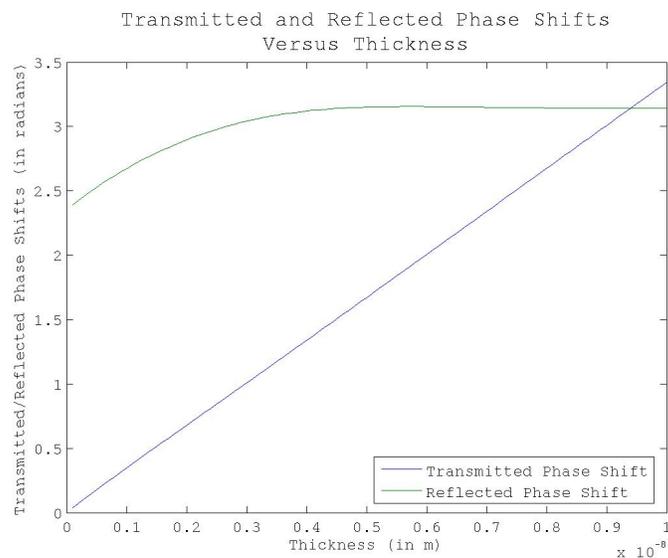


Figure 1.6: Phase shifts of the transmitted (lower) and reflected (upper) waves versus thickness of the conducting layer.

smaller than the features' transverse size. Huygen's principle can be used to estimate the distance the reflected wave propagates before there is significant smearing due to transverse diffusion of the phase gradient.

### 1.4.1 Huygen's Principle

Huygen's principle states that each point on a wavefront (surface of constant phase in a propagating wave) can be treated as the source of an outgoing spherical wave. When a plane wave is incident on a curved surface, the shape of the surface is imprinted on the wavefronts of the reflected beam, but the details of the shape diffuse as the reflected wave propagates. For nearly flat surfaces this diffusion will only occur gradually, with fine structures being

smear out first and larger-scale structures persisting to longer distances.

In a quasi-planar model of the diamond surface, Huygen's principle can be used to determine the minimum feature size on the surface profile that can be resolved using light of wavelength  $\lambda$ , after the reflected beam has traveled a distance  $L$ . The quasi-planar model of the diamond surface states that the feature height of the diamond's surface is much smaller than the characteristic transverse size of surface features. Under this assumption, the diamond surface is approximated by a mosaic of flat tiles that are joined together. Light waves reflected from these tiles can be approximated as the superposition of many beams of light (beamlets) emitted by the individual tiles. The divergence angle  $\theta$  of each beam is given by the Rayleigh diffraction limit

$$\theta d = \lambda \tag{1.8}$$

where  $d$  is the diameter of the tiles. As each beamlet travels away from the surface, its radius grows in a conical fashion. Two tiles of equal size that are side-by-side will no longer be able to be resolved after the edge of one beam begins to overlap with the center of its neighboring tile. The distance  $L$  up to which a feature of size  $d$  can still be resolved is given by

$$\theta = \frac{d}{L} \tag{1.9}$$

From Eq. 1.8 and Eq. 1.9 a relationship can be obtained that gives the size of resolvable features on the diamond surface that remain after the reflected wave has traveled a distance  $L$ .

$$d = \sqrt{L\lambda} \tag{1.10}$$

For example, if the total distance between the sample and the detector in the Michelson interferometer is  $10\text{ cm}$  and the wavelength of the light used is  $600\text{ nm}$ , then the size of resolvable features,  $d$ , of the diamond surface is  $2.5 \times 10^{-4}\text{ m}$ . Features below this length scale will not be resolved. X-ray diffraction measurements as described in Sect. 1.1.4 have shown that the important structural features of the diamond wafers under study in this project have a typical length scale of  $1\text{ mm}$ , indicating that an interferometer with  $L \leq 40\text{ cm}$  is suitable for these measurements.

# Chapter 2

## Methods

### 2.1 The Michelson Interferometer

Fringe patterns from the synthetic diamond wafers were created in University of Connecticut Professor George Gibson's optics lab using his Michelson interferometer. The diamonds were mounted in place of one of the optical mirrors and images were taken with a CCD camera. The CCD camera had a pixel size of  $20\ \mu\text{m} \times 20\ \mu\text{m}$ . The light source used was a helium-neon laser with a peak wavelength of  $633\text{nm}$ . The distance from the sample to the camera was  $10\text{cm}$ . Within the quasi-planar approximation described in Sect. 1.4, surface features of transverse size  $250\ \mu\text{m}$  or larger are resolved in the images.

Three sets of images were taken. The first set included only the reference beam and displayed a nearly uniform intensity profile with a few minor optical defects coming from dust particles on the optical elements. The second set of images were taken with the reference beam removed and exhibits in-

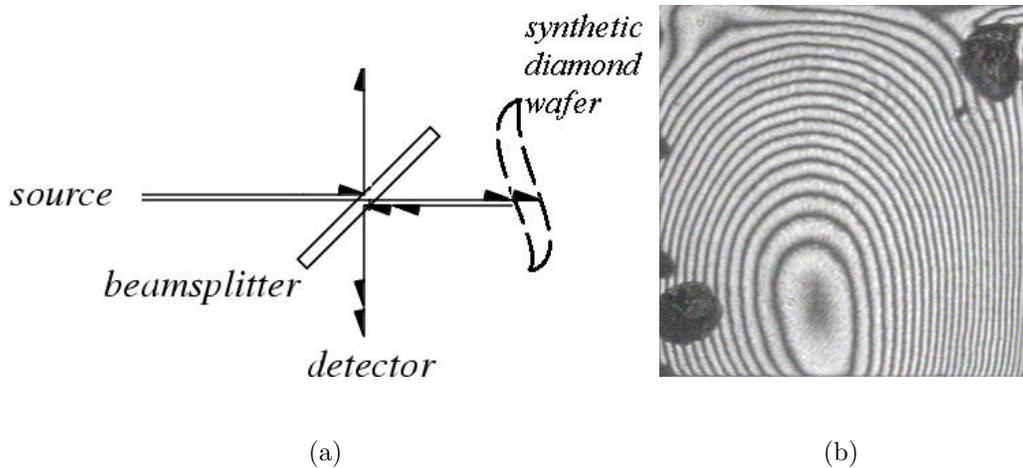


Figure 2.1: The Michelson interferometer with the reference mirror removed (a), and a fringe pattern obtained with a real diamond crystal in this configuration (b).

