NONLINEAR ELECTROMAGNETIC RADIATION FROM METAL-INSULATOR-METAL TUNNEL JUNCTIONS

Thesis

Submitted to

The School of Engineering of the

UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree of

Master of Science in Electro-Optics

By

Mallik Mohd Raihan Hussain

UNIVERSITY OF DAYTON

Dayton, Ohio

May, 2017

NONLINEAR ELECTROMAGNETIC RADIATION FROM

METAL-INSULATOR-METAL TUNNEL JUNCTIONS

Name: Hussain, Mallik Mohd Raihan

APPROVED BY:

Joseph W. Haus, Ph.D.
Advisor Committee Chairman
Professor, Department of Electro-Optics

Andrew Sarangan, Ph.D.
Committee Member
Professor, Department of Electro-Optics

Imad Agha, Ph.D.
Committee Member
Assistant Professor, Department of
Electro-Optics

Robert J. Wilkens, Ph.D., P.E.
Associate Dean for Research and
Innovation
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P. E.
Dean, School of Engineering

ABSTRACT

NONLINEAR ELECTROMAGNETIC RADIATION FROM METAL-INSULATOR-METAL TUNNEL JUNCTIONS

Name: Hussain, Mallik Mohd Raihan
University of Dayton

Advisor: Dr. Joseph W. Haus

Our goal was to experimentally detect nonlinear electromagnetic (EM) radiation (in the far field) from a metal-insulator-metal (MIM) tunnel junction where the insulator thickness lies in the nanometer to subnanometer range and the metals in the junction are coupled to the electro-magnetic field of incident photons. The radiation from an MIM junction originated from the photon-induced tunneling current passing through it. The phenomenon is elegantly described by photon-assisted-tunneling (PAT) theory that introduces transfer Hamiltonians in the uncoupled (when two metals are at infinite distance from each other) system Hamiltonian. This theory predicts the contribution of additional conductivity terms in the MIM interface (due to tunneling inside the junction) and ushered the development of quantum conductivity theory (QCT), as a consequence. In this thesis, we reviewed QCT from the perspective of many-body formulation and designed careful experiments to detect the nonlinear electromagnetic radiation from MIM junctions that can be attributed to photon assisted tunneling of electrons. In our experiment, first, an insulator layer was put on the metal surface using atomic layer deposition (ALD) technique. The number of layers were varied to produce MI samples with different insulator thickness in the subnanometer range. Then, we set the

background signal strength by measuring the second harmonic (SH) and third harmonic (TH) signal due to the bulk material and the surface of metal-insulator (MI) interface. Next, we spin-coated the MI sample with Au nanospheres (diameter $\approx$ 10 nm) to construct MIM interfaces and measured SH and TH signals from them again. Without any bias voltage across the MIM, QCT predicts an increase in TH signal only. Experimentally, we observed an increase in TH signal strength. The increase was modest which is partially attributed to the fact that we could not reliably produce MIM samples with subnanometer insulator thickness and uniform coverage. We intend to improve the surface coverage and uniformity of the insulator layer, in future, and measure SH and TH from the improved samples. Detection of such radiation would support QCT and validate the extension of transfer Hamiltonian approach from the realm of superconducting tunnel junctions to normal MIM tunnel junctions.

To my loving wife, Ishti

ACKNOWLEDGMENTS

Research of any sort comes with the inherent risk of failures. My research on metal-insulator-metal tunnel junctions was no exception. I would like thank Dr. Joseph Haus for believing in me during those initial hard times. His kindness has always inspired me. The long discussions and debates that I had with him on different aspects of the theory of MIM have pushed me to dive deeper into the subject. I can never thank him enough for giving me the space to put forward my ideas and how I see the problem. I believe, I learned a lot in the process of it. I feel blessed to have Dr. Imad Agha as one of my mentors. His suggestions during the setup and result analysis stages of the experiment meant a lot to me. I wish to express my gratitude to Dr. Andrew Sarangan for helping me with spin-coating and plasma treatment. The simulation work of Dr. Domenico de Ceglia and Dr. Michael Scalora helped me analyze my results. I am grateful to our collaborators at Washington University at St. Louis i.e. Dr. Parag Banerjee and his graduate student, Zhengning Gao for performing ALD and AFM. I would like to thank Dr. Cong Deng for sharing his expertise on optical design and ZEMAX. He showed me how to simplify a setup and how to gradually introduce complexity in it. I would like to thank Dr. Andy Chong and Dr. Partha Banerjee for letting me borrow (from their labs) different components required for my setup. I also gratefully acknowledge the support from Army Research Office (Grant number: W911NF-15-1-0178). Lastly, I would like to thank my family for their continuous encouragement and support. They are the ones who give meaning to my whole life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

The introduction of the 'transfer Hamiltonian' approach in the theory of superconducting tunnel junctions [1, 2, 3] was a breakthrough. Since it was developed for single quasi-particle tunneling[†], extension of the theory to the realm of MIM tunnel junction [4] was intriguing. With the development of PAT theory [5, 6], the effect of photons on tunneling electrons became clearer. When a photon is incident on a metal-insulator-metal junction,

1. it acts as a particle by perturbing the system Hamiltonian and, thus, inducing the transition of electrons through the insulator region (quantum tunneling).

2. it acts as a classical field and modulates the Fermi energy level of the system.

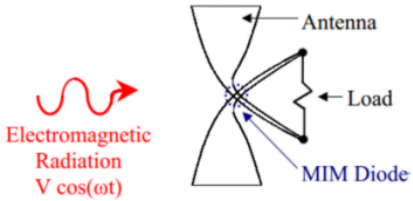


Figure 1.1: Photon field coupling to an antenna and the MIM junction.

[†]In the context of normal metals, electron-electron interactions through phonon (Cooper pairs) are not prominent. So, Josephson tunneling or pair tunneling is negligible and only non-Josephsonian part of the tunneling current (termed as, single quasi-particle tunneling current or normal tunneling current) becomes important. Therefore, the review of the theory is done for single quasi-particle tunneling only and the pertinent radiation is always implied as 'non-Josephsonian'.

If the dimension of the metal sides (see Figure 1.1) of the MIM junction is comparable to the photon wavelength than it can act as antennas and couple to a photon's oscillating electromagnetic field. It is very interesting to note that the field of photons can cause, at the same time, 'harmonic perturbation' and 'harmonic modulation' in the system Hamiltonian of an MIM junction. The terms in the system Hamiltonian that are responsible for photon assisted tunneling are called, 'transfer Hamiltonian's. The constituent of transfer Hamiltonians are as follows,

$$\hat{H}_T = \underbrace{|\hat{H}_T|}_{\substack{\text{magnitude} \\ \text{term}}} \underbrace{\times e^{i\angle_0 \hat{H}_T}}_{\substack{\text{time} \\ \text{independent} \\ \text{phase term}}} \underbrace{\times e^{i\angle_t \hat{H}_T}}_{\substack{\text{time} \\ \text{dependent} \\ \text{phase term}}}, \tag{1.1}$$

- **Magnitude of transfer Hamiltonian,** $|\hat{H}_T|$:

  $|\hat{H}_T|$ consists of terms that are directly related to 'harmonic perturbation' in the MIM junction. In fact, when two systems are coupled through quantum tunneling, the photon-induced harmonic modulation on the input-side can acts as 'harmonic perturbation' on the output-side. Such perturbation causes the transition of quasi-particles from input-side to output-side energy levels (the transition rate can be calculated using Fermi's golden rule, discussed in Appendix A) and generate a DC current. So, it is reasonable to infer that the magnitude of transfer Hamiltonian, $|\hat{H}_T|$, accounts for the particle effect of photon.

- **Phase of transfer Hamiltonian,** $\angle\hat{H}_T$:

  The phase of transfer Hamiltonian, $\angle\hat{H}_T = \angle_0\hat{H}_T + \angle_t\hat{H}_T$ accounts for the classical electromagnetic field aspect of photons. This will be further discussed later in Chapter III. In the absence of tunneling (or any coupling between the junctions), any voltage, AC or DC, including the voltage due to the electromagnetic field of photons, simply causes a phase modulation on the (Fermionic) creation and annihilation operator. In a coupled system, phases of annihilation and creation operators can form the phase part of the transfer Hamiltonian. The phase part is responsible for the nonlinear radiations coming out of the MIM junction [5, 7].

Microscopically, the effect of photons on MIM tunnel junction is analogous to photo-detection where a DC bias voltage across the MIM junction plays the same role as band gap in semiconductor photo-detectors. The result is a DC current through the junction. Meanwhile, the phase part of the transfer Hamiltonian causes back and forth 'sloshing' of electrons through the junction, generating AC currents at the input photon frequency and its harmonics. These AC current terms can be related back to the DC current terms for the corresponding DC photon voltages of higher frequencies. Consequently, the macroscopic effects of photons on tunnel junctions are two fold. Their oscillating electric field can (1) generate a DC current and, (2) AC currents of the same and higher-order harmonic frequencies. These AC currents, in turn, may couple to electromagnetic waves and radiate from the tunnel junction. 'Quantum conductivity theory' builds on that notion and couples the higher frequency current terms with the local electric field (due to the incident photon) through few constitutive relations. These constitutive relations give us a set of linear and nonlinear quantum conductivity coefficients which can be used to calculate the higher harmonic current terms easily. The goal of this thesis was to experimentally detect the harmonic electromagnetic radiation from the tunnel junction due to the corresponding current terms.

To make this thesis as self-consistent as possible, the theory is developed from first principles and gradually brought to a point where additional nonlinear conductivity terms can be introduced that explains the coupling of nonlinear radiation with tunneling current. To understand QCT and its implication, one needs to contemplate the following topics step-by-step:

Step 1 : Tunneling probability current density operator in quantum mechanics,

Step 2 : Time dependent perturbation and Fermi's second golden rule,

Step 3 : Bardeen's picture of microscopic tunneling current and introduction of transfer Hamiltonian formulation,

Step 4 : Generalized PAT theory,

Step 5 : Finally, QCT theory and predictions.

The step-by-step derivation of theoretical equations predicting radiation from MIM junction is fascinating and gives one confidence on the possibility of experimental detection of such junction-induced radiation. In Figure 1.2, we linked the concepts required to describe QCT and PAT theory from many-body formulation. The linked diagram is partitioned into boxes representing different chapters of the thesis. This figure may serve as a map of all the required concepts for future referral.

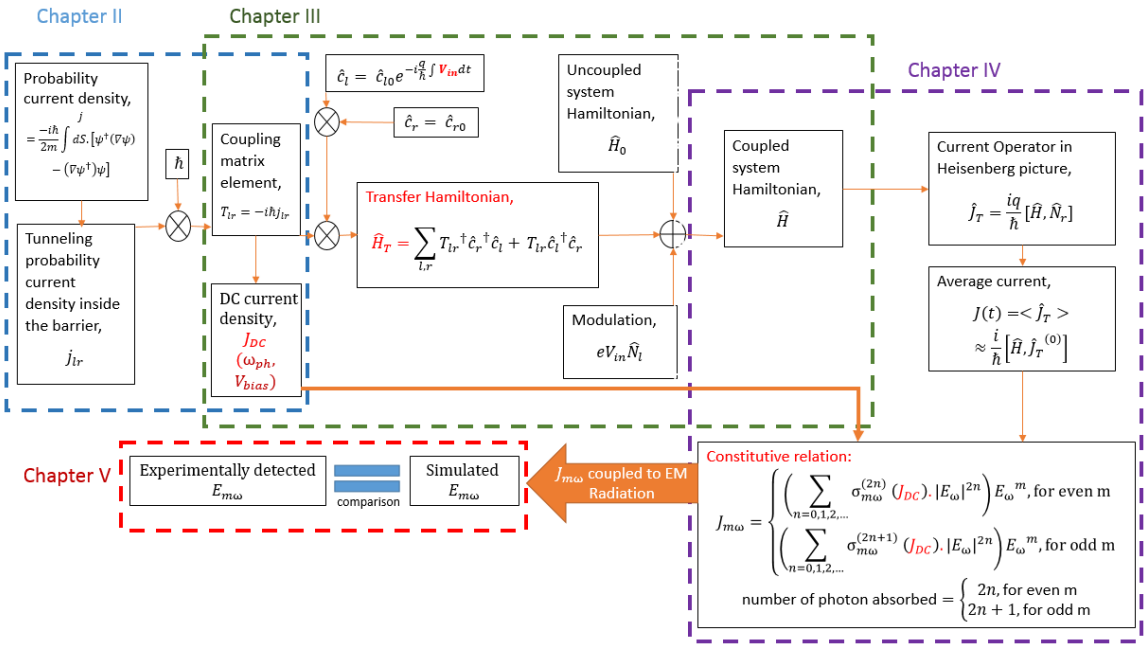


Figure 1.2: Strategic development of concepts of quantum conductivity theory.

In Chapter II, we deduced the tunneling probability current density inside a barrier by utilizing one of the 'matching boundary condition' methods, namely, 'transfer matrix method' (TMM) [8]. Chapter III is a treatment of Bardeen's picture of microscopic tunneling [1, 3]. Our goal in Chapter

III is to introduce additional terms in the uncoupled system Hamiltonian, namely transfer Hamiltonians, that account for tunneling electrons and can be used to measure currents in Heisenberg picture [9]. It turns out that the transfer Hamiltonian operator between any two energy levels of the two ends of the junction is a product of: i) a matrix element that incorporates coupling between those energy levels, ii) an annihilation operator on side, and, iii) a creation operator on the other side. Overall, the transfer Hamiltonian operators annihilate particle on one side and create it on the other side and the process is dependent on the coupling matrix element between the two energy levels. The concept of probability current density inside a barrier, discussed in Chapter II, becomes handy for the calculation of these coupling matrix elements. In fact, the coupling matrix elements are tunneling probability current densities scaled by reduced Planck's constant, $\hbar$ and has units of energy [1]. The transmitted probability current density, calculated through TMM (in Chapter II) and transfer Hamiltonian approach (in Chapter III) are shown to be one and equal [8]. In Chapter IV, we explained the PAT theory by Tien and Gordon [5]. The derivation of total current and quantum conductivity coefficients [10] are discussed later in this chapter. For the experimental part, discussed in Chapter V, we focused intense laser pulses on MIM junction and relayed an image of the radiating junction on a sensitive detector. To guide us through the experiment, we used the quantum conductivities, derived from QCT in Chapter IV, to simulate the radiation pattern. The simulation also helped us to determine the wavelength of incident photons for which the field enhancement is the most and the resonance region of the metal could be avoided. In Chapter VI, we conclude by discussing the implications and potential applications of this theory.

CHAPTER II

PROBABILITY CURRENT DENSITY

The wide-spread use of the word 'current' in different contexts is a source of confusion in the literature of quantum mechanics and one needs to be very careful. Cross-checking the units is always a good way to physically understand the context of the word 'current'. For example, 'current' may represent a particle flux (with or without charge). In some literature, 'current' is the transition rate of a particle from one state to another. Yet, sometimes, 'current' stands for current density normalized over some volume and as the context of volume changes in different number of dimensions, the unit of 'current' changes accordingly. In many-body formulation of quantum mechanics, the 'current' operator is the rate of change of 'number' operator, or, in Heisenberg picture, 'current' operator is the commutation between the system Hamiltonian and the 'number' operator. The macroscopic current density is related to the ensemble average of microscopic many-body 'current' operator. All of these 'current' concepts will become handy as we develop the theory. Having this discussion in mind, we will move forward to deduce the current, that will, eventually, couple to electromagnetic radiation. At first, we will derive equation for microscopic probability current density operator, $j(x)$. The most important take away from this chapter will be the concept that, probability current density from energy level l on one side of the junction to energy level r on the other side, i.e. $j_{lr}(x)$, multiplied by reduced Planck's constant, $\hbar$, gives us a quantity that has dimension of energy and forms the energy matrix element, T_{lr} that couples different energy levels

on both side of the junction.

$$T_{lr} = \hbar j_{lr}(x). \tag{2.1}$$

The physics behind calculating $j_{lr}(x)$, and, in turn, T_{lr} will be discussed in this chapter. In later chapters, we will discuss that how this coupling energy, i.e. T_{lr}, plays an important role in the coupled system Hamiltonian and when multiplied by (creation and annihilation) operators that contribute to the transfer of electrons, they form the 'transfer Hamiltonian' operators.

2.1 Derivation of probability current density, $\mathbf{j}(\mathrm{x})$

Quantum particles, charged or neutral, incorporates a wavefunction, $\psi(x)$ (in 1-D) that holds all the information regarding the particle, such that $\rho(x) = |\psi(x)|^2 = \psi(x)^\dagger\psi(x)$ represents the probability density and the probability of finding the particle in a small volume dV is ρdV. If the wave function is properly normalized, then the probability of finding the particle throughout all space should be 1, i.e. $\int_{-\infty}^{\infty} \rho dV = 1$. Since, probability is a unit-less term that varies from 0 to 1, the dimension of probability density ρ should be $[L]^{-n}$, where $[L]$ is the dimension of length and n is the number of dimension. For 1-D case, the unit of ψ is $m^{-1/2}$ and accordingly, the unit of ρ is m^{-1}.

Operators of any sort (i.e. position, momentum, energy etc) may act on the wavefunction and reproduce $\psi(x)$ as a linear combination of the corresponding operator's eigenfunctions, weighted by the respective eigenvalues. For a non-relativistic particle with charge q_e, mass m, momentum operator $\hat{p}$, position operator $\hat{x}$, the Hamiltonian is,

$$\hat{H} = \frac{\hat{p}^2}{2m} + U(x). \tag{2.2}$$

Here, $U(x)$ is the potential landscape. In space representation, the momentum operator is, $\hat{p} = -i\hbar\nabla$, where ∇ is a differentiation operator. By using Dirac's continuity equation that conserves

probability current, we get

$$\frac{\partial}{\partial t}\rho(x) + \nabla.j(x) = 0. \tag{2.3}$$

where, $j(x)$ is the current density (note: definition of Volume changes according to dimension). Now, differentiation of ρ with respect to time t, is,

$$\frac{\partial}{\partial t}\rho(x) = \frac{\partial}{\partial t}|\psi|^2 = \frac{\partial}{\partial t}\psi^\dagger.\psi = \dot{\psi^\dagger}.\psi + \psi^\dagger.\dot{\psi}. \tag{2.4}$$

The time evolution of a wave function follows the Schrödinger equation i.e.

$$i\hbar\dot{\psi} = \hat{H}\psi. \tag{2.5}$$

By putting this in Equation 2.4, we get,

$$\frac{\partial}{\partial t}\rho(x) = \frac{-1}{i\hbar}[(\hat{H}^\dagger\psi^\dagger)\psi - \psi^\dagger(\hat{H}\psi)]. \tag{2.6}$$

By putting Equation 2.2 and Equation 2.6 in Equation 2.3, we get,

$$\nabla.j(x) = \frac{1}{i\hbar}\frac{(i\hbar)^2}{2m}[(\nabla^2\psi^\dagger)\psi - \psi^\dagger(\nabla\psi)] = \nabla.\left\{\frac{-i\hbar}{2m}[\psi^\dagger(\nabla\psi) - (\nabla\psi^\dagger)\psi]\right\},$$

$$\boxed{\Rightarrow j(x) = \frac{-i\hbar}{2m}[\psi^\dagger(\nabla\psi) - (\nabla\psi^\dagger)\psi]}. \tag{2.7}$$

Equation 2.7 gives us the probability current density which will be used for later development.

2.2 Remarks on probability current density

Some noticeable features of Equation 2.7 will be discussed now.

Feature 1. Classical resemblance:

Notice that, the probability current can be represented in terms of momentum operator $\hat{p}$. This relation closely resembles to its classical counterpart,

$$\vec{j} = \frac{1}{2m}[\psi^\dagger(\hat{p}\psi) + (\hat{p}\psi)^\dagger\psi] = \frac{1}{m}Re\{\psi^\dagger(\hat{p}\psi)\}. \tag{2.8}$$

Classically, current density (more appropriately, particle flux density) in 1-D is, $\vec{j}_{cl} = \rho_{cl}.\vec{v}_{cl} = \rho_{cl}\frac{\vec{p}_{cl}}{m}$. The classical analog of the current density matches its quantum representation for wavefunctions with real wavevectors. In Dirac notation, the complete current density operator in 3D can be written from Equation 2.8 as [11],

$$\hat{\vec{j}}(r) = \frac{1}{2m}\left[|r\rangle\langle r|\hat{\vec{p}} + \hat{\vec{p}}^{\dagger}|r\rangle\langle r|\right], \tag{2.9}$$

where, density of states operator, $\hat{\rho} = |r\rangle\langle r|$.

Feature 2. Boundary matching conditions and probability current in free particle and particle with decaying wavefunction:

The general solution of Equation 2.2 in 1-D has the following form,

$$\psi(x) = ae^{ikx} + be^{-ikx}, \tag{2.10}$$

$$k = \pm\left(\frac{2m}{\hbar^2}(E-U)\right)^{1/2}, \tag{2.11}$$

where, the particle is placed on the potential landscape, $U = eV(x) + U_{int}$ designed by an external potential $eV(x)$ and any internal potential U_{int} due to crystal lattice, and interaction with other particles. When $E - U > 0$, the wavefunction $\psi(x)$ resembles a free propagating wave. On the other hand, when $E - U < 0$ the wavefunction represents an evanescent wave. Let's find $j(x)$ from Equation 2.7 for different kinds of potential landscape:

Case 1 $U(x)$ = Constant Function

When $E - U > 0$, from Equation 2.10,

$$\psi(x) = ae^{ikx} + be^{-ikx}, \qquad \psi(x)^{\dagger} = ae^{-ikx} + be^{+ikx},$$

$$\frac{d}{dx}(\psi(x)) = a(ik)e^{ikx} - b(ik)e^{-ikx}, \qquad \frac{d}{dx}(\psi(x)^{\dagger}) = -a(ik)e^{-ikx} + b(ik)e^{+ikx}.$$

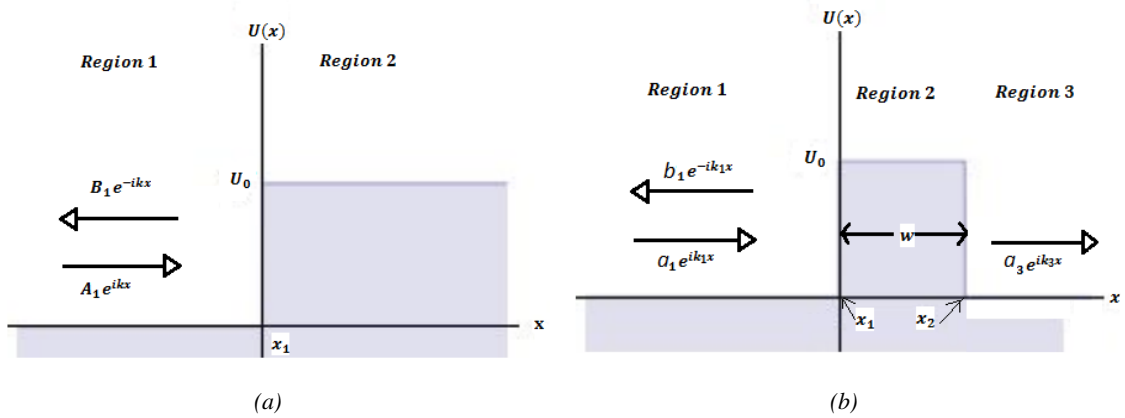


Figure 2.1: Propagating wavefunctions for (a) a step potential and (b) a rectangular potential.

From Equation 2.7

$$\begin{aligned} j(x) = \frac{-i\hbar}{2m}[a^2(ik) - ab(ik)e^{i*2k} + ab(ik)e^{i*2k} - b^2(ik) \\ +a^2(ik) - ab(ik)e^{i*2k} + ab(ik)e^{i*2k} - b^2(ik)], \\ j(x) = \frac{\hbar k}{m}(a^2 - b^2). \end{aligned} \tag{2.12}$$

Notice $v = \frac{p}{m} = \frac{\hbar k}{m}$. So, for a freely propagating particle, the probability current density corresponds to the classical result.

When $E - U < 0$, $\psi = \psi^\dagger$ and so,

$$j(x) = 0. \tag{2.13}$$

Therefore, a decaying particle with constant starting amplitude, (at $x = 0$) i.e. constant a or b carries no current.

Case 2 $U(x)$ = Step Function

Let's consider a step potential function [see Figure 2.1(a)],

$$U(x) = \begin{cases} 0, & \text{for } x < x_1, \text{ Region 1} \\ U_0, & \text{for } x \geq x_1, \text{ Region 2} \end{cases}$$

When $U(x)$ in Equation 2.2 has a discontinuity or jump at some point x_1, boundary conditions need to be matched at that point in space. Careful observation of Equation 2.2 and 2.5 brings out the matching condition. Equation 2.5 equates ψ and $\nabla^2\psi$. If ψ or $\frac{d\psi}{dx}$ has a discontinuity anywhere, $\nabla^2\psi$ ends up having $\delta(x)$ function and Equation 2.5 is not satisfied. So, ψ and $\frac{d\psi}{dx}$ have to be equal at x_1. To find the matching condition, it is convenient to express wavefunction (from Equation 2.10) of region 1 and region 2 [see Figure 2.1(a)]. in two-dimensional vector format, i.e. $\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$ and $\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$. The matching conditions can be described by a 2×2 matrix, R, as follows:

$$R_1(x = x_1) = \frac{1}{2k_1} \begin{pmatrix} (k_1 + k_2).e^{[i(-k_1+k_2)x]} & (k_1 - k_2).e^{[i(-k_1-k_2)x]} \\ (k_1 - k_2).e^{[i(k_1+k_2)x]} & (k_1 + k_2).e^{[i(k_1-k_2)x]} \end{pmatrix}, \tag{2.14}$$

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = R_1 \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}, \tag{2.15}$$

Because Equation 2.12 and 2.13 of Case 1 covers $U(x) > U_0$ and $U(x) < 0$, for Case 2 our region of interest is, $0 < U(x) < U_0$. Let's assume, a free particle in this energy band is incident on the interface from left of region 1 with probability current density $j_{in} = \frac{\hbar k_1}{m}|a_1|^2$. Since no particle is incident from the right of region 2, $b_2 = 0$ in Equation 2.15. From Equation 2.14 and 2.15, we can calculate the $\psi, \psi^\dagger, \frac{d\psi}{dx}, \frac{d\psi^\dagger}{dx}$, in region 2 after tunneling as,

$$\psi_2 = \frac{2k_1(k_1 - i\kappa_2)}{(k_1)^2 + (\kappa_2)^2} e^{ik_1x_1} e^{-\kappa_2(x-x_1)}, \qquad {\psi_2}^\dagger = \frac{2k_1(k_1 + i\kappa_2)}{(k_1)^2 + (\kappa_2)^2} e^{-ik_1x_1} e^{-\kappa_2(x-x_1)}, \tag{2.16}$$

$$\frac{d\psi_2}{dx} = \frac{-2k_1\kappa_2(k_1 - i\kappa_2)}{(k_1)^2 + (\kappa_2)^2} e^{ik_1x_1} e^{-\kappa_2(x-x_1)}, \quad \frac{d{\psi_2}^\dagger}{dx} = \frac{-2k_1\kappa_2(k_1 + i\kappa_2)}{(k_1)^2 + (\kappa_2)^2} e^{-ik_1x_1} e^{-\kappa_2(x-x_1)}. \tag{2.17}$$

By putting the above values in Equation 2.7, we calculate the tunneling probability current density as, $j_t = 0$. So, for a step function i.e. one potential discontinuity, the current never tunnels. This changes for a rectangular potential, as we will see next.

Case 3 $U(x)$ = Rectangular Function

Now, let's consider a rectangular potential function of w width [see Figure 2.1(b)],

$$U(x) = \begin{cases} 0, & \text{for } x < x_1, \text{ Region 1} \\ U_0, & \text{for } x_2 \geq x \geq x_1, \text{ Region 2} \\ U(< U_0), & \text{for } x > x_2, \text{ Region 3} \end{cases}$$

For such potential barrier problems, Equation 2.15 can be written as,

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = R_1 R_2 \begin{pmatrix} a_3 \\ b_3 \end{pmatrix}, \tag{2.18}$$

where, R_1 is the transfer matrix at left interface $(x = x_1)$ [see Equation 2.14], R_2 is, similarly, the transfer matrix at right interface $(x = x_2)$. The energy band which offers something new for Case 3 is, $U < U(x) < U_0$. A particle, with energy that lies in this band and with probability current density, $j_{in} = \frac{\hbar k_1}{m}|a_1|^2$, is incident on the left edge of the barrier and it transmits from the right edge with probability current density $j_{tr} = \frac{\hbar k_3}{m}|a_3|^2$. Since no particle is incident on the right edge of the barrier, $b_3 = 0$. By following the same trick we used for Case 2, we write,

$$\begin{aligned} a_1 &= (R_1 R_2)_{11} a_3, \\ \Rightarrow a_3 &= \frac{4k_1\kappa_2 e^{-\kappa_2 w}}{({k_1}^2 + {\kappa_2}^2)^{1/2}({k_3}^2 + {\kappa_2}^2)^{1/2}} \phi a_1, \end{aligned} \tag{2.19}$$

where, $\phi = e^{[i(\frac{\pi}{2} - \alpha) + (k_1 x_1 - k_3 x_3)]}$, $\alpha = tan^{-1}(\kappa_2/k_1) + tan^{-1}(\kappa_2/k_3)$ and $\kappa_2 = ik_2 = \frac{2m(U_0 - E)}{\hbar^2}^{1/2}$. At this point, we can figure out the tunneling probability, W,

$$\boxed{W = \frac{j_{tr}}{j_{in}} = \frac{16k_1 k_3 {\kappa_2}^2}{({k_1}^2 + {\kappa_2}^2)({k_3}^2 + {\kappa_2}^2)} e^{-2\kappa_2 w}.} \tag{2.20}$$

Notice that the tunneling probability is symmetric with respect to region 1 and 3. We can also determine the tunneling probability current density, $j(x)$ inside the barrier region. At any point inside the barrier i.e. $x_1 < x < x_2$, the wavefunction has the form,

$$\psi_2(x) = a_2 e^{-\kappa_2 x} + b_2 e^{+\kappa_2 x}. \tag{2.21}$$

By putting ψ in Equation 2.7, we find,

$$\boxed{j(x) = \frac{i\hbar\kappa}{m}(a_2 b_2{}^{\dagger} - b_2 a_2{}^{\dagger}).} \tag{2.22}$$

Notice, the difference between Case 2 (step potential) and Case 3 (rectangular potential) is that, $b_2 \neq 0$ for Case 3. So, to have a sustaining tunneling probability current density, one must introduce at least one barrier like potential with two interfaces.

Case 4 $U(x)$ = Arbitrary Function

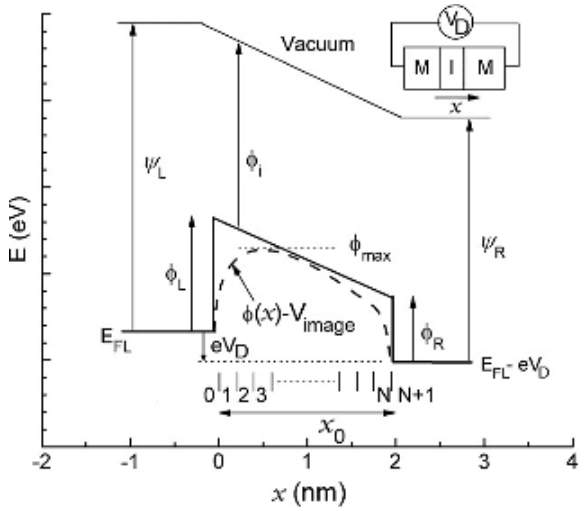


Figure 2.2: Transfer matrix method for arbitrary potential barrier [12].

The method we used so far is called transfer matrix method (TMM). We can use it to determine the tunneling probability current density through barrier of any arbitrary shape by representing the barrier as a collection of very thin rectangular barriers (see Figure 2.2). For arbitrary potentials, we can also use Wentzel, Kramers, and Brillouin (WKB) approximation. This is a simple result covered in many texts and papers. For example, tunneling through

parabolic potential is extensively developed by Simmons [13, 14] using WKB approximation. Calculation of tunneling probability current density through different cases of potential barrier using WKB is also discussed in Ghatak's book [15]. Shooting method is another technique to calculate the tunneling probability current density [16]. Since the boundary conditions at the potential barrier are utilized to develop these methods, they can be collectively called 'matching boundary condition' method.

2.3 Summary

Tunneling probability current density only exists inside a barrier potential with at least two boundaries. Inside the barrier (i.e. when, $E - U < 0$), the particle must have an evanescent wavefunction both in forward and backward direction. If either of a or b is 0, the probability current density drops to zero. This sort of current density that exists inside a barrier was used by Bardeen to couple two metal systems through a barrier [1, 3].

CHAPTER III

BARDEEN'S PICTURE OF MICROSCOPIC TUNNELING CURRENT

A good grasp over Fermi's golden rule (discussed in Appendix A) is required to understand Bardeen's picture of tunneling. In fact, the picture relates the tunneling probability current density[‡], j_{lr} (or j_{rl}), inside the barrier, to the transition rate, Γ_{lr} (or Γ_{rl}), calculated from Fermi's (second) golden rule. The dependency is shown in Equations 3.1-3.3 (subscripts 'r' and 'l' can be exchanged to produce Equations describing tunneling in the opposite direction),

$$\boxed{\Gamma_{lr}} = \frac{2\pi}{\hbar} \times |\underbrace{\boxed{T_{lr}}}_{\downarrow}|^2 \times \rho_r(E_r), \tag{3.1}$$

$$T_{lr} = -i\hbar \times \underbrace{\boxed{j_{lr}}}_{\downarrow}, \tag{3.2}$$

$$j_{lr} = \frac{-i\hbar}{2m} \int dS.(\psi_l{}^{\dagger} \nabla \psi_r - \psi_r \nabla \psi_l{}^{\dagger}). \tag{3.3}$$

Equation 3.1 is discussed in Appendix A as Fermi's golden rule, whereas, Equation 3.3 is discussed in Chapter II as tunneling probability current density inside a barrier. In Section 3.1, we will describe how Bardeen's approach to the tunneling problem gives us Equation 3.2 (energy coupling matrix element, T_{lr}) that links Equation 3.1 and 3.3. In Section 3.2, we will show that Equation 3.2 is equivalent to the simple 'matching boundary condition' approach [discussed in Section 2.2] taken by Simmons [13, 14] to solve the same tunneling problem.