interference between reflections coming from the front and back surfaces of the diamond wafer (Fig. 2.1). The third set of images exhibit interference between both the reference mirror reflection and those from the front and back surfaces of the diamond wafer (Fig. 2.2). In accordance with the quasi-planar approximation, each of the reflecting surfaces appear at the detector as a plane wave with its surface height multiplied by the laser wavenumber imprinted on the phase of the wave, amplified by a factor of two. This amplification results from the wave having to travel both to the surface and back from it following reflection. Surfaces of constant phase in the reflected beam can thus be considered to be a map of the physical surface

$$\phi(x, y) = 2S(x, y) \tag{2.1}$$

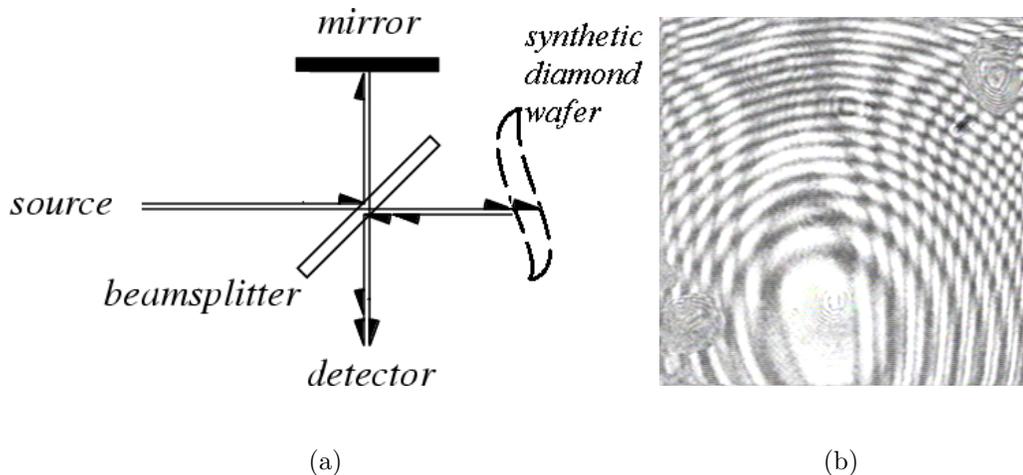


Figure 2.2: The Michelson interferometer including the reference mirror (a), and a fringe pattern obtained with a real diamond crystal in this configuration (b).

where  $\phi$  is the local phase of the quasi-plane wave and  $S$  is the shape of the physical surface. Each quasi-plane wave also has a corresponding amplitude ( $A = A(x, y)$ ) to accompany the phase.

## 2.2 Two-Wave Interference

The fringe pattern shown in Fig. 2.1 exhibits the interference of the reflected waves from the front and back surfaces of the diamond wafer. This 2-wave interference pattern can be expressed as

$$I_2 = \left| A_{front} e^{i\phi_{front}} + f A_{front} e^{i\phi_{back}} \right|^2 \quad (2.2)$$

where the factor  $f = 0.9$  takes into account the loss of intensity in the beam

reflected from the back surface relative to the front due to the additional two boundary crossings. A fringe pattern of this type is not sensitive to the two surface profiles individually, but only to variations in the thickness of the diamond wafer. While this quality does not help in determining the absolute shape of the diamond wafer, it does provide important information on the uniformity of the wafer-thickness resulting from the thinning process.

### 2.3 Three-Wave Interference

Fig. 2.2 shows the interference of the light waves reflected from the front and back surfaces of the diamond, as well as the reference mirror. Its interference pattern has the form

$$I_3 = \left| A_{ref} e^{i\phi_{ref}} + A_{front} e^{i\phi_{front}} + f A_{front} e^{i\phi_{back}} \right|^2 \quad (2.3)$$

which includes an additional term within the absolute value squared. Unlike the two-wave pattern in Sect. 2.2, the three-wave fringe pattern contains information on the non-planarity of the individual diamond surfaces in addition to the non-uniformity of its thickness.

# Chapter 3

## Fringe Pattern Analysis

### 3.1 Introduction

Creating simulated fringe patterns is a relatively easy task. All one needs are two or three model surfaces and a computer program to interfere the simulated wavefronts. Each combination of surfaces makes a single fringe pattern. However, the task of going from a fringe pattern back to the original surfaces is not as easy and, in general, not unique. Many different sets of surfaces can lead to the same fringe pattern. Only by imposing certain smoothness criteria on the shapes of the surfaces can the inverse problem be made well-defined.

Methods of fringe pattern analysis can be sorted into two main categories: “temporal (phase-shifting) methods and spatial methods” [6]. According to Ref. [6], interferogram analysis began in the 1960’s with the work of Carre, Rowley and Harmon, who pioneered the temporal method. However, this

form of analysis relies on the availability of many images taken while the sample is shifted by a fractional number of wavelengths precisely along the interferometer beam axis. To circumvent this problem, a number of spatial analysis methods were developed which were capable of analyzing single fringe patterns to obtain the underlying surface shapes. With advances in both experimental control and computing power, current methods of fringe pattern analysis utilize combinations of both experimental and computer-based procedures to produce more accurate results than were previously attainable [6].

## 3.2 Fringe Pattern Analysis Methods

### 3.2.1 phase-shifting Technique

The phase-shifting technique relies on being able to adjust the offset position of the sample in the interferometer in such a way that the difference in the phase between two interference patterns is known. For each pixel in the image, the intensity varies sinusoidally with  $k \delta z$  where  $\delta z$  is the offset coordinate and  $k$  is the laser wavenumber. Generally the maximum and minimum of the interference fringes are *a-priori* unknown, together with the phase offset  $\phi$  at  $\delta z = 0$ . Thus a minimum of three of these offset images (“frames”) is needed to uniquely extract the values  $\phi$  for each pixel.

With the availability of improved precision mirror mounts, the phase-shifting technique with sub-micron translation stages, has re-emerged as a practical

method, as the ability to measure and implement the phase shift comes under much greater control. However, this technique does require a more expensive experimental setup, and requires some experience in optical alignment to carry out correctly.

### 3.2.2 Fourier-transform Method

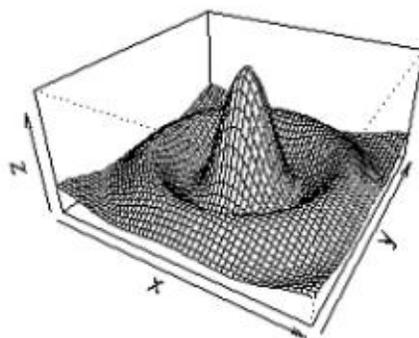


Figure 3.1: The Mexican hat function

The Fourier-transform method of interferograms was created in 1982 by the authors of Ref. [7]. It was originally intended as an alternative to Moire Topography and the phase-shifting technique [7], [8]. However, as originally proposed this method was ineffective at analyzing closed fringe patterns (i.e. those whose intensity map contain closed contours). A revision to the method solved this problem by utilizing a Cartesian-to-polar coordinate transform [9], the result of which could then be analyzed using the original method proposed.

The revised Fourier Analysis method does have several limitations. The first requirement is that the “measurement wave front be a monotonic function in the direction of the carrier frequency” [9]. The notion of a “carrier frequency” refers to the regular patterns of nearly straight equi-distant fringes that appear in an interferogram when the normal vectors of the sample and mirror are deliberately displaced in angle. In this case, departures from straight and equidistant fringes are the signature of the sample surface profile. For instance, if a surface resembling the Mexican hat function shown in Fig. 3.1 (courtesy of Ref. [10]) were analyzed by the revised Fourier Analysis method, the result would look like an inverted cone with a rounded top. In order to analyze a fringe pattern generated by such a surface, an additional fringe pattern giving the carrier frequency must be given as an additional input.