[‡]Subscript definition: 'r' = right, '1' = left, 'lr' = right to left, 'rl' = left to right. The direction of electron transfer from right to left (subscript 'lr') is positive. These are the convention followed in the literature and in this thesis.

The merits of Bardeen's approach are:

1. many-body effects can be easily included,

2. the technique naturally lends itself to time-dependent phenomena.

The success of Bardeen's approach in explaining the results of Giaever's experiment on superconductor tunnel junction [17] makes it attractive. Josephson extended this approach for pair tunneling [18], which brilliantly explained the AC and DC Josephson effect in Shapiro's experiment [19]. It is intriguing to describe tunneling in MIM junctions using Bardeen's approach and, thus, extend its use in regular metals. In Section 3.3, we will derive the total probability current density, J_{DC} from transition rate, Γ_{lr}.

3.1 Introduction of transfer Hamiltonian, $\hat{\mathbf{H}}_{\mathbf{T}}$

The objective of Bardeen's approach is to introduce a transfer Hamiltonian $\hat{H}_T$ to the system that accounts for tunneling.

$$\hat{H} = \hat{H}_{l0} + \hat{H}_{r0} + \hat{H}_T, \tag{3.4}$$

where, $\hat{H}_0 = \hat{H}_{l0} + \hat{H}_{r0}$ is the uncoupled Hamiltonian of the system, when the metals are at infinite distance from each other and no electron can tunnel from one to another. So, they are two separate systems $\hat{H}_{l0}$ and $\hat{H}_{r0}$ and their Eigenfunctions are $\psi_l(x)$ and $\psi_r(x)$ and Eigenvalues are E_l and E_r respectively. $\hat{H}_T$ is the unknown addition to the Hamiltonian that needs to be calculated. The Eigenfunctions $\psi_l(x)$ and $\psi_r(x)$ of uncoupled, unperturbed system will be used as trial function to, eventually, calculate $\hat{H}_T$ of the coupled system. Notice, the usual implementation of time-dependent perturbation technique involves the introduction of small changes in the Hamiltonian and, then, determination of the new exact wavefunction. But Bardeen took a different approach. He introduced an approximate trial wavefunction and solved for the exact transfer Hamiltonian, $\hat{H}_T$.

This way, the calculated transfer Hamiltonian does not have to be "assumed" as small, it will always be small. To couple the systems through tunneling, metals are now brought very close to each other.

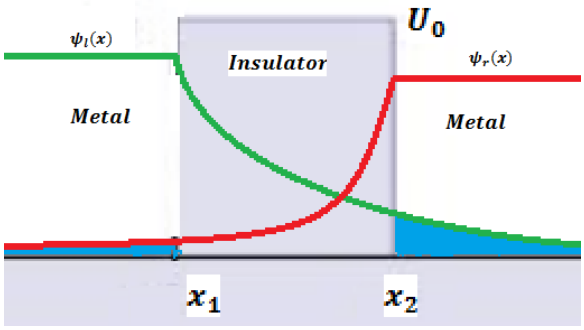


Figure 3.1: Bardeen's picture of tunneling in MIM junction.

Now, as they get closer (see Figure 3.1), the decaying tail of electron wavefunction in the barrier region starts to overlap with the other metal. But 'overlap' does not mean that electron wavefunction is still decaying after it reached the metal system on the other side. Rather, decaying tail of the wavefunction of one metal system acts as a source of perturbation in the other system. Now, let's assume the decaying part of the wavefunction in the uncoupled system are:

$$\psi_l(x) = a_{2l}e^{-\kappa_2 x}; \qquad \text{for } x \geq x_1, \tag{3.5}$$

$$\psi_r(x) = b_{2r}e^{\kappa_2 x}; \qquad \text{for } x \leq x_2, \tag{3.6}$$

where, κ_2 is the attenuation constant inside the barrier. The wavefunction envelops of the uncoupled system are shown in Figure 3.1 with $\psi_l(x)$ in green and $\psi_r(x)$ in red. When the coupling is weak, the uncoupled solution, $\psi_l(x)$ is still a good solution of the new Hamiltonian, $\hat{H}$ for $x < x_2$. For $x > x_2$, $\psi_l(x)$ is not a correct solution of $\hat{H}$, but it is taken as an ansatz for this region. Similarly, $\psi_r(x)$ is a good solution for $x > x_1$, but for $x < x_1$ it is incorrect but an approximate solution.

In Figure 3.1, the part of the curves that have the blue shaded regions underneath it, represent the part of the corresponding wavefunction that is not a solution of the new coupled system. Since, inside the barrier both $\psi_l(x)$ and $\psi_r(x)$ are good solution, we can use Equation 2.7 to calculate the probability current density at any point inside the barrier (i.e. $x_1 < x_B < x_2$).

$$\begin{aligned} j(x_B) &= -\frac{i\hbar}{2m}\left[(\psi_l+\psi_r)^\dagger(\frac{d}{dx})(\psi_l+\psi_r) - (\psi_l+\psi_r)(\frac{d}{dx})(\psi_l+\psi_r)^\dagger\right], \\ &= \underbrace{-\frac{i\hbar}{2m}\left(\psi_r^\dagger\frac{d\psi_l}{dx} - \psi_l\frac{d\psi_r^\dagger}{dx}\right)}_{=j_{rl}} \underbrace{+\frac{i\hbar}{2m}\left(\psi_l^\dagger\frac{d\psi_r}{dx} - \psi_r\frac{d\psi_l^\dagger}{dx}\right)}_{=j_{lr}}. \end{aligned} \tag{3.7}$$

Let's assume, initially, that the electron is in left-side metal and we want to calculate the effective transfer Hamiltonian. The new mixed state can be written as,

$$\psi(t) = c(t)\psi_l e^{-iE_l t/\hbar} + d(t)\psi_r e^{-iE_r t/\hbar}. \tag{3.8}$$

By substituting $\psi(t)$ in Schödinger's equation i.e. $\hat{H}\psi = i\hbar\frac{\partial\psi}{\partial t}$, we get,

$$\begin{aligned} i\hbar\dot{c}\psi_l e^{-iE_l t/\hbar} + c\psi_l E_l e^{-iE_l t/\hbar} + i\hbar\dot{d}\psi_r e^{-iE_r t/\hbar} + d\psi_r E_r e^{-iE_r t/\hbar} \\ = ce^{-iE_l t/\hbar}\hat{H}\psi_l + de^{-iE_r t/\hbar}\hat{H}\psi_r. \end{aligned} \tag{3.9}$$

Since the electron is assume to be sitting at left side wavefunction initially, $c \approx 1, d \approx 0$. Now, the particle must be found either in left or right side metal. Therefore, if Equation 3.8 is normalized, then,

$$cc^\dagger + dd^\dagger = 1. \tag{3.10}$$

By differentiating Equation 3.10 with respect to time, t,

$$\begin{aligned} \frac{d}{dt}(cc^\dagger + dd^\dagger) = 0, \\ \Rightarrow \dot{c} = 0. \end{aligned} \tag{3.11}$$

By putting the values of $c, d, \dot{c}$ in Equation 3.9, we get,

$$i\dot{d}\psi_r e^{-iE_r t} = \frac{(\hat{H} - E_l)}{\hbar}\psi_l e^{-iE_l t}. \tag{3.12}$$

By multiplying both sides by ${\psi_r}^\dagger$ and then integrating over all volume, we get,

$$\dot{d} = \frac{1}{i\hbar}\int_{-\infty}^{\infty} {\psi_r}^\dagger(\hat{H} - E_l)\psi_l e^{-i(E_l - E_r)t}dx. \tag{3.13}$$

Notice, we used a time-dependent perturbation technique which is similar to the technique used to derive Fermi's golden rule (discussed in Appendix A) and Equation 3.13 looks very similar to Equation A.6. Therefore, by using same deduction, we can write,

$$T_{rl} = \int_{-\infty}^{\infty} {\psi_r}^\dagger(\hat{H} - E_l)\psi_l dx. \tag{3.14}$$

Let's focus on the limits of the integration of Equation 3.14. For $x < x_2$ the integrand in zero, because, as already discussed above, $\psi_l(x < x_2)$ is a solution of new system Hamiltonian $\hat{H}$. So, for any point inside the barrier x_B, i.e. $x_1 < x_B < x_2$.

$$T_{rl} = \int_{x_B}^{\infty} {\psi_r}^\dagger(\hat{H} - E_l)\psi_l dx. \tag{3.15}$$

Similarly, it can be shown,

$$T_{lr} = \int_{-\infty}^{x_B} {\psi_l}^\dagger(\hat{H} - E_r)\psi_r dx. \tag{3.16}$$

Equation 3.15 can be further developed by adding the term $-\psi_l(\hat{H} - E_r){\psi_r}^\dagger dx$ to the integrand. The addition changes nothing because within the integration limit $\psi_r(x)$ is the correct solution of $\hat{H}$ and the value of the added term is 0. We may write,

$$T_{rl} = \int_{x_B}^{\infty} \left[{\psi_r}^\dagger(\hat{H} - E_l)\psi_l - \psi_l(\hat{H} - E_r){\psi_r}^\dagger\right] dx. \tag{3.17}$$

Notice, $T_{rl} = -{T_{lr}}^\dagger$. So, $\boxed{|T_{rl}|^2 = |T_{lr}|^2}$. Integrating Equation 3.17 by parts gives,

$$T_{rl} = -\frac{\hbar^2}{2m}\left({\psi_r}^\dagger\frac{d\psi_l}{dx} - \psi_l\frac{d{\psi_r}^\dagger}{dx}\right). \tag{3.18}$$

Similarly,

$$T_{lr} = -\frac{\hbar^2}{2m}\left({\psi_l}^\dagger\frac{d\psi_r}{dx} - \psi_r\frac{d{\psi_l}^\dagger}{dx}\right). \tag{3.19}$$

From Equation 3.7, 3.18 and 3.19, we can deduce that,

$$\boxed{T_{rl} = -i\hbar j_{rl},} \tag{3.20}$$

and,

$$\boxed{T_{lr} = +i\hbar j_{lr}.} \tag{3.21}$$

Equation 3.20 and 3.21 are very important as well as very interesting results achieved by Bardeen. At the beginning of the chapter (see Equation 3.2), we shortly discussed the implications of these equations. They will be utilized in two ways:

1. Introduction of transfer Hamiltonian, $\hat{H}_T$,

2. Formulation of total probability current density, J_{DC}.

It is evident from Equation 3.12 (similar deductions are also discussed in Equation A.8) that, the energy coupling matrix elements, T_{rl}, carry a phase part with it, i.e. $e^{-i(E_l - E_r)t}$, that represents the external modulation of the Fermi level of the system. Equation 3.15 and 3.16, with their phase part, can be presented using many-body formulation [20, 21], as follows:

$$\hat{H}_T^- = \sum_{l,r} \hat{T}_{rl} \hat{c}_r^\dagger \hat{c}_l, \tag{3.22}$$

$$\hat{H}_T^+ = \sum_{l,r} \hat{T}_{lr} \hat{c}_l^\dagger \hat{c}_r, \tag{3.23}$$

where, $\hat{c}_l^\dagger$ and $\hat{c}_l$ represents the Fermionic annihilation and creation operators on the left side of the barrier and $\hat{c}_r^\dagger$ and $\hat{c}_r$ represents the Fermionic annihilation and creation operators on the right side barrier. $\hat{H}_T^-$ represents the perturbation that causes the electrons to transfer from left to right, while $\hat{H}_T^+$ does the opposite. As we can see, they are the complex conjugate of each other. So, $\hat{H}_T$ can be written as,

$$\boxed{\hat{H}_T = \hat{H}_T^+ + \hat{H}_T^- = \sum_{l,r} \left[\hat{T}_{rl} {c_r}^\dagger c_l + \hat{T}_{lr} {c_l}^\dagger c_r \right].} \tag{3.24}$$

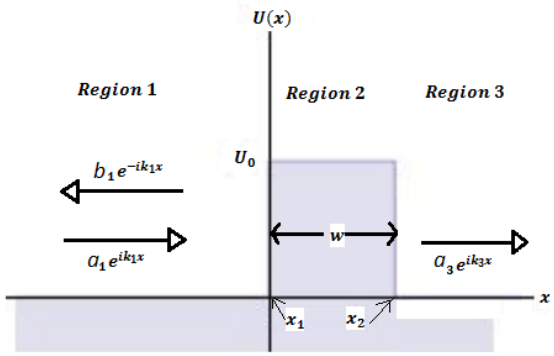


Figure 3.2: Comparison of transfer Hamiltonian approach and boundary condition matching approach for rectangular potential.

This is the transfer Hamiltonian we are looking for. Notice, the indices 'l' and 'r' used to represent just the left and right side of the barrier respectively. But in Equation 3.24 they are used to represent the number of Bloch states in their corresponding sides. So, they have integer values now. This transfer Hamiltonian approach was successful in describing the linear relationship of tunneling current with density of state in superconductor junction. But the question remains, whether this approach can be extended to metals. In Section 3.2 this approach is justified by showing that tunneling probability remains the same as determined from 'matching boundary condition' approach. Experimental validation of QCT (to be discussed in Chapter IV) can further consolidate the claim that the transfer Hamiltonian approach is applicable to regular metals.

3.2 Justification of transfer Hamiltonian approach

Let's assume that an electron has tunneled from left metal to right metal and is now a free electron on the right metal [Region 1 to Region 3 of Figure 3.2]. The amplitude of the incident wavefunction is a_1. Since no particle is incident from the right side, $b_3 = 0$. The transmitted wavefunction amplitude is a_3. Putting these in Equation 2.12, the probability current densities on

each side of the barrier can be written as,

$$j_{in}(x) = \frac{\hbar k_1}{m} {a_1}^2, \qquad \text{(in Region 1)}$$

$$j_{tr}(x) = \frac{\hbar k_3}{m} {a_3}^2. \qquad \text{(in Region 3)}$$

The tunneling probability is, then, calculated in Equation 2.20 by using $j_{tr}(x)$. Let's rewrite the tunneling probability equation here:

$$W_1 = \frac{j_{tr}}{j_{in}} = \frac{16 k_1 k_3 {\kappa_2}^2}{({k_1}^2 + {\kappa_2}^2)({k_3}^2 + {\kappa_2}^2)} e^{-2\kappa_2 w}. \tag{3.25}$$

To justify the transfer Hamiltonian approach, we will determine the transmitted current $\Gamma_{tr}(x)$ from left to right metal, again, but this time by using the tunneling probability current density, j_{rl}, i.e. current density living inside the barrier (Region 2 of Figure 3.2). We will then show that the tunneling probability, W_2 is the same as W_1. Here are the steps to calculate W_2:

1. By putting the value of $\psi_l(x)$ and $\psi_r(x)$ [from Equation 3.5 and 3.6] in Equation 3.7, we get,

$$j_{rl} = {b_{2r}}^{\dagger} a_{2l}. \tag{3.26}$$

2. From Equation 3.20, we get,

$$T_{rl} = -\frac{\hbar^2 \kappa}{m} {b_{2r}}^{\dagger} a_{2l}. \tag{3.27}$$

3. Since $\psi_l(x)$ and $\psi_r(x)$ are exact solutions of the unperturbed Hamiltonian H_0, $\psi_l(x)$ is matched to the wavefunction at the left interface of the barrier. Similarly, $\psi_r(x)$ is matched to the wavefunction at the right interface of the barrier. From matching conditions, we put the values of a_{2l} and b_{2r}, in terms of a_{1l} and a_{3r} respectively. Since the coupling is weak, $a_{2r} \approx b_{2l} \approx 0$. By taking the square of the perturbation matrix element, i.e. $|T_{rl}|^2$,

$$|T_{rl}|^2 = \frac{\hbar^4 {\kappa_2}^2}{m^2} \frac{16 {k_1}^2 {k_3}^2 |a_{1l}|^2 |a_{3r}|^2 e^{-2\kappa_2 w}}{({k_1}^2 + {\kappa_2}^2)({k_3}^2 + {\kappa_2}^2)}. \tag{3.28}$$

4. The density of states in k-space is, $\rho(k) = \frac{1}{2\pi|a|^2}$, where a is the normalization amplitude of the wavefunction. Now, the energy density of states, $\rho(E)$is,

$$\rho(E) = \rho(k).\frac{dk}{dE} = \frac{1}{2\pi|a|^2}\frac{1}{\hbar}\frac{m}{\hbar k} = \frac{m}{2\pi\hbar^2 k|a|^2}. \tag{3.29}$$

From Equation 3.29 it is easy to show that the energy density of final states in the right-side metal, ρ_r is,

$$\rho_r = \frac{m}{2\pi\hbar^2 k_3|a_{3r}|^2}. \tag{3.30}$$

5. Putting the values of $|T_{rl}|^2$ and ρ_r in Equation 3.21, let's calculate the tunneling current, Γ_{tr},

$$\Gamma_{tr} = \frac{\hbar{\kappa_2}^2}{m}\frac{16{k_1}^2 k_3|a_{1l}|^2 e^{-2\kappa_2 w}}{({k_1}^2+\kappa_2{}^2)({k_3}^2+\kappa_2{}^2)}. \tag{3.31}$$

6. Finally, the tunneling probability, W_2, is

$$W_2 = \frac{\Gamma_{tr}}{j_{in}} = \frac{j_{tr}}{j_{in}} = \frac{16k_1k_3{\kappa_2}^2}{({k_1}^2+\kappa_2{}^2)({k_3}^2+\kappa_2{}^2)}e^{-2\kappa_2 w}. \tag{3.32}$$

We can immediately see from Equation 3.25 that, $W_1 = W_2$. The tunneling probability remains the same in 'matching boundary condition' approach and 'transfer Hamiltonian' approach. From this we can deduce that,

$$\boxed{j_{tr} = \Gamma_{tr}.} \tag{3.33}$$

So, the transmitted current density in output-side (region 3) does not change , no matter what approach are we taking to deduce it. But, the upside of choosing 'transfer Hamiltonian' approach is that, it gave us the framework to extend the theory for time-dependent phenomena e.g. modulation of the Fermi level on the input side of the junction. This extension is called 'photon assisted tunneling' theory.