### **3.2.3 Regularization Algorithms**

The regularization method was created for the specific purpose of automatically demodulating “noisy” fringe patterns. Regularization algorithms involve evaluating the estimated phase field with a cost function against the actual image and then imposing the smoothness criterion. This method is begun in a region of the image where the intensity is locally flat, and then extended into adjacent regions until the entire image is covered and a global minimum is reached in the cost function [11], [12].

The algorithm is comparable to the way crystal growth occurs, starting from a seed. One drawback of the method is that it is unable to take advantage of global information to resolve local ambiguities in whether the surface is

locally increasing or decreasing height.

### 3.2.4 Artificial Neural Network Method

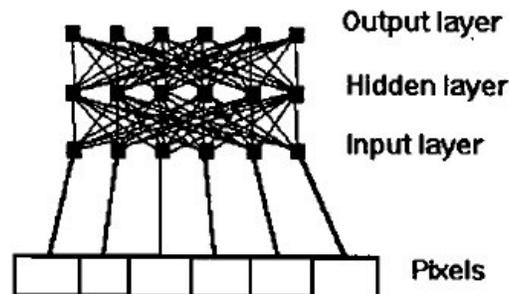


Figure 3.2: Schematic of a neural network

The human brain is composed of nearly 5 billion neurons, each of which have the apparently simple job of receiving, integrating and transmitting nerve pulses. The complex interconnections between many functional units give rise to the functionality of the brain. This principle was postulated in the 1940's by McCulloch and Pitts who theorized that a similar approach could be applied to computing [13]. Using a network of simple independent computing units (“neurons”), one might be able to mimic the brain and thereby create complex behavior from simple components.

Differing from most computer programs, artificial neural network software requires a “learning” phase to adapt itself to the problem being undertaken. The neural network organizes the neurons into three distinct layers: input neurons, output neurons, and hidden neurons, as depicted in Fig. 3.2 (cour-

tesy of Ref. [13]). It is important that neurons in any layer communicate only with neurons in an adjacent layer. After their creation, the network of neurons is “trained” under one of three training regiments: supervised, reinforced or unsupervised learning. Supervised learning involves constant feedback being given to the neural network during the training sequence. Reinforced learning can be defined by the use of simple “good” and “bad” evaluation to the program after each run. A neural network with unsupervised learning is given an output goal, but receives no feedback from the trainer during the learning phase. Given time constraints and the desire for a reasonable output, supervised or reinforced learning schedules are usually adopted. The result of the learning phase is a set of weight factors by which each neuron computes the average of its inputs to form its output [13].

Neural networks are frequently employed for problems in complex pattern recognition, inspired by the impressive powers of the brain in this area. The method can either involve a system of many neurons linked together to analyze an entire image, or a small number of neurons can be used to analyze the image section by section. The former requires both a long learning period and a large number of neurons. Thus it is both computationally expensive and time intensive. Sub-image analysis with artificial neural networks can be accomplished with much more limited computer resources, but is suitable for only a restricted set of problems [13].

# Chapter 4

## Simulated Annealing

### 4.1 Optimization

Optimization can be defined as the “identification of global extremities of mathematical functions” [14]. The task of finding such a global extremity can range from very easy to very difficult, as the complexity of the function in question grows. Optimization problems arise in a vast array of fields that include biology, chemistry, engineering, computer science, physics and many others. Simple optimization problems can be solved using elementary calculus, and, as the difficulty increases, variational principles may be employed. However, at a certain point, analytical solutions may no longer be possible and computer algorithms must be employed. There exist an extremely large number of these computer algorithms, ranging from deterministic search algorithms to various Monte-Carlo techniques [15].

## 4.2 History

The first paper on simulated annealing was published in 1983 by Kirkpatrick, Gelatt, and Vecchi, who sought to emphasize the “deep and useful connection between statistical mechanics . . . and multivariate or combinatorial optimization” [16].

## 4.3 Method

Simulated annealing is based on the fact that in physical systems whose ground states are highly ordered are able to self-assemble starting from disordered states if their temperature is lowered gradually, starting from a high enough initial temperature. In the case of physical annealing, it is the Boltzmann factor  $e^{-E/k_bT}$  that governs the relative probability that any state of free energy  $E$  is attained by the system in equilibrium at temperature  $T$ . In more general optimization problems, the state of the system is represented by a candidate solution called a configuration, and the energy of the configuration is represented by a “cost function” whose global minimum defines the true solution. In statistical mechanics, the system undergoes a random walk in configuration space, so that it visits all possible states of the system with some probability, but as the temperature is reduced, states with energies near the ground state become increasingly probable. In simulated annealing, a similar random walk takes place, whose steps are governed by the Boltzmann factor so that it minimizes the physical system, and at low temperature is increasingly confined to configurations of the lowest cost.

## 4.4 The Metropolis Algorithm

The Metropolis Algorithm was created in 1959, by Metropolis et al. in an effort to model a collection of atoms in contact with a heat bath [16]. The Metropolis algorithm is based on the representation of a state consisting of  $N$  interacting atoms by a point in  $6N$ -dimensional phase space. The temporal evolution of the system in phase space is replaced with a chain of points (configurations) whose sequence is the result of a Monte Carlo selection procedure. The energy of a configuration in the chain is compared with the energy resulting from a random perturbation. If the perturbed energy is lower, that displacement is accepted and the chain continues from the new point. However if the value of the perturbed energy is higher, the value of the Boltzmann factor ( $e^{-\Delta E/k_b T}$ ) is compared to a uniform random number between 0 and 1. If the random number is greater than the value of the Boltzmann factor, the displacement is accepted, and if not it is discarded [16].

A variation of this algorithm is employed by Ref. [16], where the energy is replaced by a quantity known as the cost function (the value to be optimized) and the temperature is replaced by a controlled quantity that mimics the physical process of annealing. Larger increases in the cost function, particularly when the temperature is low, result in a Boltzmann factor closer to 0, and thus are less likely to be accepted.

## 4.5 Parallel Simulated Annealing

Simulated annealing has been shown to be an effective means for finding a near-global minimum for large combinatorial optimization problems [16]. Compared with deterministic searches, simulated annealing typically requires much greater computing resources. Since users of optimization algorithms want to find solutions in the most time-efficient manner without sacrificing the quality of the solution, simulated annealing is often overlooked. However, if one can parallelize the algorithm to run on multiple processors, the amount of time that it takes to find an acceptable solution is decreased.

The Parallelized Simulated Annealing (ParSA) library created by Georg Kliewer and Karsten Klohs [17] provides a general framework in C++ for implementing simulated annealing on a parallel computing platform. Two modules present within the ParSA package are the *SA\_Scheduler* and the *SA\_Solver*. Within the *SA\_Scheduler* module, the user is able to choose between a variety of annealing strategies, each of which is supported by a *SA\_Solver* module which interfaces to the user's classes where the specifics of the optimization problem are implemented.

The parallel processing strategy used by ParSA is as follows. At each temperature setting, several independent chains are generated, each of which is worked on by a set of cooperating processes called a cluster. Each processor in the cluster is called a node. One special node known as the head node, collects cost function information from the other nodes, called slave nodes. The head node communicates with the slave nodes using the MPI library,

which is a “specification for message-passing, proposed as a standard by a broadly based committee of vendors, implementors, and users. . . designed for high performance on both massively parallel machines and on workstation clusters” [18].

A simulated annealing run consists of two phases, the warming-up phase and the cooling phase. Each of these periods are punctuated by the equilibrium point and the frozen point, respectively. The warming-up phase ends when the system is said to reach equilibrium. The cooling phase ends when the system is frozen. The purpose of the warming-up phase is to ensure that the final solution does not depend on the initial configuration, while the cooling phase determines the quality of the solution that results.

### 4.5.1 Aarts Strategy

#### Scheduler

Aarts temperature scheduling is done completely adaptively, that is, the temperature of each portion of the run is controlled by feedback from the history of the run. In the warming-up phase, an initial acceptance ratio  $\chi_0$  is chosen by the user, together with a starting configuration. The region around the starting configuration is explored by taking many perturbations of the initial configuration, and comparing the cost function with the initial value. At the end of the warming phase, an initial temperature  $T_0$  is computed as

$$T_0 = \overline{\Delta C^{(+)}} \left( \ln \frac{m_2}{m_2 \chi_0 - (1 - \chi_0) m_1} \right)^{-1} \quad (4.1)$$

where  $m_1$  and  $m_2$  are the number of neighboring configurations with better and worse cost function values, respectively, and  $\overline{\Delta C^{(+)}}$  is the mean value of the differences between the cost function of all worse solutions [17]. After a stable value is achieved for  $T_0$  equilibrium is reached and phase two begins. The temperature in phase two is reduced according to

$$T_n = T_{n-1} \left( 1 + \frac{\ln(1 + \delta)T_{n-1}}{3\sigma T_{n-1}} \right)^{-1} \quad (4.2)$$

where  $\sigma(T_{n-1})$  is the standard deviation of the values of the cost function encountered during the current temperature step and  $\delta$  is the so-called *distance parameter*. The size of  $\delta$  determines the speed of the reduction of the temperature [17]. An Aarts run is said to be frozen when the derivative of the smoothed mean value of the cost function is smaller than the user-defined error  $\epsilon$  [17].

## Solver

The solver portion of the Aarts scheduler follows a divide-and-conquer strategy in which the head node keeps track of all of the best solutions presented by each of the slave nodes, and at each temperature step the slave nodes each follow an independent Metropolis chain. At the end of each job, the best solution cataloged by the head node is given as the solution of the run. It should be noted that at the beginning of each temperature step, all of the slaves adopt a common configuration as their starting point.

## 4.5.2 MIR

### Scheduler

The warming-up phase of the MIR scheduler proceeds in an adaptive manner similar to the Aarts scheduler. At the end of this period, the starting and ending temperature is defined according to the following formulas:  $T_{start} = \frac{\Delta C_{max}}{\chi_0}$  and  $T_{end} = \frac{\Delta C_{min}}{\chi_0}$ . If the user sets the *end temperature* parameter, this value is used in place of the calculated one. After this, the job enters phase two. In phase two, the temperature is cooled according to a geometric sequence defined as:  $T_n = (\alpha) \times T_{n-1}$ , where  $\alpha \in (0, 1)$  is selected by the user. Higher values of  $\alpha$  result in the temperature being lowered more slowly, whereas lower values of  $\alpha$  result in quicker cooling.

Both higher and lower values of  $\alpha$  have their potential merit. Higher values allow more of the solution space to be searched while holding the number of steps at each temperature constant. This ensures that the initial configuration is not reflected in the final configuration and that a greater portion of the configuration space is searched. However, the downside of this lies in the fact that if the temperature is kept relatively high for the duration of the run, the algorithm does not get a chance to “down-climb” into a “valley” in solution space and settle upon a potentially good solution. On the other hand, lower values of  $\alpha$  result in the temperature being lowered very quickly and less hill-climbing being done. Thus, the run descends in solution space more quickly than its cousin with a higher  $\alpha$ . This strategy has the downside that too much quenching of the temperature may result in not

enough of the solution space being explored.

For very large or complex problems it is often more effective to use multiple runs where the temperature is rapidly quenched, as opposed to using a single run which cools more slowly. There are two reasons why quicker multiple runs may be more effective. The first reason is that slowly cooled runs might spend too much time wandering before settling into the neighborhood of a good solution. The second reason is that search spaces are often highly segmented, and success in a search can often depend on taking the correct branch in the search sequence. By using many runs that are cooled quickly, a large number of solutions will be found, some of which will be within the set of what are deemed good solutions i.e. solutions with a cost function value less than  $Cost_{min}$ . The probability that a solution  $\chi_n$  found by an MIR run will not be within this set is given by

$$P(\chi_n \notin Cost_{min}) \sim \left(\frac{K}{n}\right)^\alpha \quad (4.3)$$

$n$  is the number of steps in the run, and  $K$  and  $\alpha$  are parameters that depend on the problem. Once these parameters are determined empirically, it is then possible to estimate the optimum run length required to obtain an acceptable solution.

## **Solver**

The number of runs completed in phase two depends on the parameters *Betta\_Runtime*, *Minimum\_Runlength*, *Maximum\_Runlength* and *Samples*. The first three of these parameters determine the number of runs per sam-

ple. The first run in a sample is of length *Minimum\_RunLength*, which is a user-defined quantity which serves as a reasonable approximation of the neighborhood size in the solution space. For subsequent runs in the sample, the run length is increased by the factor *Betta\_RunTime* until *Maximum\_RunLength* is reached, at which point the sample is complete. This sequence is repeated *Samples* times to complete the optimization job.

# Chapter 5

## Testing

### 5.1 Solutions

Now that both the problem (deciphering interferograms) and a proposed method of analysis (simulated annealing) have been presented, a parametrized model of an interferogram must be created so that ParSA can be used to find the interfering surfaces. As mentioned in Sect. 1.4, the most interesting features of the diamond wafers are of order 1 *mm* in transverse size. Since the diameter of the diamond is around 5 *mm*, these features are not discontinuous, but smooth and gradual. This physical insight can help determine what mathematical basis set to use to describe the diamond surface (i.e. the amplitude and phase terms in Eq. 2.3) so that simulated annealing may be used to extract the fringe pattern. One such basis is the Legendre polynomials. Solutions to each of the amplitude and phase terms can then be thought of as a weighted sum of the elements of a matrix of the Legendre polynomials

$$P_{ij} = a_{ij}P_i(x)P_j(y) \tag{5.1}$$

where  $a_{ij}$  is the coefficient of the  $ij^{th}$  element of the matrix  $P_{ij}$ ,  $P_i(x)$  and  $P_j(y)$  are Legendre polynomials of order  $i$  and  $j$  in the  $x$  and  $y$  directions, respectively, and  $i, j = 1, \dots, n$  where  $n \in \mathbf{N}$ .