3.3 Total probability current density, $\mathbf{J_{DC}}$

Electrons in a metal have a collective wavefunction with exchange symmetry. Fermionic exchange symmetry confirm that no two electron occupy the same state. For simplicity, we are suppressing any forms of interaction between the electrons and the host crystal (i.e. phonon interaction) and the electrons themselves, (i.e. Coulomb interaction). To describe a collection of interaction-free particle we need some statistical dependence that quantifies the probability of finding a particle at certain energy level. At any temperature, $\mathcal{T}$ the occupation of the states is determined by the Fermi distribution:

$$f(E) = \frac{1}{1 + e^{[(E-E_F)/k\mathcal{T}]}}, \tag{3.34}$$

where, Fermi level, E_F is an energy level below which the probability of finding particle is 1 when temperature, $\mathcal{T} = 0$. In second quantization formulation, the probability of finding an empty state follows this Fermi distribution. So, the probability of a particle occupying a state with energy between E and $E + dE$ is $\rho(E)f(E)dE$. On the other hand, the probability of a particle NOT occupying in a state with energy between E and $E + dE$ is $\rho(E)\left(1 - f(E)\right)dE$. We need to include Fermi distribution $f(E)$ to account for many-body effects for our calculation of probability current density. If we assume electrons are moving from left metal to right metal, then by following the steps mentioned in Table 3.1 :

Table 3.1: Calculation of output probability current density from Region 1 to 3, J^-

	Region 1 (left) to Region 3 (right)
Input wavefunction amplitude	$\|a_1\|$
Free electron Velocity, $v = p/m$	$\hbar k_1/m$
Probability current density per state, j	$j_1 = \|a_1\|^2 \hbar k_1/m$
Fermi occupancy factor	$f(E_l)(1 - f(E_r))$
Energy density of states	$\rho_l(E_l)$
Input probability current density, J_{in}^-	$J_{in}^- = \int_{\infty}^{-\infty} j_1 \rho_l(E_l) f(E_l)(1 - f(E_r)) dE_l$
Tunneling probability	W_{31}
Output probability current density, J^-	$J^- = W_{31} \int_{\infty}^{-\infty} j_1 \rho_l(E_l) f(E_l)(1 - f(E_r)) dE_l$

Table 3.2: Calculation of output probability current density from Region 3 to 1, J^+

	Region 3 (right) to Region 1 (left)
Input wavefunction amplitude	$\|b_3\|$
Free electron Velocity, $v = p/m$	$\hbar k_3/m$
Probability current density per state, j	$j_3 = \|b_3\|^2 \hbar k_3/m$
Fermi occupancy factor	$f(E_r)(1 - f(E_l))$
Energy density of states	$\rho_r(E_r)$
Input probability current density, J_{in}^+	$J_{in}^+ = \int_{\infty}^{-\infty} j_3 \rho_r(E_r) f(E_r)(1 - f(E_l)) dE_r$
Tunneling probability	W_{13}
Output probability current density, J^+	$J^+ = W_{13} \int_{\infty}^{-\infty} j_3 \rho_r(E_r) f(E_r)(1 - f(E_l)) dE_r$

From Table 3.1, the transmitted probability current density oozing out of the right barrier inside Region 3 (see Figure 3.2) is,

$$\begin{aligned}
J^{-} &= W_{31}\sum_{l,r}\int_{-\infty}^{\infty} j_1\rho_l(E_l)f_l(E_l)(1-f(E_r))dE_l,\\
&= \sum_{l,r}\int_{-\infty}^{\infty} {\Gamma_{tr}}^{-}\rho_l(E_l)f_l(E_l)(1-f(E_r))dE_l,\\
&= \sum_{l,r}\int_{-\infty}^{\infty} \underbrace{\left[\int_{-\infty}^{\infty} dE_r\delta(E_l-E_r-E_\omega)(\frac{2\pi}{\hbar})|T_{rl}|^2\rho_r(E_r)\right]}_{{\Gamma_{tr}}^{-}} \rho_l(E_l)f_l(E_l)\big(1-f_r(E_r)\big)dE_l,\\
&= \sum_{l,r}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[(\frac{2\pi}{\hbar^2})|T_{rl}|^2A_r(\omega_r)A_l(\omega_l)\right]\delta(\omega_l-\omega_r-\omega)f_l(\hbar\omega_l)\big(1-f_r(\hbar\omega_r)\big)d\omega_l d\omega_r,
\end{aligned} \tag{3.35}$$

where, $E_{l,r}=\hbar\omega_{l,r}$, $E_\omega=\hbar\omega$ (photon energy) and $A_{l,r}(\omega_{l,r})=\hbar\rho_{l,r}(E_{l,r})$, is the single particle spectral density. Similarly, From Table 3.2, the transmitted probability current density oozing out of the left barrier inside Region 1 (see Figure 3.2) is,

$$\begin{aligned}
J^{+} &= W_{13}\sum_{l,r}\int_{-\infty}^{\infty} j_3\rho_r(E_r)f_r(E_r)(1-f(E_l))dE_r,\\
&= \sum_{l,r}\int_{-\infty}^{\infty} {\Gamma_{tr}}^{+}\rho_r(E_r)f_r(E_r)(1-f(E_l))dE_r,\\
&= \sum_{l,r}\int_{-\infty}^{\infty} \underbrace{\left[\int_{-\infty}^{\infty} dE_l\delta(E_l-E_r-E_\omega)(\frac{2\pi}{\hbar})|T_{lr}|^2\rho_l(E_l)\right]}_{{\Gamma_{tr}}^{+}} \rho_r(E_r)f_r(E_r)\big(1-f_l(E_l)\big)dE_r,\\
&= \sum_{l,r}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[(\frac{2\pi}{\hbar^2})|T_{lr}|^2A_r(\omega_r)A_l(\omega_l)\right]\delta(\omega_l-\omega_r-\omega)f_r(\hbar\omega_r)\big(1-f_l(\hbar\omega_l)\big)d\omega_l d\omega_r.
\end{aligned} \tag{3.36}$$

From Equation 3.17, we noticed, $|T_{rl}|^2=|T_{lr}|^2$. Therefore, from Equation 3.35 and 3.36, the total probability current density can, now, be written as,

$$\begin{aligned}
J_{DC}(\omega) &= J^{-}-J^{+},\\
&= \sum_{l,r}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\frac{2\pi}{\hbar^2})|T_{rl}|^2A_l(\omega_l)A_r(\omega_r)\delta(\omega_l-\omega_r+\omega)\\
&\quad\big[\big(1-f_r(\hbar\omega_r)\big)f_l(\hbar\omega_l)-\big(1-f_l(\hbar\omega_l)\big)f_r(\hbar\omega_r)\big]\,d\omega_l d\omega_r,
\end{aligned}$$

$$\boxed{\Rightarrow J_{DC}(\omega) = \sum_{l,r} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\frac{2\pi}{\hbar^2}) |T_{rl}|^2 A_l(\omega_l) A_r(\omega_r) \delta(\omega_l - \omega_r + \omega) \big(f_l(\hbar\omega_l) - f_r(\hbar\omega_r)\big) d\omega_l d\omega_r.} \tag{3.36}$$

This DC current is the **total probability current density**. This expression for J_{DC} is provided in Tucker's paper [6]. Notice, $J_{DC}(\omega)$ should not be confused as a Fourier transform of a time domain function. The ω in $J_{DC}(\omega)$ represents the input photon frequency,

$$\omega = q_e V_{photon}. \tag{3.37}$$

3.4 Summary

The chapter reveals the inner working of the Bardeen's 'transfer Hamiltonian' approach. Bardeen's goal was to explain linear dependency of tunneling current, J_{DC} to energy density of states $\rho(E)$ in superconducting tunnel junction. Fermi's second golden rule provided a clue towards that linear dependency. The concept of transfer Hamiltonian and its external voltage dependent phase lays the foundation of PAT theory. We will develop this in the next chapter.

CHAPTER IV

GENERALIZED PHOTON ASSISTED TUNNELING THEORY AND QUANTUM CONDUCTIVITY THEORY

'Photon assisted tunneling' theory stand for exactly what the name suggests. The oscillating field of photons helps electrons in one side of the junction to tunnel on the other side. Tien and Gordon [5] proposed the photon assisted tunneling theory, which was generalized by the work of Rogovin, Scalapino [22] and Tucker [23]. To set the stage for PAT theory, let's discuss Figure 4.1. For two metal systems placed very close to each other, if harmonic perturbation occurs in

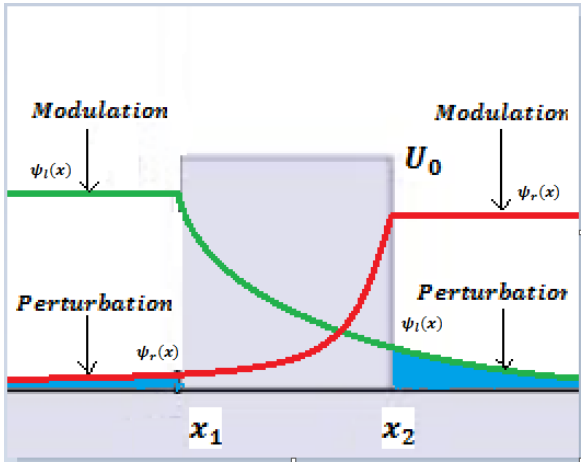


Figure 4.1: Harmonic 'modulation' vs harmonic 'perturbation' in MIM.

one metal system (let's assume, right system in Figure 4.1) due to tunneling electron wavefunction

($\psi_l(x)$, in green) of another metal (left), electron transition occurs from left to right according to Fermi's golden rule. This harmonic perturbation is, in fact, the effect to photon absorption on the other side (left). The electric field of the photon causes harmonic modulation in its energy level. Due to this modulation, the decaying wavefunction of this electron inside the barrier modulates. The attenuation constant, κ inside the barrier is so high that at the right barrier-metal interface the wavefunction amplitude becomes small enough and the harmonic 'modulation' can be consider as harmonic 'perturbation'. The treatment of this harmonic perturbation is done in Chapter III using Bardeen's 'transfer Hamiltonian' approach. The procedure introduced $\hat{H}_T$ in the system Hamiltonian that accounts for tunneling current [see Equation 3.24]. But, the modulation due to photon absorption on the left side that caused this perturbation on the right side has not been treated yet. We will start discussing this modulation in Section 4.1.

4.1 Harmonic modulation of Fermi energy

Let's assume a Voltage, $V(t) = V_{DC} + V_\omega cos(\omega t)$ is introduced in the left side of the barrier, where V_{DC} is the DC voltage (bias voltage) part and $V_\omega cos(\omega t)$ is the AC voltage (modulation voltage) part. Let's assume, q_e is the charge of each electron. Now the Hamiltonian becomes,

$$\hat{H} = \hat{H}_0 + \hat{H}_T + q_e V(t)\hat{N}_l, \tag{4.1}$$

where, $\hat{N}_l = \sum_l \hat{c}_l^\dagger \hat{c}_l$, is the left side number operator. Similarly, the number operator on the right side is, $\hat{N}_r = \sum_r \hat{c}_r^\dagger \hat{c}_r$. For the uncoupled system, the only effect of $V(t)$ is the phase modulation of all the left-side single particle creation-annihilation operator, $\hat{c}_l(t)$.

$$\begin{aligned} \hat{c}_l(t) &= \hat{c}_{l0} e^{-i\frac{q_e}{\hbar}\int_0^t V(t')dt'}, \\ &= \hat{c}_{l0} e^{-i\frac{q_e V_{DC} t}{\hbar}} \int_{-\infty}^{\infty} W(\omega') e^{-i\omega' t} d\omega', \end{aligned} \tag{4.2}$$

where, $\hat{c}_{l0}$ is the Heisenberg operator in the absence of any applied potential (i.e. $V(t) = 0$) and $W(\omega')$ is the Fourier transform of the phase part of the creation operator without the DC voltage,

V_{DC}, i.e.

$$W(\omega') = \int_{-\infty}^{\infty} e^{-i\frac{q_e}{\hbar}\int_0^t [V(t')-V_{DC}]dt'} e^{i\omega' t'} dt' = \int_{-\infty}^{\infty} e^{-i\frac{q_e}{\hbar}\int_0^t [V_\omega cos(\omega t')]dt'} e^{i\omega' t'} dt'. \tag{4.3}$$

So, for the coupled system, after referring to Equation 3.22 and 4.19, we may write the transfer Hamiltonian as,

$$\hat{H}_T(t) = e^{iq_e/\hbar\int_0^t V dt'} \hat{H}_T^+ + e^{-iq_e/\hbar\int_0^t V dt'} \hat{H}_T^-. \tag{4.4}$$

Now in the Heisenberg picture, the current operator is given by [2, 24],

$$\boxed{\hat{J}(t) = q_e \dot{\hat{N}}_r = i\frac{q_e}{\hbar}\left[\hat{H}(t), \hat{N}_r\right].} \tag{4.5}$$

Only $\hat{H}_T$ in Equation 4.1 does not conserve particle number. So, Equation 4.5 becomes,

$$\begin{aligned} \hat{J}(t) &= i\frac{q_e}{\hbar}\left[\hat{H}_T(t), \hat{N}_r\right], \\ &= i\frac{q_e}{\hbar}\left[\hat{H}_T^+(t), \hat{N}_r\right] + i\frac{q_e}{\hbar}\left[\hat{H}_T^-(t), \hat{N}_r\right]. \end{aligned} \tag{4.6}$$

Now, we know Fermions follow the anti-commutation relation, i.e.

$$\left[\hat{c}_k, \hat{c}_{k'}^\dagger\right] = \hat{c}_k\hat{c}_{k'}^\dagger + \hat{c}_{k'}^\dagger\hat{c}_k = \delta_{k,k'} = \delta_{k',k}, \tag{4.7}$$

$$[\hat{c}_k, \hat{c}_{k'}] = \left[\hat{c}_k^\dagger, \hat{c}_{k'}^\dagger\right] = 0,$$

$$\hat{c}_k\hat{c}_{k'} = -\hat{c}_{k'}\hat{c}_k \text{ and } \hat{c}_k^\dagger\hat{c}_{k'}^\dagger = -\hat{c}_{k'}^\dagger\hat{c}_k^\dagger, \tag{4.8}$$

where, $\hat{c}_k$ and $\hat{c}_{k'}^\dagger$ are the Fermionic creation and annihilation operators. k, k' are any states of the system. For our MIM system, the states on left metal are represented with l, l' and the states on the right side metal are represented with r, r'. Using anti-commutation relations (Equation 4.7 and 4.8),

let's find, $\left[\hat{H}_T^+(t), \hat{N}_r\right]$ and $\left[\hat{H}_T^-(t), \hat{N}_r\right]$,