In Sect. 2.1 it was determined that the reference mirror was almost perfectly flat. This planarity makes it convenient to define the phase of the reference mirror as zero. With this simplification, only two amplitude and two phase terms remain in Eq. 2.3 to be optimized against the image data.

## 5.2 Test Problem

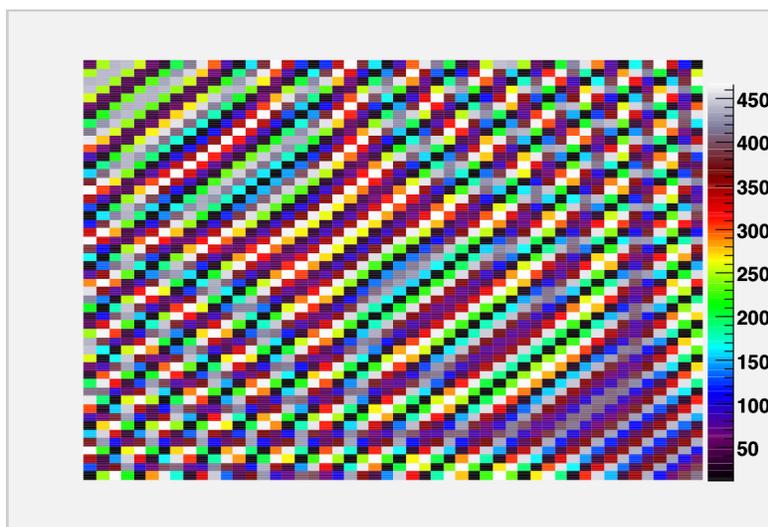


Figure 5.1: Test problem fringe pattern

A  $50\text{ px} \times 50\text{ px}$  test interferogram was created from three model surfaces, as depicted in Fig. 5.1. A configuration file was arranged for the MIR scheduler

and solver, with parameters given in Table 5.1.

<i>Betta_Runtime</i>	1.1
<i>Maximum_RunLength</i>	$RunFactor * GetLocalN()$
<i>Minimum_RunLength</i>	$GetLocalN()$
<i>Samples</i>	10
<i>RunFactor</i>	5
<i>GetLocalN()</i>	$10^5$

Table 5.1: Parameter values used in the submitted jobs

The highest allowed order of Legendre polynomial allowed for each of the amplitude and phase terms is given in Table 5.2.

$A_{ref}$	1
$A_{diamond}$	1
$\phi_{diafront}$	3
$\phi_{diaback}$	3

Table 5.2: Highest order Legendre polynomial allowed for each amplitude and phase in the test problem.

A chi-squared value was picked as the cost function for comparisons between the test interferogram and candidate solutions found by the simulated annealing algorithm.

### 5.3 Running ParSA

ParSA jobs with the *alpha* parameter set to 0.5 and 0.9 were submitted. It was predicted that the job with the more aggressive cooling schedule (i.e. the job with *alpha* = 0.5) would show better performance in finding solutions to the test problem, according to the argument presented in Sect. 4.5.2. This hypothesis could be tested by determining the  $\alpha$  and  $K$  parameters given in Eq. 4.3. Relative performance under each of the cooling parameters could then be evaluated by determining which job would converge most efficiently.

In order to create plots to determine the  $\alpha$  and  $K$  parameters, a maximum acceptable cost function value ( $Cost_{min}$ ) must be chosen. The initial configuration for all of the simulated annealing runs assumed uniform  $A_{ref}, A_{dia} = 1$  and  $\phi_{dia1}, \phi_{dia2} = 0$ , which resulted in a cost function of  $10^8$ . Solutions with cost function values less than 1% of this initial value are visually indistinguishable from the actual test problem. Thus,  $10^6$  was determined to be a valid value for  $Cost_{min}$ .

With the maximum acceptable value for the cost function defined, the probability that a run of length  $n$  will not converge can then be found using binomial statistics. The binomial distribution  $B(n, N; p) = \binom{N}{n} p^n (1-p)^{N-n}$  gives the probability of observing  $n$  successes out of  $N$  trials, given the mean success probability  $p$ . In binomial statistics, one takes  $N$  and  $n$  as given by experiment and uses them to estimate  $p$ . The Bayesian estimator for  $p$  given

$N$  and  $n$  is

$$\hat{p} = \frac{n_f + 1}{N + 2} \quad (5.2)$$

where  $N$  is the total number of runs of length  $n$  and  $n_f$  is the number of failed runs. The variance of the estimator  $\hat{p}$  is given by

$$V(\hat{p}) = \frac{(n_f + 1)(N - n_f + 1)}{(N + 2)^2(N + 3)} \quad (5.3)$$

which provides an estimate for the statistical error in the value of the estimator  $\hat{p}$  for the sample of  $N$  runs.

A log-plot of  $\hat{p}$  versus  $n$  can be obtained and  $\alpha$  and  $K$  can be determined from the slope and y-intercept of a linear fit. The standard deviation of each measurement can be determined by the following relationship

$$\sigma = \sqrt{V(\hat{p})} \quad (5.4)$$

where  $\sigma$  is the standard deviation on the measured values for  $\hat{p}$ . The linear fit returns the best values for  $\alpha$  and  $K$  based on these data, together with their respective errors.

# Chapter 6

## Results and Conclusions

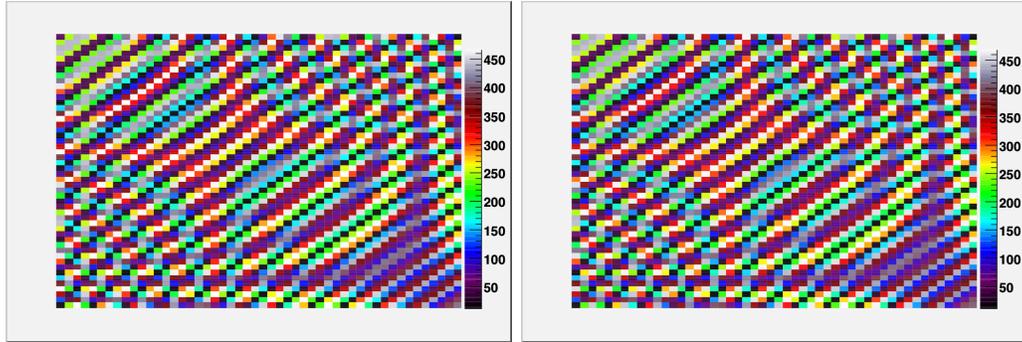
### 6.1 Visual Comparison

#### 6.1.1 Interferograms

Fig. 6.1(b) shows the best solution found during several simulated annealing runs on the three-wave test problem described in Chap. 5. A visual comparison with the exact solution depicted in Fig. 6.1(a) shows that the solution is a good one. Solutions with cost functions of the same order of magnitude were repeatedly found. No two exactly were alike, but all had a similar resemblance to the input image.

#### 6.1.2 Surfaces

Fig. 6.2(a) and Fig. 6.3(a) show the surface profiles for the front and back of the test diamond, respectively, where the incident light travels down along the vertical axis in the figures. Fig. 6.2(b) and Fig. 6.3(b) show the corre-



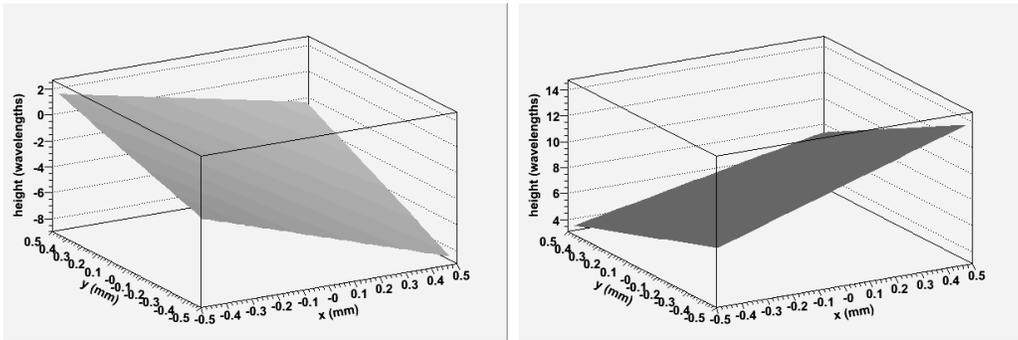
(a)

(b)

Figure 6.1: The test interferogram created from three random surfaces (a), and the best solution found amongst several simulated annealing runs (b).

sponding solutions found during the best simulated annealing run described in Sect. 6.1.1. The test diamond surfaces are the same shape and give the diamond a uniform thickness. The surfaces appear to be planar, but in actuality they are not, as they have a small contribution from the  $P_{22}$  element in the Legendre matrix sum (Eq. 2.3), which gives it some curvature. This curvature can be seen in the test interferogram in the form of the complex light and dark band structure. If the surfaces were to have been completely planar, the resulting interferogram would only have featured a series of straight equally-spaced parallel light and dark bands. As noted in Sect. 6.1.1, the interferogram solutions found by simulated annealing were nearly identical to the test interferograms. This is surprising because the solutions for the surfaces are actually sloped in the opposite direction! This sign error results from the ambiguity present in the problem due to the fact that a mirror transformation of the diamond through a horizontal plane results in an unchanged

interferogram. The spacing between the two solution diamond surfaces also features an ambiguity resulting from the periodic nature of light. As depicted in the solution, the surfaces are separated by about 12 wavelengths. If the algorithm were to have found a solution with the same surface features but with either of the surfaces displaced in height by an integral number of wavelengths, the solutions would be equivalent to the ones shown.



(a)

(b)

Figure 6.2: The front diamond surface used to generate the interferogram in the test problem (a), and the best solution found amongst several simulated annealing runs (b).

## 6.2 Run Length Optimization

For each of the jobs described in Sect. 5.3, the Bayesian estimator  $\hat{p}$  for the probability of non-convergence versus run length  $n$  was obtained, together with its statistical error. Run lengths  $n = 10^5$  to  $n = 4.5 \times 10^5$  were analyzed for two different values of the cooling parameter  $alpha$ . A linear fit was

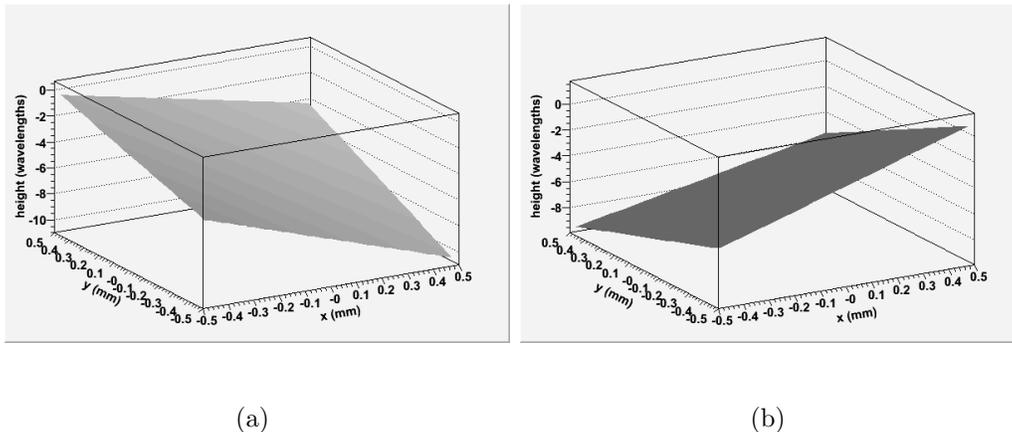


Figure 6.3: The back diamond surface used to generate the interferogram in the test problem (a), and the best solution found amongst several simulated annealing runs (b).

obtained for each of the log-log graphs, of the form

$$\ln \hat{p} = \alpha \ln K - \alpha \ln n \quad (6.1)$$

where  $\hat{p}$ ,  $n$ ,  $K$  and  $\alpha$  are as described in Sect. 5.3. By finding the slope and  $y$ -intercept of the linear fit, a best estimate for  $K$  and  $\alpha$  were obtained, together with their errors.

### 6.2.1 MIR performance

The plot for  $alpha = 0.5$  is shown in Fig. 6.4. The linear least-squares fit has a slope of  $-0.0273 \pm 0.0103$  and a  $y$ -intercept equaling  $0.297 \pm 0.127$  with a reduced  $\chi^2$  value of 1.0. Using Eq. 6.1,  $\alpha$  and  $K$  were determined to be  $0.0273 \pm 0.0103$  and  $(5.31 \pm 2.4) \times 10^4$ .

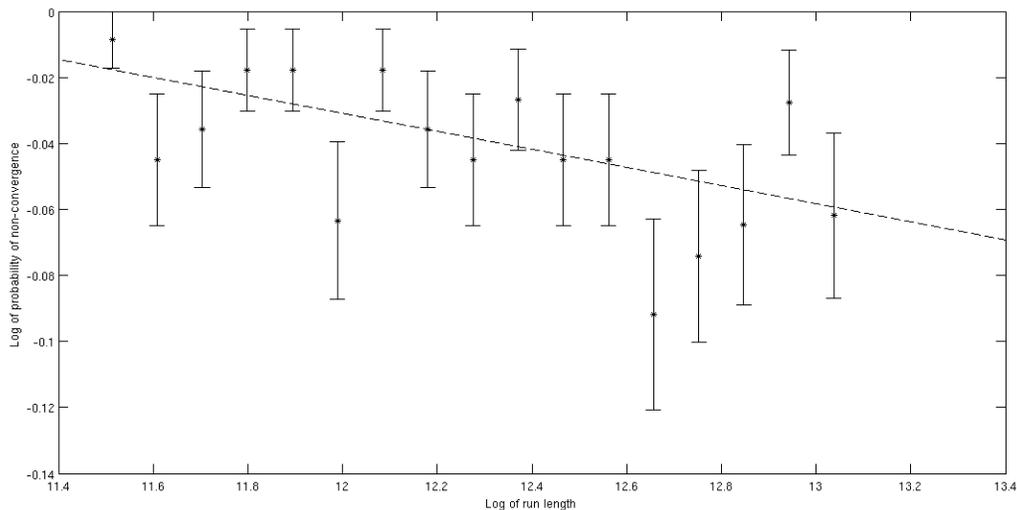


Figure 6.4: Log of probability of non-convergence versus log of run length for  $\alpha = 0.5$ .