$$\begin{aligned}
\left[\hat{H}_T^-(t), \hat{N}_r\right] &= \sum_{l,r,r'} T_{rl} \left(\hat{c}_r^\dagger \hat{c}_l \hat{c}_{r'}^\dagger \hat{c}_{r'} - \hat{c}_{r'}^\dagger \hat{c}_{r'} \hat{c}_r^\dagger \hat{c}_l\right), \\
&= \sum_{l,r,r'} T_{rl} \left(\hat{c}_r^\dagger \left(\delta_{l,r'} - \hat{c}_{r'}^\dagger \hat{c}_l\right) \hat{c}_{r'} - \hat{c}_{r'}^\dagger \hat{c}_{r'} \hat{c}_r^\dagger \hat{c}_l\right), \\
&= \sum_{l,r,r'} T_{rl} \left(\delta_{l,r'} \hat{c}_r^\dagger \hat{c}_{r'} - \hat{c}_{r'}^\dagger \hat{c}_r^\dagger \hat{c}_{r'} \hat{c}_l - \hat{c}_{r'}^\dagger \hat{c}_{r'} \hat{c}_r^\dagger \hat{c}_l\right), \\
&= \sum_{l,r,r'} T_{rl} \left(\delta_{l,r'} \hat{c}_r^\dagger \hat{c}_{r'} - \hat{c}_{r'}^\dagger \left(\delta_{r,r'} - \hat{c}_{r'} \hat{c}_r^\dagger\right) \hat{c}_l - \hat{c}_{r'}^\dagger \hat{c}_{r'} \hat{c}_r^\dagger \hat{c}_l\right), \\
&= \sum_{l,r,r'} T_{rl} \left(\delta_{l,r'} \hat{c}_r^\dagger \hat{c}_{r'} - \delta_{r,r'} \hat{c}_{r'}^\dagger \hat{c}_l + \cancel{\hat{c}_{r'}^\dagger \hat{c}_{r'} \hat{c}_r^\dagger \hat{c}_l} - \cancel{\hat{c}_{r'}^\dagger \hat{c}_{r'} \hat{c}_r^\dagger \hat{c}_l}\right), \\
&= \sum_{l,r,r'} T_{rl} \delta_{l,r'} \hat{c}_r^\dagger \hat{c}_{r'} - \sum_{l,r,r'} T_{rl} \delta_{r,r'} \hat{c}_{r'}^\dagger \hat{c}_l, && \text{[Since, } \delta_{l,r'} = 0\text{]} \\
&= -\sum_{l,r} T_{rl} \hat{c}_r^\dagger \hat{c}_{l0} e^{-i\frac{q_e}{\hbar}\int_0^t V(t')dt'}, && \left[\hat{c}_l = \hat{c}_{l0} e^{-i\frac{q_e}{\hbar}\int_0^t V(t')dt'}\right] \\
&= -\hat{H}_T^- e^{-i\frac{q_e}{\hbar}\int_0^t V(t')dt'}. && (4.9)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\left[\hat{H}_T^+(t), \hat{N}_r\right] &= \sum_{l,r,r'} T_{lr} \left(\hat{c}_l^\dagger \hat{c}_r \hat{c}_{r'}^\dagger \hat{c}_{r'} - \hat{c}_{r'}^\dagger \hat{c}_{r'} \hat{c}_l^\dagger \hat{c}_r\right), \\
&= \sum_{l,r,r'} T_{lr} \left(\hat{c}_l^\dagger \hat{c}_r \hat{c}_{r'}^\dagger \hat{c}_{r'} - \hat{c}_{r'}^\dagger \left(\delta_{r',l} - \hat{c}_l^\dagger \hat{c}_{r'}\right) \hat{c}_r\right), \\
&= \sum_{l,r,r'} T_{lr} \left(\hat{c}_l^\dagger \hat{c}_r \hat{c}_{r'}^\dagger \hat{c}_{r'} + \hat{c}_l^\dagger \hat{c}_{r'}^\dagger \hat{c}_r \hat{c}_{r'} - \delta_{r',l} \hat{c}_{r'}^\dagger \hat{c}_r\right), \\
&= \sum_{l,r,r'} T_{lr} \left(\hat{c}_l^\dagger \hat{c}_r \hat{c}_{r'}^\dagger \hat{c}_{r'} + \hat{c}_l^\dagger \left(\delta_{r,r'} - \hat{c}_r \hat{c}_{r'}^\dagger\right) \hat{c}_{r'} - \delta_{r',l} \hat{c}_{r'}^\dagger \hat{c}_r\right), \\
&= \sum_{l,r,r'} T_{lr} \left(\cancel{\hat{c}_l^\dagger \hat{c}_r \hat{c}_{r'}^\dagger \hat{c}_{r'}} - \cancel{\hat{c}_l^\dagger \hat{c}_r \hat{c}_{r'}^\dagger \hat{c}_{r'}} + \delta_{r,r'} \hat{c}_l^\dagger \hat{c}_{r'} - \delta_{r',l} \hat{c}_{r'}^\dagger \hat{c}_r\right), \\
&= \sum_{l,r,r'} T_{lr} \delta_{r,r'} \hat{c}_l^\dagger \hat{c}_{r'} - \sum_{l,r,r'} T_{lr} \delta_{r',l} \hat{c}_{r'}^\dagger \hat{c}_r, && \text{[Since, } \delta_{r',l} = 0\text{]} \\
&= \sum_{l,r,r'} T_{lr} \hat{c}_{l0}^\dagger \hat{c}_{r'} e^{i\frac{q_e}{\hbar}\int_0^t V(t')dt'}, && \left[\hat{c}_l^\dagger = \hat{c}_{l0}^\dagger e^{i\frac{q_e}{\hbar}\int_0^t V(t')dt'}\right] \\
&= \hat{H}_T^+ e^{i\frac{q_e}{\hbar}\int_0^t V(t')dt'}. && (4.10)
\end{aligned}$$

Therefore, putting Equation 4.9 and 4.10 in Equation 4.6, we get,

$$\begin{aligned}\hat{J}(t) &= i\frac{q_e}{\hbar}\left(e^{i\frac{q_e}{\hbar}\int_0^t V(t')dt'}\hat{H_T}^{+} - e^{-i\frac{q_e}{\hbar}\int_0^t V(t')dt'}\hat{H}_T^{-}\right), \\ &= -\hat{J}^{+}(t) + \hat{J}^{-}(t), \end{aligned} \tag{4.11}$$

where,

$$\begin{aligned}\hat{J}^{\pm}(t) &= \mp i\frac{q_e}{\hbar}\hat{H}_T^{\pm}e^{\pm i\frac{q_e}{\hbar}\int_0^t V(t')dt'}, \\ &= \hat{J}_0^{\pm}(t)e^{i\phi t}\int_{-\infty}^{\infty} d\omega W(\omega)e^{-i\omega t}, \end{aligned} \tag{4.12}$$

where, again,

$$\hat{J}_0^{\pm}(t) = \pm i\frac{q_e}{\hbar}\hat{H}_T^{\pm}, \tag{4.13}$$

$$\phi = \frac{q_e V_{DC}}{\hbar}, \tag{4.14}$$

and $W(\omega)$ is given in Equation 4.3.

Now, $\hat{J}(t)$ can be expanded in Taylor series (up to first order term), i.e.

$$\begin{aligned}\hat{J}^{(1)}(t) &= \hat{J}^{(0)}(t) + \int_0^t \frac{d}{dt'}\hat{J}^{(0)}(t)dt' + \dots, \\ &= \hat{J}^{(0)}(t) + \frac{i}{\hbar}\int_0^t [\hat{H}_T(t'), \hat{J}^{(0)}(t)]\theta(t-t')dt' + \dots, \quad \text{[Heisenberg picture]} \end{aligned} \tag{4.15}$$

where $\theta(t-t')$ is Heaviside function that enforces the fact that, the coupled system was in thermal equilibrium and $\hat{H}_T$ was adiabatically turned on at some point. According to 'fluctuation dissipation theorem', at thermal equilibrium, voltage across an electrical resistor would statistically show a value with mean 0. But the measurement will have a certain standard deviation. This voltage is called thermal noise or Johnson noise. The process can be assumed as a weak sense stationary (WSS) process, i.e. their auto-correlation between different time-step, only depends on the time difference. $J(t)$ in an MIM can be thought of as a WSS process, in a similar sense. So, the average of the zeroth order term of Equation 4.15 will give us a zero value, i.e. $\left\langle \hat{J}^{(0)}(t) \right\rangle = 0$. Therefore,

average current is the average of the first order term in Equation 4.15. We may write,

$$J(t) = \left\langle \hat{J}(t) \right\rangle = \left\langle \hat{J}^{(1)}(t) \right\rangle = \frac{i}{\hbar} \int_0^t \left\langle [\hat{H}_T(t'), \hat{J}^{(0)}(t)] \right\rangle \theta(t-t')dt'. \tag{4.16}$$

Now let's calculate, $\left\langle [\hat{H}_T(t'), \hat{J}^{(0)}(t)] \right\rangle$ by using Equation 4.11.

$$\begin{aligned} \left\langle \left[\hat{H}_T(t'), \hat{J}^{(0)}(t) \right] \right\rangle &= i\frac{q_e}{\hbar} \left\langle \left[\hat{H}_T^+(t') + \hat{H}_T^-(t'), \hat{H}_T^+(t) - \hat{H}_T^-(t) \right] \right\rangle, \\ &= i\frac{q_e}{\hbar} \left\langle \left[\hat{H}_T^+(t'), \hat{H}_T^+(t) \right] - \left[\hat{H}_T^-(t'), \hat{H}_T^-(t) \right] \right. \\ &\quad \underbrace{\left. + \left[\hat{H}_T^-(t'), \hat{H}_T^+(t) \right] - \left[\hat{H}_T^+(t'), \hat{H}_T^-(t) \right] \right\rangle}_{\text{Quasi-paricle terms}}. \end{aligned} \tag{4.17}$$

We are only interested in quasi-particle terms. So,

$$\begin{aligned} \left\langle \left[\hat{H}_T(t'), \hat{J}^{(0)}(t) \right] \right\rangle_{qp} &= i\frac{q_e}{\hbar} \left\langle \left[\hat{H}_T^-(t'), \hat{H}_T^+(t) \right] - \left[\hat{H}_T^+(t'), \hat{H}_T^-(t) \right] \right\rangle, \\ &= \frac{\hbar}{q_e} \left\langle \left[i\frac{q_e}{\hbar} \hat{H}_T^-(t'), i\frac{q_e}{\hbar} \hat{H}_T^+(t) \right] - \left[i\frac{q_e}{\hbar} \hat{H}_T^+(t'), i\frac{q_e}{\hbar} \hat{H}_T^-(t) \right] \right\rangle, \\ &= -\frac{2\hbar}{q_e} Re \left\{ \left\langle \left[\hat{J}^-(t'), \hat{J}^+(t) \right] \right\rangle \right\}. \end{aligned} \tag{4.18}$$

By putting this value in Equation 4.16, we get,

$$J(t) = \left\langle \hat{J}(t) \right\rangle = -\frac{2i}{q_e} Re \left\{ \int_{-\infty}^{\infty} \theta(t-t') \left\langle \left[\hat{J}^-(t'), \hat{J}^+(t) \right] \right\rangle dt' \right\}. \tag{4.19}$$

From Equation 4.19, 4.12 and 4.14, we get,

$$\boxed{\begin{aligned} J(t) = \left\langle \hat{J}(t) \right\rangle = \frac{-2i}{q_e} Re \Bigg\{ \int_{-\infty}^{\infty} dt' \theta(t-t') \left\langle \left[\hat{J}_0^-(t'), \hat{J}_0^+(t) \right] \right\rangle e^{i\phi(t-t')} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\omega' d\omega'' W(\omega') W^*(\omega'') e^{-i(\omega'-\omega'')t} \Bigg\}. \end{aligned}} \tag{4.20}$$

This is the average current passing through MIM [6].

4.2 Self coupling and higher harmonic generation

In Equation 4.20, when the input across the MIM junction is a DC voltage, i.e. $V(t) = V_{DC}$, from Equation 4.3, we can infer $W(\omega') = \delta(\omega')$. Equation 4.20 then becomes,

$$\left\langle \hat{J}_{DC}(t) \right\rangle = \frac{2}{q_e} Im \Bigg\{ \int_{-\infty}^{\infty} \underbrace{-i\theta(t-t') \left\langle \left[\hat{J}_0^-(t'), \hat{J}_0^+(t) \right] \right\rangle}_{\chi(t-t')} \Bigg\}. \tag{4.21}$$

The value of this $\left\langle \hat{J}_{DC}(t) \right\rangle$ is derived in Equation 3.36. So, $\left\langle \hat{J}_{DC}(t) \right\rangle = J_{DC}(q_e V_{DC}/\hbar)$. The term inside the integrand of Equation 4.21 is noted in the literature as $\chi(t-t')$ [6]. $\chi(t-t')$ is a current correlation function, which is a parameter of WSS process and depends on the time difference only. It can be derived from single particle Green's function from many-body point of view [24].

On the other hand, when $V(t)$ has a frequency component, ω in it, so,

$$V(t) = V_{DC} + V_{\omega} cos(\omega t). \tag{4.22}$$

Equation 4.3 produces Bessel function of different order, n, i.e.

$$\begin{aligned} W(\omega') &= \int_{-\infty}^{\infty} e^{-i\frac{q_e}{\hbar}\int_0^t V_{\omega} cos(\omega t')dt'} e^{i\omega' t'} dt', \\ &= \int_{-\infty}^{\infty} e^{-i\frac{q_e V_{\omega}}{\hbar\omega} sin(\omega t')dt'} e^{i\omega' t'} dt', \\ &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \mathbf{J_n}(\alpha) e^{in\omega t} e^{i\omega' t'} dt', \\ &= \sum_{n=-\infty}^{\infty} \mathbf{J_n}(\alpha)\delta(\omega' - n\omega). \end{aligned} \tag{4.23}$$

where $\mathbf{J_n}$ is a n^{th} order Bessel function, $\alpha = \frac{q_e V_{\omega}}{\hbar\omega}$, $V_{\omega} = E_{\omega} w$, w is the width of the MIM junction and E_{ω} is the local electric field inside junction. From Equation 4.20 we may write,

$$\begin{aligned} J(t) =& Re\left\{ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\omega' d\omega'' W(\omega') W^*(\omega'') e^{-i(\omega'-\omega'')t} J_{DC}(\omega' + q_e V_{DC}/\hbar) \right\}, \\ =& Re\Bigg\{ \sum_{n,m=-\infty}^{\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\omega' d\omega'' \mathbf{J_n}(\alpha)\delta(\omega' - n\omega) \\ & \mathbf{J_{n+m}}^*(\alpha)\delta(\omega'' - (n+m)\omega) e^{-i(\omega'-\omega'')t} J_{DC}(\omega' + q_e V_{DC}/\hbar) \Bigg\}, \\ =& Re\Bigg\{ \sum_{n,m=-\infty}^{\infty} \mathbf{J_n}(\alpha)\mathbf{J_{n+m}}(\alpha) e^{im\omega t} J_{DC}(n\omega + q_e V_{DC}/\hbar) \Bigg\}, \qquad (4.24) \\ =& \sum_{m=0}^{\infty} \left(\frac{J_{m\omega}}{2} e^{-im\omega t} + \frac{J_{m\omega}}{2} e^{im\omega t} \right). \qquad (4.25) \end{aligned}$$

From Equation 4.24 the AC current density, $J_{m\omega}$ with frequency at m^{th} harmonic of the fundamental modulation frequency, ω, can be written as,

$$\boxed{J_{m\omega} = \sum_{n=-\infty}^{+\infty} \mathbf{J_n}(\alpha)[\mathbf{J_{n+m}}(\alpha) + \mathbf{J_{n-m}}(\alpha)]J_{DC}(n\omega + \frac{q_e V_{DC}}{\hbar}).} \tag{4.26}$$

We will expand the Bessel function only for $n = -1, 0, 1$. From Equation 4.26, we will gradually generate values for, $m = 0, 1, 2, 3$ by Haus [10].

For $m = 0$, Equation 4.26 may be written as,

$$\begin{aligned} J_{rect}(\omega, V_{DC}) &= [\mathbf{J_0}]^2 J_{DC}(\frac{q_e V_{DC}}{\hbar}) + [\mathbf{J_1}]^2 \left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar}) \right), \\ &= \left(1 - \frac{\alpha^2}{2}\right) J_{DC}(\frac{q_e V_{DC}}{\hbar}) + \frac{\alpha^2}{4} \left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar}) \right). \end{aligned} \tag{4.27}$$

For $m = 1$, Equation 4.26 may be written as,

$$\begin{aligned} J_{\omega}(\omega, V_{DC}) =& \mathbf{J_1}\left(\mathbf{J_0} + \mathbf{J_2}\right) \left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar}) \right) + \\ & \mathbf{J_1 J_2} \left(J_{DC}(2\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-2\omega + \frac{q_e V_{DC}}{\hbar}) \right), \\ \approx & \frac{\alpha}{2} \left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) - J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar}) \right) \\ & + \frac{\alpha^3}{16} \Big(2J_{DC}(2\omega + \frac{q_e V_{DC}}{\hbar}) - 2J_{DC}(-2\omega + \frac{q_e V_{DC}}{\hbar}) \\ & - J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar}) \Big). \end{aligned} \tag{4.28}$$

For $m = 2$, Equation 4.26 may be written as,

$$\begin{aligned} J_{2\omega}(\omega, V_{DC}) =& 2\mathbf{J_0 J_2} J_{DC}(\frac{q_e V_{DC}}{\hbar}) + \mathbf{J_1}\left(\mathbf{J_3} - \mathbf{J_1}\right) \left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar}) \right) \\ & + \mathbf{J_2}\left(\mathbf{J_4} + \mathbf{J_0}\right) \left(J_{DC}(2\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-2\omega + \frac{q_e V_{DC}}{\hbar}) \right), \\ \approx & \frac{\alpha^2}{4} J_{DC}(\frac{q_e V_{DC}}{\hbar}) - \frac{\alpha^2}{4} \left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar}) \right) \\ & + \frac{\alpha^2}{8} \left(J_{DC}(2\omega + \frac{q_e V_{DC}}{\hbar}) + J_{DC}(-2\omega + \frac{q_e V_{DC}}{\hbar}) \right). \end{aligned} \tag{4.29}$$

For $m = 3$, Equation 4.26 may be written as,

$$\begin{aligned} J_{3\omega}(\omega, V_{DC}) =& \mathbf{J_1}\left(\mathbf{J_4} - \mathbf{J_2}\right)\left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) - J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar})\right) \\ &+ \mathbf{J_2}\left(\mathbf{J_5} - \mathbf{J_1}\right)\left(J_{DC}(2\omega + \frac{q_e V_{DC}}{\hbar}) - J_{DC}(-2\omega + \frac{q_e V_{DC}}{\hbar})\right) \\ &+ \mathbf{J_3}\left(\mathbf{J_6} - \mathbf{J_0}\right)\left(J_{DC}(3\omega + \frac{q_e V_{DC}}{\hbar}) - J_{DC}(-3\omega + \frac{q_e V_{DC}}{\hbar})\right), \\ =& \frac{\alpha^3}{16}\left(J_{DC}(\omega + \frac{q_e V_{DC}}{\hbar}) - J_{DC}(-\omega + \frac{q_e V_{DC}}{\hbar})\right) \\ &- \frac{\alpha^3}{16}\left(J_{DC}(2\omega + \frac{q_e V_{DC}}{\hbar}) - J_{DC}(-2\omega + \frac{q_e V_{DC}}{\hbar})\right) \\ &- \frac{\alpha^3}{48}\left(J_{DC}(3\omega + \frac{q_e V_{DC}}{\hbar}) - J_{DC}(-3\omega + \frac{q_e V_{DC}}{\hbar})\right). \end{aligned} \tag{4.30}$$

Since we have calculated $J_{m\omega}$, we can move on to QCT to further develop the theory.