The  $\alpha = 0.9$  plot is depicted in Fig. 6.5. The slope and y-intercept of the fit have values of  $0.0047 \pm 0.0060$  and  $-0.070 \pm 0.074$  with a reduced  $\chi^2$  of 0.55.  $K$  and  $\alpha$  were determined to be  $(0.03 \pm 1.2) \times 10^8$  and  $-0.0047 \pm 0.0060$ . Runs of greater lengths would be needed to determine these parameters with any degree of precision. This serves as an indicator to use more aggressive cooling strategies on problems of this nature, as much more frequent convergence was found for the  $\alpha = 0.5$  case.

## 6.2.2 Sequential Performance

Given the  $\alpha$  and  $K$  parameters found from the  $\alpha = 0.5$  case, the number of steps that are required for a single run to achieve 50% probability of success

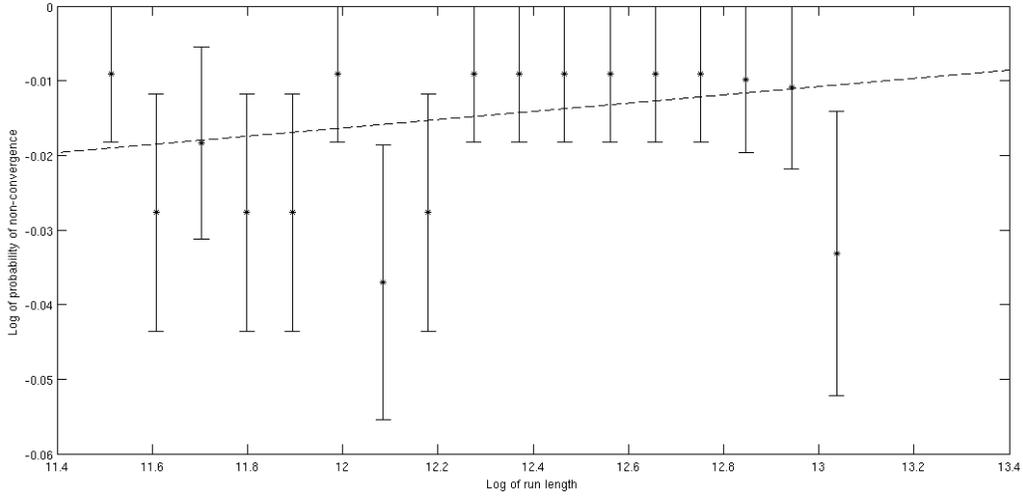


Figure 6.5: Log of probability of non-convergence versus log of run length for  $\alpha = 0.9$

is on the order of  $10^{19}$ , according to Eq. 6.1. It would take a single processor on the order of 500 million years to take this number of steps! The sheer size of this number demonstrates the power of the MIR strategy, when compared to sequential runs.

### 6.2.3 Optimizing Run Length

The run length that maximizes the probability of convergence can be determined using the method of Lagrange multipliers. The total number of steps  $T$  is given by

$$T = Rn \tag{6.2}$$

where  $R$  is the number of runs and  $n$  is the run length. The probability of obtaining at least one successful run out of  $R$  runs is given by

$$P_s = 1 - (P_f)^R = 1 - \left(\frac{K}{n}\right)^{R\alpha} \quad (6.3)$$

where  $P_f$  is the probability of non-convergence on any given run. Optimizing  $P_s$  under the constraint of fixed  $T$  is equivalent to finding an extremum of the function  $L(n, R, \lambda)$

$$L(n, R, \lambda) = 1 - \left(\frac{K}{n}\right)^{R\alpha} + \lambda(T - Rn) \quad (6.4)$$

where  $\lambda$  is a Lagrange multiplier. Setting the partial derivatives of  $P_s$  with respect to  $R$  and  $n$  equal to zero,

$$\frac{\partial L}{\partial R} = -\alpha \ln\left(\frac{K}{n}\right) \left(\frac{K}{n}\right)^{R\alpha} - \lambda n = 0 \quad (6.5)$$

$$\frac{\partial L}{\partial n} = R\alpha K^{R\alpha} n^{-R\alpha-1} - \lambda R = 0 \quad (6.6)$$

and solving for  $n$

$$\ln\left(\frac{K}{n}\right) = -1 \quad (6.7)$$

$$n = Ke \quad (6.8)$$

the optimal number of steps can be found to be  $(1.44 \pm 0.65) \times 10^5$ .

### 6.3 Conclusions

In this thesis, the problem of determining diamond surface deformities using laser interferometry is addressed. A Michelson interferometer comprised of

a helium-neon laser, a reference mirror at the end of one leg and a diamond wafer at the end of the other produced interferograms containing information about the diamond surface. The interferograms result from three-wave interference between the reference mirror and both sides of the diamond. These types of interferograms cannot be solved using traditional methods. A novel method for their analysis was presented in the form of a simulated annealing algorithm. This problem is a challenging one for numerical search methods because of the very large search space and the large number of local minima in the cost function. Additionally, Monte Carlo techniques suffer from very long run times on these large problems. This obstacle was overcome by utilizing a parallelized multiple-run strategy which distributed the job amongst a mid-sized cluster of 24 processors. A  $50px \times 50px$  test interferogram was created, and visually indistinguishable solutions were found repeatedly. Convergence for different cooling parameters was tested. The superiority of using many shorter runs over a single run of greater length was shown by demonstrating that single runs would take on the order of 500 million processor-years to achieve a 50% probability of convergence. The optimal run length for the multiple-runs strategy was found to be on the order of  $10^5$ , which happened to coincide with the value used in runs on the test problem.

Future work should be devoted to testing the effects of increasing the size of the test problem from  $50px \times 50px$  until its size is comparable to the actual diamond three-wave interferograms ( $300px \times 300px$ ). Low resolution solutions (i.e. ones found with lower maximal orders of Legendre polynomials) should be used to find an initial configuration for runs with higher resolu-

tion. Temperature scheduling may have to be modified during these runs, as the warming-up algorithm in the existing program is designed to ensure that the final configuration does not depend on the starting configuration. This test would investigate the claim made by the authors of Ref. [16] that coarse features are determined at higher temperature values whereas fine features are exposed at lower values.