4.3 Quantum conductivity from PAT theory

In Equation 4.26, notice that, J_{DC} is now a function of $n\omega$. That means, multiple photons can be absorbed or re-emitted while the Fermi levels are modulated by the field of the photon itself. On the other hand, this phenomenon is inducing current density of higher harmonics. Naturally, it is intriguing to make the connection between them through the classical nonlinear constitutive relationship. The generalized classical constitutive relation between is as follows,

$$J_{m\omega} = \begin{cases} \left(\sum_{n=0,1,2,..} \sigma_{m\omega}^{(2n)}(J_{DC}).|E_\omega|^{2n}\right) {E_\omega}^m, & \text{for even m} \\ \left(\sum_{n=0,1,2,..} \sigma_{m\omega}^{(2n+1)}(J_{DC}).|E_\omega|^{2n}\right) {E_\omega}^m, & \text{for odd m} \end{cases} \tag{4.31}$$

where, number of photons involved, N_{photon} is,

$$N_{photon} = \begin{cases} 2n, & \text{for even m} \\ 2n+1, & \text{for odd m} \end{cases} \tag{4.32}$$

Equations 4.27-4.30 are already in the format presented in Equation 4.31, once it is realized that, for a tunneling barrier of width $= w$, $\alpha = \underbrace{\frac{q_e w}{\hbar\omega}}_{constant} \times E_\omega$. All the J_{DC} in Equation 4.27-4.30 can be put into a single parameter, $\sigma_{m\omega}^{(N_{photon})}$ as shown on Equation 4.31. These $\sigma_{m\omega}$ parameters are the

so-called **'quantum conductivity coefficients'**.

For $m = 0, 1, 2, 3$ the generalized constitutive relations of Equation 4.31 reduce to the following:

$$J_{rect} = J_{dc}(q_e V_{DC}) + {\sigma_0}^{(2)} |E_\omega|^2, \tag{4.33}$$

$$J_\omega = {\sigma_\omega}^{(1)} E_\omega + {\sigma_\omega}^{(3)} |E_\omega|^2 E_\omega, \tag{4.34}$$

$$J_{2\omega} = \sigma_{2\omega} {E_\omega}^2, \tag{4.35}$$

$$J_{3\omega} = \sigma_{3\omega} {E_\omega}^3. \tag{4.36}$$

By using Equations 4.27-4.30 and Equations 4.33-4.36, we can derive all the conductivity coefficients arising due to tunneling in MIM junction,

$${\sigma_0}^{(2)} = \left(\frac{q_e w}{2\hbar\omega}\right)^2 \left[J_{DC}(\hbar\omega + q_e V_{DC}) + J_{DC}(-\hbar\omega + q_e V_{DC}) - 2J_{DC}(q_e V_{DC})\right], \tag{4.37}$$

$${\sigma_\omega}^{(1)} = \left(\frac{q_e w}{2\hbar\omega}\right) \left[J_{DC}(\hbar\omega + q_e V_{DC}) - J_{DC}(-\hbar\omega + q_e V_{DC})\right], \tag{4.38}$$

$$\begin{aligned} {\sigma_\omega}^{(3)} = \left(\frac{q_e w}{2\hbar\omega}\right)^3 \big[& J_{DC}(2\hbar\omega + q_e V_{DC}) - J_{DC}(-2\hbar\omega + q_e V_{DC}) \\ & - J_{DC}(\hbar\omega + q_e V_{DC}) + J_{DC}(-\hbar\omega + q_e V_{DC})\big], \end{aligned} \tag{4.39}$$

$$\begin{aligned} \sigma_{2\omega} = \left(\frac{q_e w}{2\hbar\omega}\right)^2 \big[& \frac{1}{2} J_{DC}(q_e V_{DC}) - \left(J_{DC}(\hbar\omega + q_e V_{DC}) + J_{DC}(-\hbar\omega + q_e V_{DC})\right) \\ & + \frac{1}{2}\left(J_{DC}(2\hbar\omega + q_e V_{DC}) + J_{DC}(-2\hbar\omega + q_e V_{DC})\right)\big], \end{aligned} \tag{4.40}$$

$$\begin{aligned} \sigma_{3\omega} = & \frac{1}{2}\left(\frac{q_e w}{2\hbar\omega}\right)^3 \big[\left(J_{DC}(\hbar\omega + q_e V_{DC}) - J_{DC}(-\hbar\omega + q_e V_{DC})\right) \\ & - \left(J_{DC}(2\hbar\omega + q_e V_{DC}) - J_{DC}(-2\hbar\omega + q_e V_{DC})\right) \\ & + \frac{1}{3}\left(J_{DC}(3\hbar\omega + q_e V_{DC}) - J_{DC}(-3\hbar\omega + q_e V_{DC})\right)\big]. \end{aligned} \tag{4.41}$$

Using quantum conductivity coefficient, 'σ's, we have essentially coupled the E_ω to $J_{m\omega}$. The $J_{m\omega}$ will, in turn, radiate $E_{m\omega}$. Calculation of 'σ's for different geometry is done by Haus [10, 25]. This coefficient was used to simulate the radiation field of the MIM junction. In Chapter V we will try to experimentally detect this radiation.

4.4 Summary

According to PAT theory, the effect of incident photons on MIM is two-fold. First, they change (or assist) the DC tunneling current, J_{DC}. Second, there electric field modulates the Fermi level generating higher harmonics of tunneling current, $J_{m\omega}$, m = 1, 2, 3 etc. This tunneling current should couple to a radiative field according to Maxwell's equation. This coupling of E_{ω} and $J_{m\omega}$ is discussed by QCT. Detection of this higher-order harmonics would validate the existence of such higher-order tunneling current. In this experiment, we are trying to detect this radiation.

CHAPTER V

EXPERIMENTS

Our inspiration for this experiment came from the analysis of QCT and its prediction of nonlinear EM radiation. As discussed in Section 4.3, the quantum conductivity terms are functions of E-field of the fundamental photon frequency (FF), E_ω and produce nonlinear tunneling current, $J_{m\omega}$ in the MIM system. The constitutive relation between $E_{m\omega}$ and $J_{m\omega}$ causes the nonlinear tunneling currents to further couple to EM fields and radiate EM waves at harmonic frequencies. Such radiation could be very short-lived due to the geometry and material properties of MIM. If not designed properly, they are absorbed before reaching at the far field. Through our experiments, we intended to detect second harmonic (SH) and third harmonic (TH) signals at the far field generated on MIM samples. Since far field detection was expected to be hard, the input photon frequency (FF) and the geometry and material of MIM were chosen accordingly. The work-flow to experimentally detect SH and TH from MIM at the far field consists of the following steps:

- Simulate the field enhancement at pump (or, input) frequency inside the insulator region of the MIM geometry of our choice i.e. Au nanospheres - plane layer of Al_2O_3 of subnanometer thickness - plane Au surface.

- Use the enhanced pump field amplitude data to calculate the quantum conductivity coefficients for SH and TH frequencies by using Equation 4.34-4.35.

- Introduce this quantum conductivities as insulator parameter to simulate the tunneling-induced SH and TH signals near the MIM interface and, also, at the far field.

- Experimentally detect the SH and TH signal from a physical MIM sample at far field and compare them with simulated results .

The development of simulation models of MIM and experiments on physical MIM were simultaneous and their evolution influenced each other. We sought guidance from the simulation results from time to time. But, For this thesis, we focused mostly on the experimental setup and attempted to physically measure SH and TH signals from MIM samples at the far field. The experimental procedure consists of stages shown in Figure 5.1:

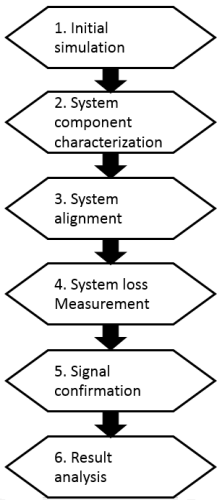
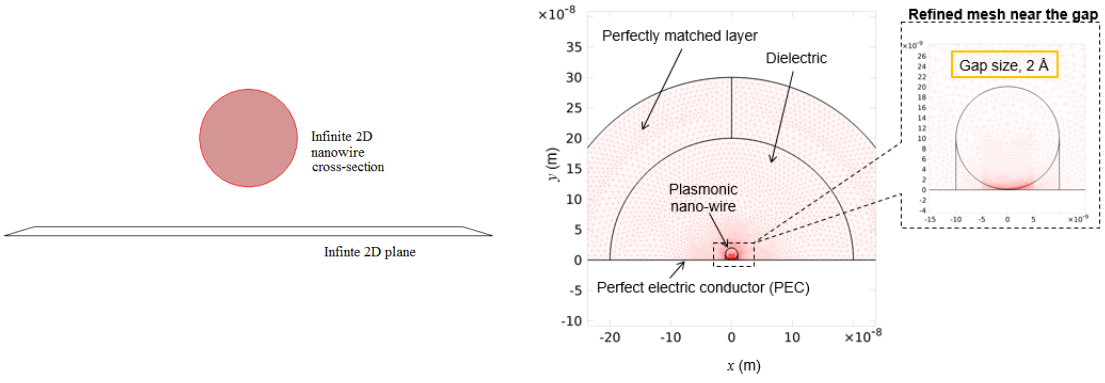


Figure 5.1: Experimental work flow diagram

5.1 Initial simulation

As our first step, we assumed a 2-D geometry consisting of a circle next to a line shown in Figure 5.2(a). In 3-D, the geometry is essentially an infinite metal nano-wire sitting on a thin insulating

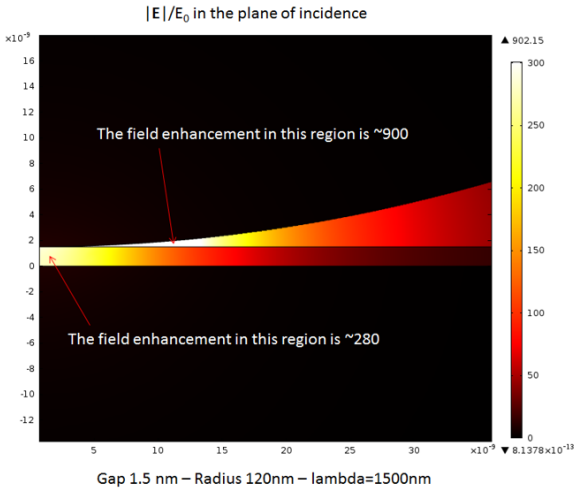
oxide layer that was deposited on a metal surface. This is the simplest geometry for which the metal layer can interact with FF photons and, at the same time, the SH and TH coupled to the tunneling current inside the insulator can radiate into free space and become accessible for detection. For our simulation, we chose $Au - Al_2O_3 - Au$ as our metal-insulator-metal combinations. But they could have been a combination of any other metal and insulator. We intend to use other relevant oxides (TiO_2, Ta_2O_5, HfO_2 etc.) for simulation in future. Simulation was performed using a finite element method (FEM) software called COMSOL. The above-stated geometry of MIM ($Au - Al_2O_3 - Au$) was studied for different radii and gap sizes. It was noted that a field enhancement occurs for all wavelength. The field enhancement inside the gap between the cylinder and the plane is responsible for the quantum tunneling effect. Figure 5.3(b) shows that wavelength around $1500nm$ gave us maximum (around $900\times$) field enhancement[†]. But we decided to use Ti:Sapphire laser (tuned at 810 nm) as our pump because it was available to us and the intensity was measured to be high enough for nonlinear effects.



(a) *(b)*

Figure 5.2: (a) Cross-section of 2D simulation geometry, (b) meshing of simulation geometry.

[†]The simulation was done by Dr. Domenico de Ceglia and shared through private communication.



(a)

Resonance wavelength (nm) as a function of radius and gap size

Radius (nm) \ Gap size (nm)	0.25	0.5	0.75	1	1.5	2
90	1400	1280	1100	1070	1040	1020
100	1470	1360	1250	1180	1160	1140
110	1540	1380	1360	1300	1280	1280
120	1620	1490	1450	1420	1360	1360
130	1690	1580	1540	1520	1500	1480
140	1840	1660	1640	1620	1600	1580

(b)

Figure 5.3: Field enhancement in simulation geometry: (a) color-map representation of normalized electric field between the insulator gap, (b) field enhancement chart for different gap size, cross-section radii and input pump wavelength. The darker the green the larger the field enhancement.

5.2 System component characterization:

To better organize our discussion on the system components, the experimental setup is divided into five major blocks: **input, input path, sample, output path, detector**. Figure 5.4 shows the block diagram of the setup and its components in tandem. We will characterize those components of each block and discuss their positioning and alignment procedures.

5.2.1 Input

The input beam (pump) is a tunable Ti:Sapphire laser tuned at 810 nm. From its spectrum, the -3dB bandwidth was measured to be 10 nm. The temporal pulse width as well as the size, shape and intensity of the beam at the focal length of the first lens L1 was measured and cross-checked with the laser manual.

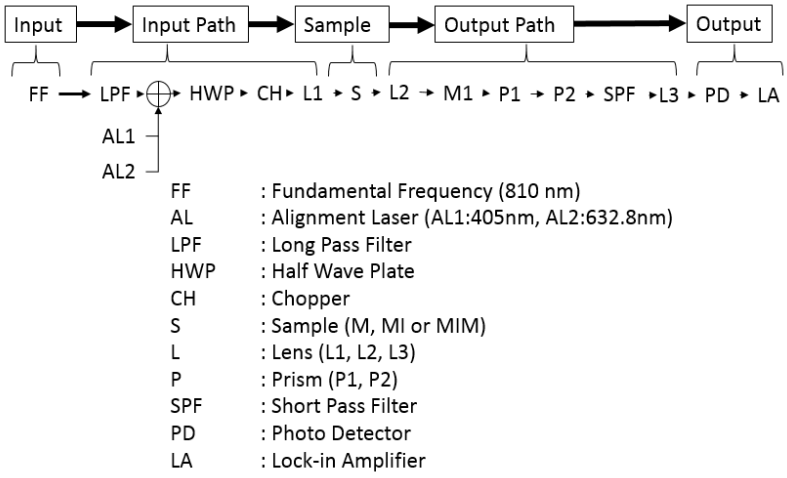


Figure 5.4: Block diagram of the experimental setup.

Pulse width characterization

The pulse width of our laser was characterized using an autocorrelation method. The pulse width is inferred to be $78fs$ [†], which is very close to the value ($t_p = 80fs$) given in the manual. The repetition rate of the pulse is $f_p = 80MHz$.

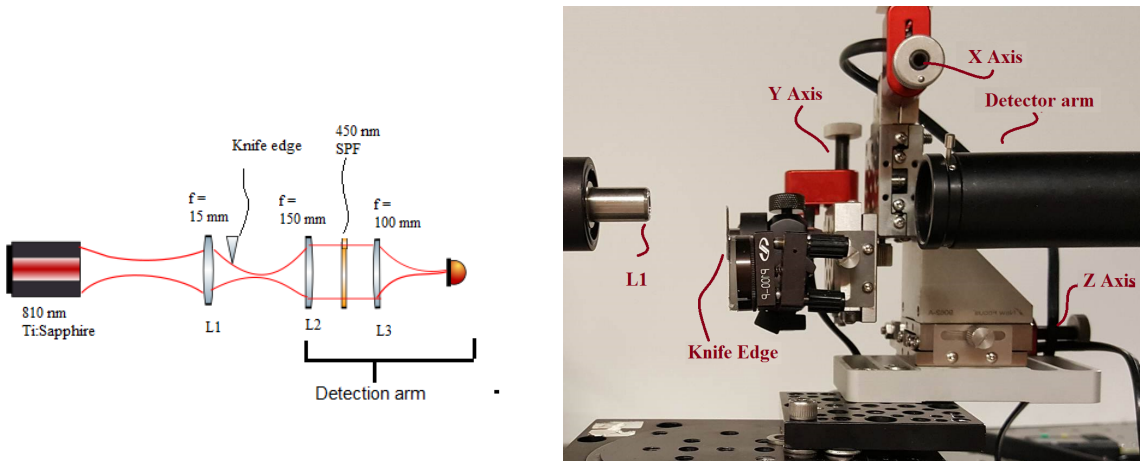
Spot size characterization

The spot size of the beam after focused by L1 is measured by 'knife-edge' experiment. Figure 5.5 shows the sketch and the actual experimental setup. Since the beam shape is elliptical, the experiment was done along both transverse directions (i.e. X axis and Y axis).