In this paper, simulated annealing was shown to be an effective method for diamond wafer interferogram analysis. Before this work was done, the only means for obtaining surface information from the diamonds was through x-ray diffraction measurements, which required much planning, time and human resources. However, with the effectiveness of simulated annealing demonstrated, surface information can now be obtained by simply analyzing interferograms with a computer cluster. This flexibility will help to expedite the construction of the beam line at Jefferson Lab's Hall D. Once Hall D is complete, GlueX can take place and new physics can be explored!

## **Acknowledgements**

I would like to thank Dr. Richard Jones for giving me the ability to do this project and his infinite patience over the course of the last two years. Dr. George Gibson, thank you for your help in obtaining the diamond interferograms, without them this project wouldn't be possible. To Igor Senderovich and Carl Nettleton, thank you for your help with everything from charming Matlab into doing what I need it to do to helping me keep focused on the work at hand. To Mary Nolan, thank you for putting up with me for the last three years - - I'm sure my sanity would be completely gone now if it weren't for you. To my family, thank you for the opportunity to come to school here at UConn and the unrelenting support over my 21 year existence. Last but not least, to all my friends, thanks for being there!

# Bibliography

- [1] D. J. Schneider. An electromagnetic wave.  
<http://www.geo.mtu.edu/rs/back/spectrum/>.
- [2] W. H. Steel. *Interferometry*. Cambridge U. Press, 1983.
- [3] C. Candler. *Modern Interferometers*. Hilger and Watts Ltd., 1951.
- [4] Falcorian. Michelson interferometer green laser interference. [http : //commons.wikimedia.org/wiki/Image : Michelson\\_Interferometer\\_Green\\_Laser\\_Interference.jpg](http://commons.wikimedia.org/wiki/Image:Michelson_Interferometer_Green_Laser_Interference.jpg).
- [5] L. Motta. Michelson interferometer. Wolfram Science World.
- [6] K. Larkin. A self-calibrating phase-shifting algorithm based on the natural demodulation of two-dimensional fringe patterns. *Optics Express*, 9(5):236–253, 2001.
- [7] H. Ina M. Takeda and S. Kobayashi. Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. *Journal of the Optical Society of America*, 72(1):156–160, 1982.

- [8] J. Sossa-Azuela F. Cuevas and M. Servin. A parametric method applied to phase recovery from a fringe pattern based on a genetic algorithm. *Optics Communications*, 203(3–6):213–223, 2002.
- [9] Z. Ge, F. Kobayashi, S. Matsuda, and M. Takeda. Coordinate-transform technique for closed-fringe analysis by the fourier-transform method. *Applied Optics*, 40(10):1649–1657, 2001.
- [10] A. Smith. Mexican hat function. <http://www.vector.org.uk/archive/v194/finn194.htm>.
- [11] M. Servin, J. L. Marroquin, and F. J. Cuevas. Demodulation of a single interferogram by use of a two-dimensional regularized phase-tracking technique. *Applied Optics*, 36(19):4540–4548, 1997.
- [12] M. Servin, J. L. Marroquin, and F. J. Cuevas. Fringe-follower regularized phase tracker for demodulation of closed-fringe interferograms. *Journal of the Optical Society of America A*, 18(3):689–695, 2001.
- [13] M. J. Lalor D. J. Tipper, D. R. Burton. A neural network approach to the phase unwrapping problem in fringe analysis. *Nondestructive Testing and Evaluation*, 12(6):391–400, 1996.
- [14] E. Onbaşoğlu and L. Özdamar. Parallel simulated annealing algorithms in global optimization. *Journal of Global Optimization*, 19(1):27–50, 2001.
- [15] J. Pintér. Global optimization. Wolfram Math World.

- [16] C. D. Gelatt S. Kirkpatrick and M. P. Vecchi. Optimization by simulated annealing. *Science, Number 4598, 13 May 1983*, 220, 4598:671–680, 1983.
- [17] G. Kliewer and K. Klohs. Parallel simulated annealing library (parsa) user manual. <http://wwwcs.uni-paderborn.de/fachbereich/AG/monien/SOFTWARE/PARSA/>.
- [18] W. D. Gropp and E. Lusk. The message passing interface (mpi) standard. <http://www-unix.mcs.anl.gov/mpi/>.

# Appendix A

## Theory

### A.1 Maxwell's Equations

#### A.1.1 In Vacuum

$$\begin{aligned}\text{Gauss' Law :} & \quad \nabla \cdot \mathbf{E} = 0 \\ \text{Gauss' Law for Magnetism :} & \quad \nabla \cdot \mathbf{B} = 0 \\ \text{Faraday's Law :} & \quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \text{Ampere's Law :} & \quad \nabla \times \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = 0\end{aligned}\tag{A.1}$$

#### A.1.2 In the Presence of External Media

$$\begin{aligned}\text{Gauss' Law :} & \quad \nabla \cdot \mathbf{D} = \rho_{free} \\ \text{Gauss' Law for Magnetism :} & \quad \nabla \cdot \mathbf{B} = 0 \\ \text{Faraday's Law :} & \quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \text{Ampere's Law :} & \quad \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}_{free}\end{aligned}\tag{A.2}$$

Note that for linear media:

$$\begin{aligned} \mathbf{D} &= \epsilon \mathbf{E} \\ \mathbf{H} &= \frac{1}{\mu} \mathbf{B} \end{aligned} \tag{A.3}$$

### A.1.3 In Conducting Media

$$\begin{aligned} \text{Ohm's Law :} & \quad \mathbf{J}_{\text{free}} = \sigma \mathbf{E} \\ \text{Continuity Equation :} & \quad \nabla \cdot \mathbf{J}_{\text{free}} = -\frac{\partial}{\partial t} \rho_{\text{free}} \\ \text{Gauss' Law :} & \quad \nabla \cdot \mathbf{E} = 0 \\ \text{Gauss' Law for Magnetism :} & \quad \nabla \cdot \mathbf{B} = 0 \\ \text{Faraday's Law :} & \quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \text{Ampere's Law :} & \quad \nabla \times \mathbf{B} = \mu \epsilon \frac{\partial \mathbf{E}}{\partial t} + \mu \sigma \mathbf{E} \end{aligned} \tag{A.4}$$

Note:  $\sigma$  is electrical conductivity.

## A.2 Wave Equations

### A.2.1 In Vacuum and Linear Media

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{A.5}$$

$$\nabla^2 \mathbf{B} = \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} \tag{A.6}$$

### A.2.2 In Conducting Media

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu \sigma \frac{\partial \mathbf{E}}{\partial t} \tag{A.7}$$

$$\nabla^2 \mathbf{B} = \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} + \mu \sigma \frac{\partial \mathbf{B}}{\partial t} \tag{A.8}$$

### A.3 Wave Equation Solutions

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)} \quad (\text{A.9})$$

$$\mathbf{B} = \frac{1}{v} \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)} \quad (\text{A.10})$$

Note:  $v = \frac{c}{n}$  and  $n = \sqrt{\epsilon\mu}$ .