Step 1: S curve fitting of intensity, I(Z)

First, the intensity data were acquired by translating the knife edge along the X axis and gradually blocking the beam. Once the beam is completely blocked, the knife edge is translated along the Y axis, and the beam was gradually unblocked (see Figure 5.6). The intensity was fitted to S curve or

[†]The measurement was done and shared by Dr. Andy Chong



(a) *(b)*

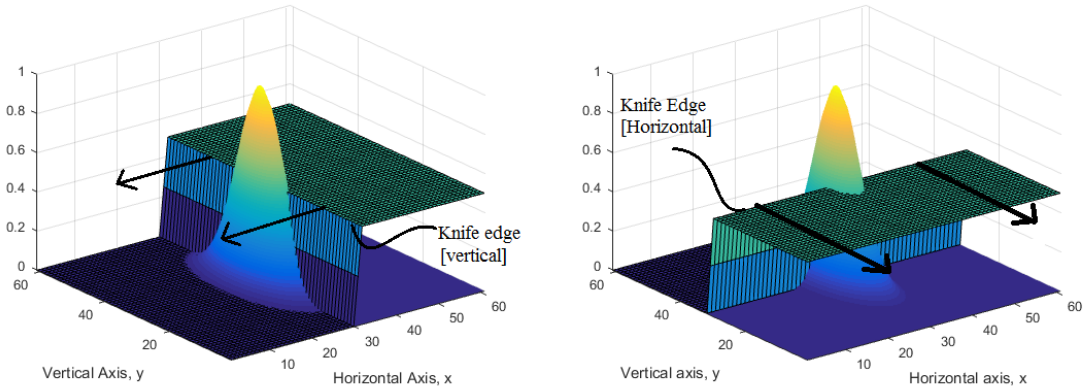
Figure 5.5: Experimental setup of knife-edge experiment: (a) sketch of the experimental setup and (b) actual setup.

$1 \pm erf()$, where $erf()$ is an error function. A sample of fitted S curve near the focal length of L1 (beam waist) is given in Figure 5.7. The $erf()$ equations used for the fitting is as follows:

$$I_{measured}(X,Z) = \frac{P_1}{2}\left(1 - erf(\frac{\sqrt{2}(X-P_2)}{P_3})\right), \tag{5.1}$$

$$I_{measured}(Y,Z) = \frac{P_1}{2}\left(1 + erf(\frac{\sqrt{2}(Y-P_2)}{P_3})\right), \tag{5.2}$$

where, P_1, P_2 and P_3 are the fit parameters. P_1 is the maximum intensity, P_2 is the starting position of the knife edge relative to the beam. P_3 corresponds to the $1/e^2$ radius of the Gaussian beam. Since translation along X axis blocks the beam, the intensity data along X axis is fitted to Equation 5.1 (i.e. $1 - erf()$). Meanwhile, the translation along Y axis releases the beam, so the intensity data along Y axis is fitted to Equation 5.2 (i.e. $1 + erf()$). This process of data acquisition and curve fitting was repeated for 21 different positions of the knife edge along Z axis. The fitted S curve for all Z positions are given in Appendix B. The number of data samples(N_{sample}) inside the beam width at each Z position was picked up from the curve parameter P_3.

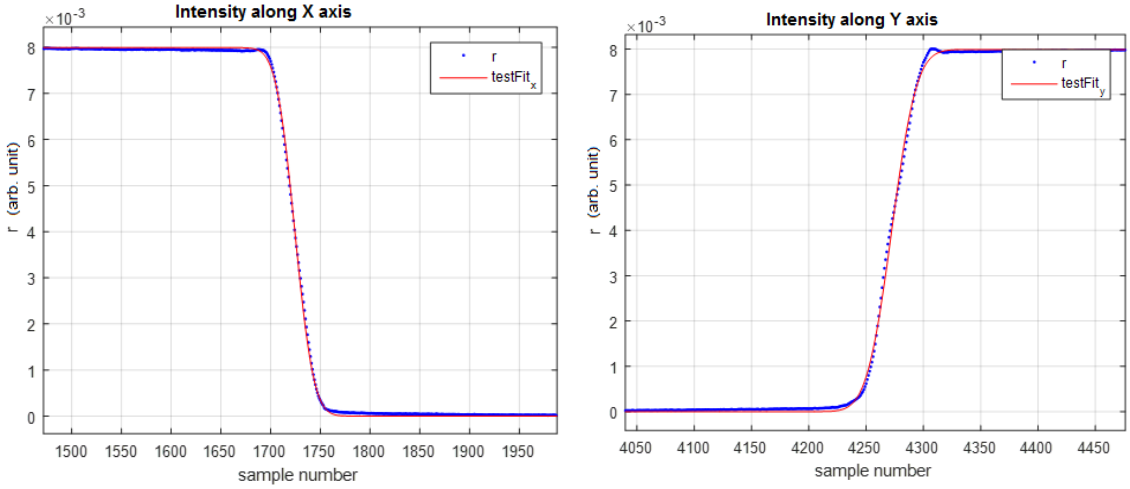


(a) *(b)*

Figure 5.6: Blocking and releasing of beam for knife-edge experiment: (a) Translation of the knife edge along the X axis and blockage of the beam, (b) translation of the knife edge along the Y axis and release of the beam.

For each position along the Z axis the velocities (V_k) of the knife edge (along both X and Y axis) were determined by noting its start and end position and time required for the knife edge to go from the start position to the end position. Table 5.1 shows the velocity values measured on different Z positions. The velocity values, divided by the sampling rate $(f_s = 10Hz)$ i.e. V_k/f_s, provided us the sample spacing, $l_s(\mu m/sample)$, i.e. the distance (μm) traveled by the knife edge per sample of data. Sample spacing, when multiplied by number of sample in the beam waist i.e. $l_s \times N_{sample}$, gives us the measurement of beam width, w.

Precautions: The spatial and temporal resolution of the stage moving the knife edge has to be fine enough to pick up the steep intensity change even at the focal position where the spot size is tiny. Care was taken to confirm that the sample spacing is small enough (by reducing the velocity of the knife edge, and by keeping the sampling frequency large) to have multiple samples even at the focal position where the spot size is the smallest. This helped us to resolve even the small focal spot. The stage used to move the knife edge was a Newfocus Picometer 8752 which has $30nm$ spatial



(a) *(b)*

Figure 5.7: S curve fitting of $I(Z)$: (a) $1 - erf()$ curve fit to intensity data along X axis and (b) $1 + erf()$ curve fit to intensity data along Y axis.

resolution. On the other hand, the time constant of the detection system was small enough ($100ms$) to confirm that the steep change in intensity near the focal spot is not lost. Figure 5.7 shows the result of such careful measurement near focal spot. The number of samples across the region where there is a steep change of intensity is around 50.

Step 2: Quadrature curve fitting of beam width squared, $w^2(Z)$

A quadratic equation that is used to fit the square of beam width i.e. $w^2(Z)$ for different values of Z. The equation of interest is as follows:

$$w^2(Z) = {w_0}^2 \left(1 + \frac{(Z - z_0)}{{z_R}^2}\right). \tag{5.3}$$

Here, w_0 is the spot size, z_0 is the relative position of the beam waist along Z axis, z_R is the Rayleigh range. We fit $w^2(Z)$ to the following parameterized quadratic equation (Equation 5.4) and determined the fit parameters P_1, P_2 and P_3.

$$w^2(Z) = P_1.Z^2 + P_2.Z + P_3. \tag{5.4}$$

Table 5.1: Velocity measurement of knife edge along X and Y axis.

Ref values ->		55	50	43	48			forward= 15.0,	reverse = 9.0			
.mat	z position	x position (start) [mm]	x position (end) [mm]	y position (start) [mm]	y position (end) [mm]	time_x [s]	time_y [s]	Velocity_ x [um/s]	Velocity_ y [um/s]	sample rate [Hz]	x axis sample spacing [um/sample]	y axis sample spacing [um/sample]
test1	29	55	50.57	43	48.37	333	555	13.3033	9.675676	10	1.33033	0.967568
test2	28	55	50.65	43	48.11	333	555	13.06306	9.207207	10	1.306306	0.920721
test3	27	55	50.57	43	48.27	333	555	13.3033	9.495495	10	1.33033	0.94955
test4	26	55	50.73	43	48.43	333	555	12.82282	9.783784	10	1.282282	0.978378
test5	25	55	50.72	43	48.37	333	555	12.85285	9.675676	10	1.285285	0.967568
test6	24	55	50.44	43	48.37	333	555	13.69369	9.675676	10	1.369369	0.967568
test7	23	55	49.84	43	48.21	333	555	15.4955	9.387387	10	1.54955	0.938739
test8	22	55	50.76	43	48.16	333	555	12.73273	9.297297	10	1.273273	0.92973
test9	21	55	50.08	43	48.3	333	555	14.77477	9.54955	10	1.477477	0.954955
test10	20	55	50.45	43	48.21	333	555	13.66366	9.387387	10	1.366366	0.938739
test11	19	55	50.43	43	48.15	333	555	13.72372	9.279279	10	1.372372	0.927928
test12	18	55	50.76	43	48.15	333	555	12.73273	9.279279	10	1.273273	0.927928
test13	17	55	50.27	43	48.45	333	555	14.2042	9.81982	10	1.42042	0.981982
test14	16	55	50.37	43	48.33	333	555	13.9039	9.603604	10	1.39039	0.96036
test15	15	55	50.44	43	48.05	333	555	13.69369	9.099099	10	1.369369	0.90991
test16	14	55	50.08	43	48.41	333	555	14.77477	9.747748	10	1.477477	0.974775
test17	13	55	50.47	43	48.47	333	555	13.6036	9.855856	10	1.36036	0.985586
test18	12	55	50.15	43	48.1	333	555	14.56456	9.189189	10	1.456456	0.918919
test19	11	55	50.47	43	48.08	333	555	13.6036	9.153153	10	1.36036	0.915315
test20	10	55	50.21	43	48.32	333	555	14.38438	9.585586	10	1.438438	0.958559
test21	9	55	50.3	43	48.42	333	555	14.11411	9.765766	10	1.411411	0.976577

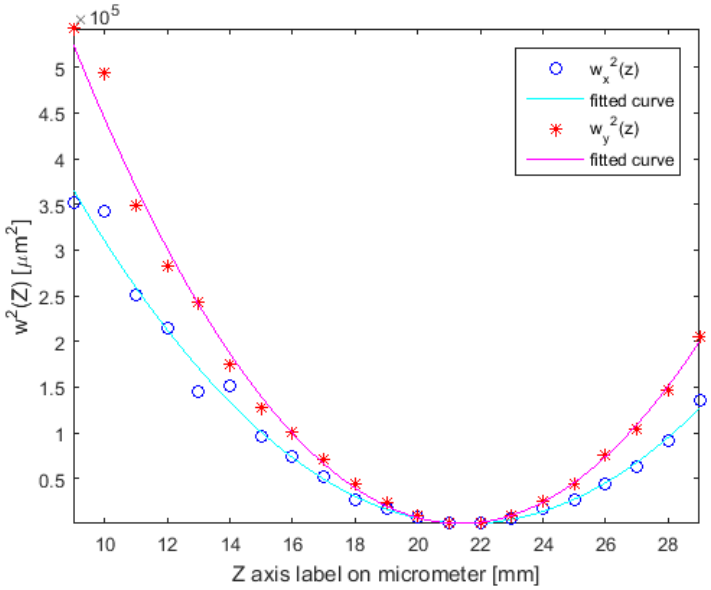


Figure 5.8: Quadrature curve fitting of $w^2(Z)$.

Table 5.2: Calculated fit parameters, P_1, P_2 and P_3.

	x	y
P1	$0.0023 * 10^6$	$0.0034 * 10^6$
P2	$-0.0993 * 10^6$	$-0.1464 * 10^6$
P3	$1.0734 * 10^6$	$1.5654 * 10^6$

Then by comparing the coefficients of Equation 5.3 and 5.4, we calculate the beam parameters of our interest, i.e.

$$w_0 = \sqrt{P_3 - \frac{{P_2}^2}{4.P_1}}, \tag{5.5}$$

$$, z_0 = -\frac{P_2}{2.P_1}, \tag{5.6}$$

$$z_R = \sqrt{\frac{P_3}{P_1} - \frac{{P_2}^2}{4.{P_1}^2}}. \tag{5.7}$$

Equation 5.5, 5.6 and 5.7 is used to characterize the beam (see Table 5.3)

Table 5.3: Characteristic parameters of the FF beam.

	Beam waist, w_0 [um]	Rayleigh range, z_R [mm]	$M^2 = \frac{\pi w_0^2}{\lambda z_R}$	Beam waist position on translation stage, z_0 [mm]
Horizontal, x	25.4891	0.5315	4.7408	21.5980
Vertical, y	23.4510	0.4007	5.3233	21.3737

Remarks: The beam profile is elliptical.

$w_{x,input} \approx w_x.(z = 35) = 0.643mm,$

$w_{y,input} \approx w_y.(z = 35) = 0.797mm.$

As expected from elliptical beam profiles (since, $w_{y,input} \neq w_{x,input}$), the beam waist positions and sizes on vertical and horizontal planes are different.

Beam power characterization

The beam power was measured as follows:

Average power, $P_{avg,FF} = 370mW$
Peak power per pulse, $P_{peak,FF} = \frac{P_{avg,FF}}{t_p \times f_p} = \frac{370[mW]}{80[fs] \times 80[MHz]} = 57.812kW$
Peak intensity at the focal spot of L1, $I_{peak,FF} = \frac{P_{peak,FF}}{A_{spot}}$
$= \frac{57.812[kW]}{(\pi/2) \times 25.4861[\mu m] \times 23.4510[\mu m]} = 6.1574GW/cm^2$

5.2.2 Input path

The input path consists of the following:

Alignment lasers, AL

Once the FF beam is completely characterized, we started designing the input path by choosing alignment lasers. Since HeNe $(632.8nm)$ is visible to naked eye, we used this wavelength to fix the optic axis of the system and the height of all optical components with respect to the optical table. Once prisms (dispersive elements) are introduced on the output path (see 5.2.4), HeNe is no longer useful. To align for SH $(405nm)$, we put a blue laser of same wavelength as a guidance. Careful manipulation of prism position and incident angle (discussed in Section 5.3.3) spares us from the need of another alignment laser at TH (which was not available to us, anyway). All the alignment lasers (FF, HeNe, SH) were made collinear to each other at the input path. A long pass filter (FELH0650) was put in front of the FF laser output aperture to ensure that SH or TH released from the laser cavity (if any) couldn't pollute the signal generated on the sample.

Half waveplate, HWP

From Fresnel's equation of transmission and reflection, we see that the phase of s-polarized light must be zero at the metal surface. So, s-polarization can not interact with metal surface. Only p-polarization with respect to metal surface can generate nonlinear effects on the metal surface. So,

we have introduced a half waveplate in the input path to make the FF beam p-polarized with respect to metal surface.

Input objective lens, L1

Since our input and output arms are separate and the sample is positioned at a relatively large angle ($\approx 70°$), we required enough room to focus the light and collect it at the specular angle. So, we cannot use lenses with high NA or small WD as our L1. Since we intended to achieve a micrometer-ranged spot size, an aspheric lens with minimum spherical aberration was required. We chose A260TM-B with AR coating at FF wavelength as our L1; $EFL(810nm) = 15.31mm$, $NA_{eff} \approx \frac{1[mm]}{15.31[mm]} = 0.0653$. The spot size at the focal length of L1 is measured by knife-edge experiment, discussed in Section 5.2.1. The results are given in Table 5.3.

5.2.3 Sample

The plain Au surface that we used has a surface roughness of following nature:

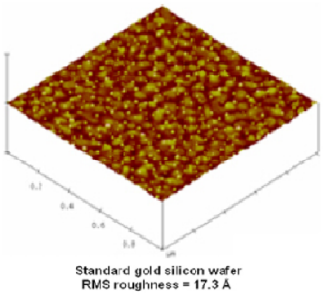
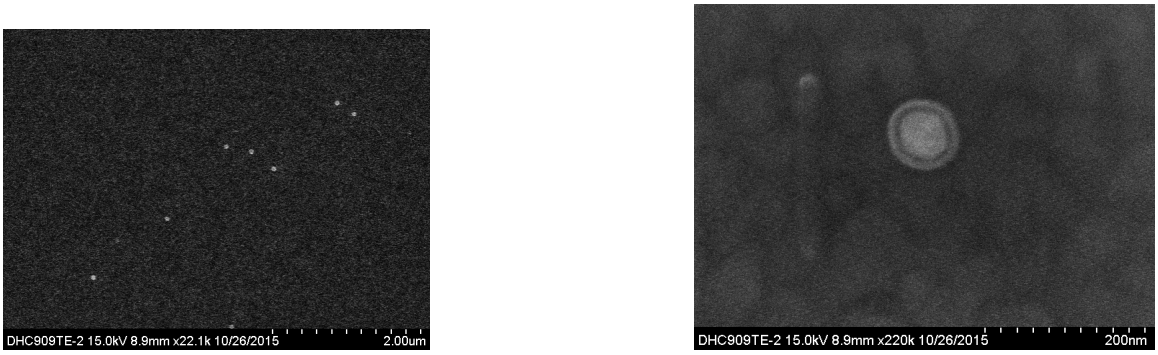


Figure 5.9: The rough surface of a Au sample.

Since MIMs require a subnanometer-ranged thickness for tunneling to take place, we chose the ALD technique to deposit a film of Al_2O_3 on the Au surface. The ALD technique uses precursors like tri-methyl-aluminum (TMA) to deposit Al_2O_3 on Au [26, 27]. Samples of different numbers of

ALD cycles, thus, different ALD thicknesses, were produced[†]. After performing the SH/TH experiments on MI samples, they were treated with UV and O_3 to decontaminate the surface and make it hydrophilic again. To produce MIM samples we spin coated Au nanoparticle solution $(10nm)$ on the MI samples. Then we imaged the surface through SEM to qualitatively get a idea of particle distribution. Here is what we saw through SEM: The particle density is about $20 particles/\mu m^2$.



(a) *(b)*

Figure 5.10: SEM images of Au NP on MI samples

Figure 5.11 illustrates a single MIM geometry on the sample with a p polarized input field.

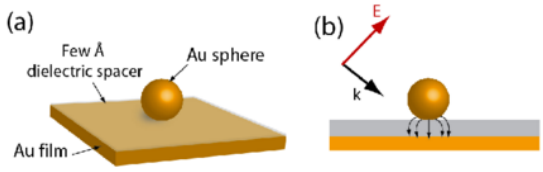


Figure 5.11: Illustration of the MIM structure: nanosphere(Au)-nano-layer of Oxide(Al_2O_3)- Au surface.

[†]The ALD was performed in Washington University in St. Louis by Zhengning Gao

5.2.4 Output path

The output path collects the signal beam and, then, collimates, filters and refocuses it on the detector. Next, we discuss each element of the output arm.

Output objective lens, L2 and tube lens, L3

If the prisms and filters are put aside, the output path, at its simplest level, is a combination of two lens systems that images the metal surface on the detector. Output objective lens (L2) collimates the signal and tube lens (L3) refocuses the signal onto the detector. The following issues were considered while choosing L2 and L3:

- **Surface conic constant:** Since we are only interested in the power of our signal, the imaging need not be perfect. Aberrations are acceptable. So aspheric lenses were not required.
- **Lens material and AR coating:** Since our signal can be of two wavelengths i.e. SH (405 nm) and TH (270 nm), the lenses need to be achromatic and the focal length shift should be as small as possible. Also, the material of the lenses and the AR coating should be such that it allows both SH and TH transmission. If it can block FF, that will be considered as added benefit. UV enhanced fused silica provides all the required material characteristics.
- **Numerical aperture:** The effective NA of L1 is approximately 0.0653. The NA of L2 must match the NA of L1.
- **Diameter:** The diameter of the lenses must be compatible to filter diameters and prism dimensions. We chose 25 mm to be our lens diameter. Since the NA has to be 0.065, the focal length of L2 turned out to be $150mm$. Additionally, the lenses need to be AR coated for SH and TH.

- **System magnification:** The L2-L3 system has to be a demagnifying system. This way, any inevitable, but slight, defocus, decenter, tilt of the sample will be demagnified and the system tolerance will significantly increase. We chose the focal length of L3 to be $100mm$.

We chose L2 and L3 to be ACA254-150-UV and ACA254-100-UV respectively. Those lenses matches all the requirements stated above. We have done a ZEMAX simulation to measure the optimum object distance (from sample to L2) while keeping the detector at the focal plane of L3. The object distance turned out to be $\approx 146mm$. The results are shown in Figure 5.12 to 5.14. The detector size is conveniently larger than the spot size. Also, on axis chromatic aberration is very small.

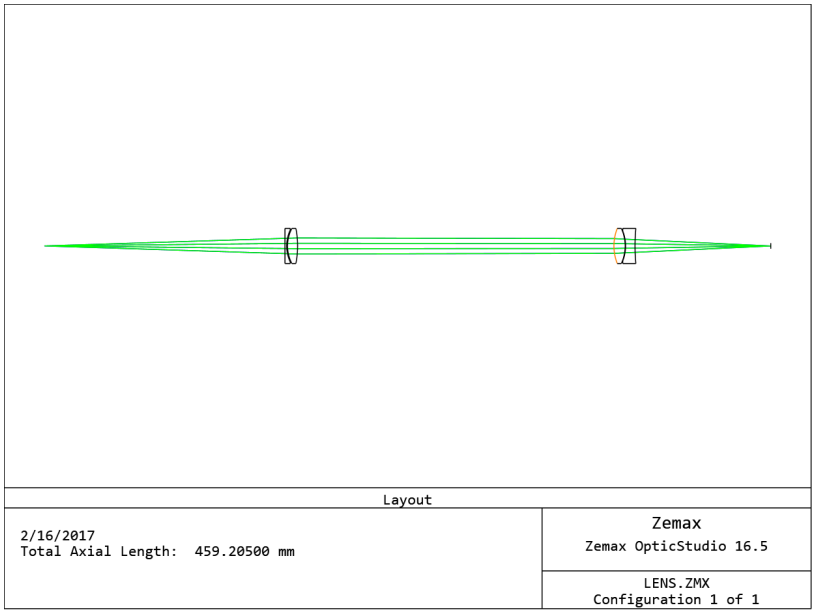


Figure 5.12: Ray diagram of output path (without dispersive elements) consisting of L2 and L3.

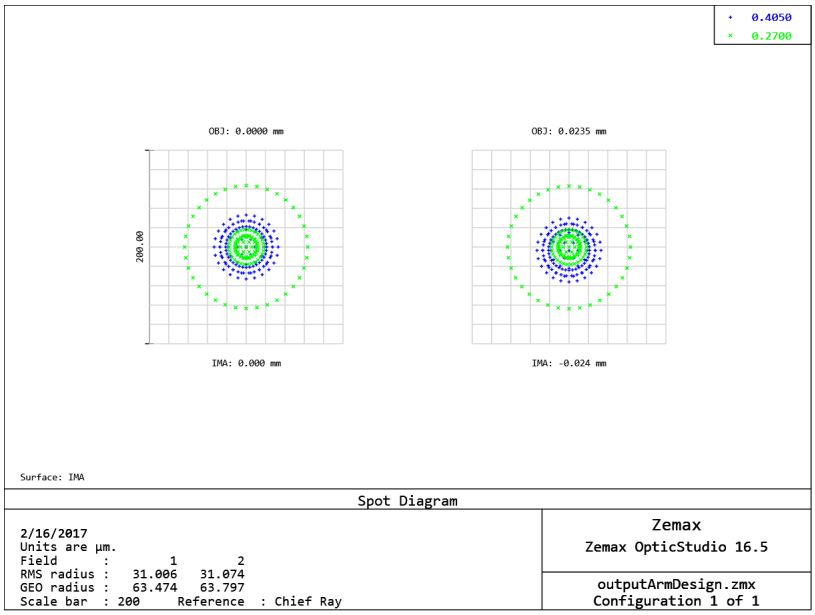


Figure 5.13: Spot diagram of output path (without dispersive elements) consisting of L2 and L3.

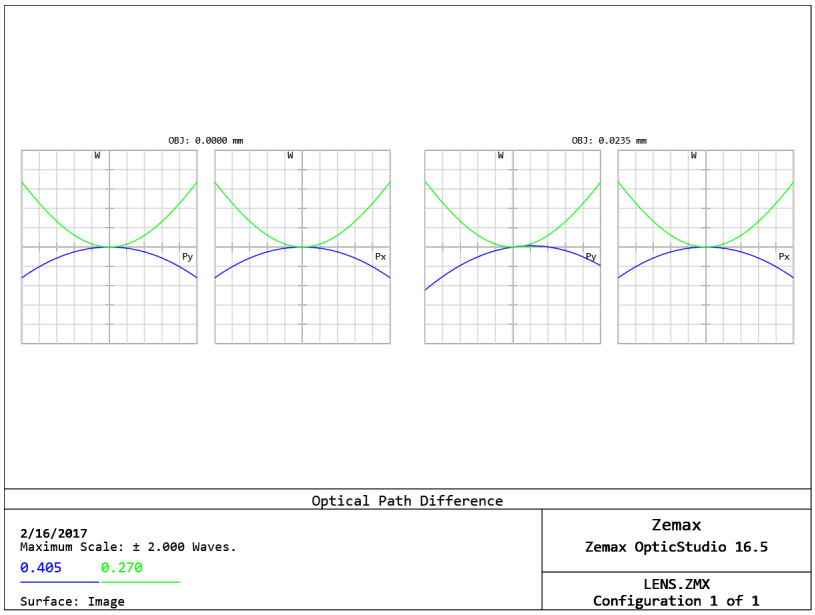


Figure 5.14: The OPD fan diagram at the detector plane. The fan suggests slight defocus of off-axis points of image plane.

Frequency filtering, SPF

The filters used to isolate the SH and TH were FESH0450 (Thorlabs) and FF01-276 (Semrock) respectively. They were put before L3. So, they filtered collimated light and signal's angle of incidence is small. This makes the filtering more efficient. The filters' transmission (normalized to 1) is given in Table 5.4 We could use a couple of FESH0450 filters to extract the SH signal directly at the

Table 5.4: Transmittance data (normalized to 1) of filters for FF, SH and TH.

Wavelength	FF = 810 nm	SH = 405 nm	TH = 270 nm
Filter 1 (FESH0450)	1.58×10^{-6}	0.98197021	5.868×10^{-4}
Filter 2 (FF01-276)	8.2191×10^{-4}	3.544×10^{-5}	0.35569

specular direction from the samples while dropping down the FF power by $(\approx -120dB)$ and submerging it below the noise level $(3.35fW)$ of the detector-lock-in combination. But using multiple FF01-276 to separate TH from the FF and SH would hurt the TH signal (since the transmittance at TH is only 35%). So, we had to introduce prisms to spatially separate the signals from each other.

Spatial filtering, P1/P2

Since we could not find a perfect filter for TH that completely kills the FF, we had to use prisms to further separate the FF from the signals. In fact, putting one prism was not enough. At least two prisms were required to remove FF adequately. The procedure of extracting the SH and TH without widely shifting the detector, requires a careful alignment of the prisms, which will be discussed in Section 5.3.3. Following considerations were crucial during the selection of prism material, incident angle on prism and prism frequency resolution:

- **Prism material:**

  Since our output signal consists of wavelength in the UV range (TH = $270nm$), we used UV graded synthetic fused silica (UVGSFS) that transmits both of our signals (Figure 5.15).

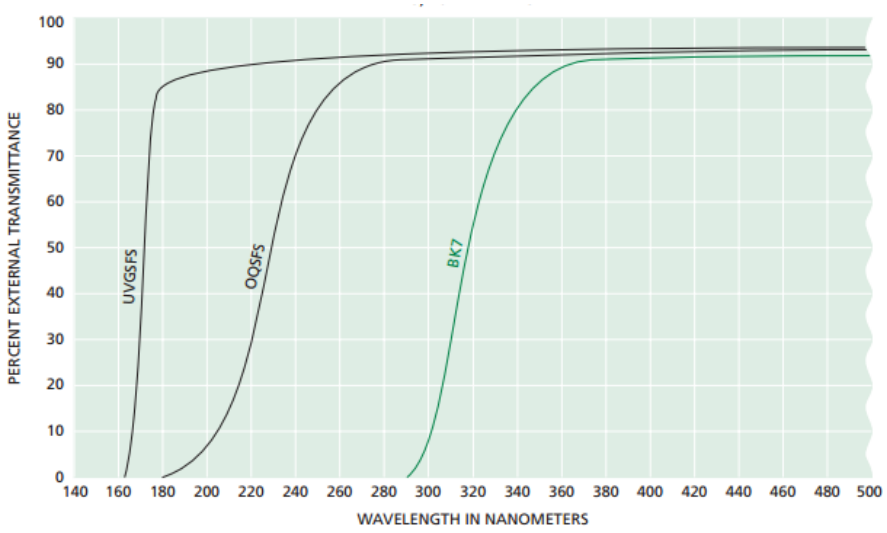


Figure 5.15: Transmittance of UV grade synthetic fused silica.[†]

The refractive index of the UVGSFS can be found using the dispersion relation given in Equation 5.8 [28].

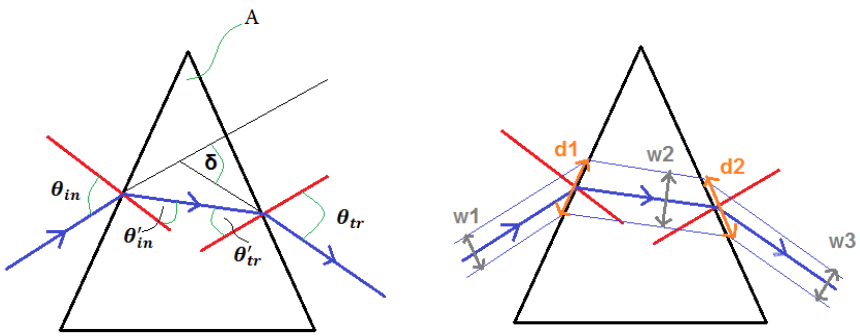
$$n^2(\lambda) - 1 = \frac{0.6961663\lambda^2}{\lambda^2 - (0.0684043)^2} + \frac{0.4079426\lambda^2}{\lambda^2 - (0.1162414)^2} + \frac{0.8974794\lambda^2}{\lambda^2 - (9.896161)^2}. \tag{5.8}$$

We chose two UVGSFS prisms with apex angles 30° (P1) and 45° (P2).

- **Incident angle on prism:**

  The choice of incident angle is bounded by three constraints. A MATLAB code was written utilizing Equation 5.8 to 5.15 to generate a range of incident angles that could be allowed in the system. These constraints influenced the choice of incident angle of the beam.

[†]From manufacturer's website: http://www.slac.stanford.edu/grp/eb/dircrd/melles-griot-fused-silica.pdf

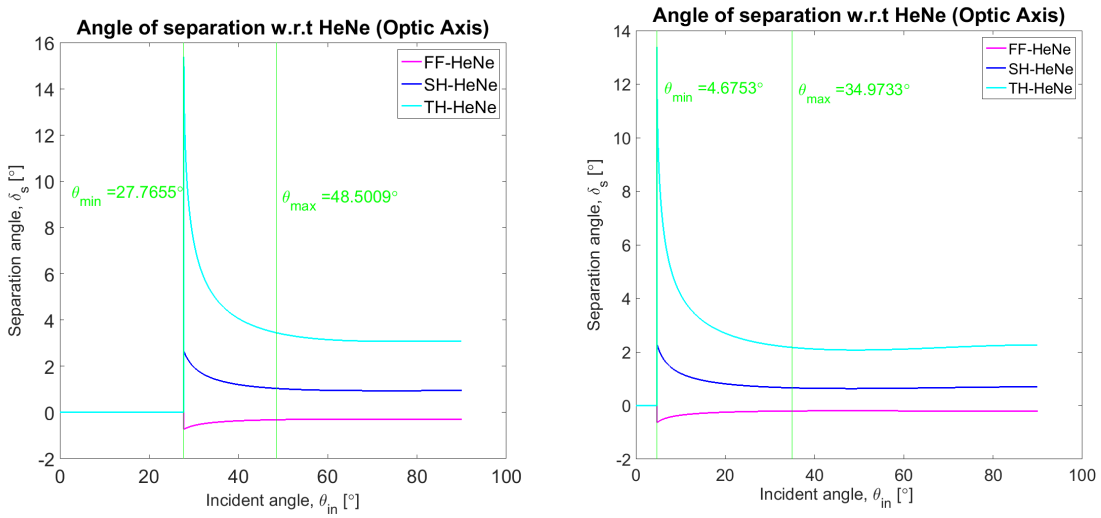


(a) *(b)*

Figure 5.16: Prism constraints: (a) Fresnel transmittance constraint and the angle of separation constraint bound the minimum incident angle and (b) beam width constraint bounds the maximum incident angle.

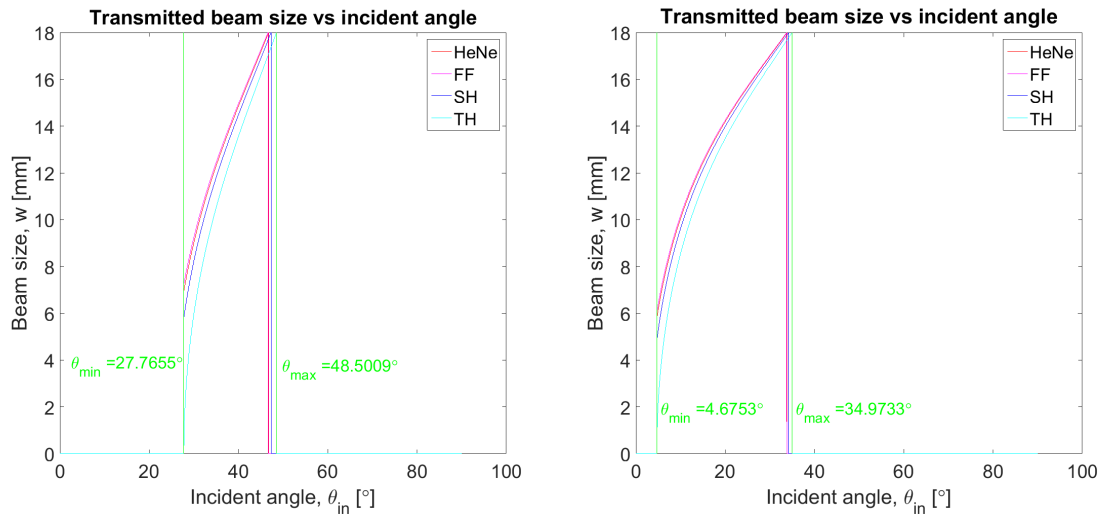
– **Constraint 1 : Angle of separation with respect to optic axis (HeNe)**

The 'angle of deviation', δ in Figure 5.16(a), quantifies the deviation of signal at a certain wavelength from the input beam, whereas, the 'angle of separation' is the difference of angle of deviation between the signal at that certain wavelength and the optic axis (HeNe). The angle of separation gives the angular separation of different wavelengths (with respect to optic axis) after the prism. Lower angular separation would require a large optical path to achieve enough spatial separation between the SH/TH and the FF. Large optical path would lower the throughput of the system. On the other hand, large optical path would lower the field of view, reducing the noise. The angle of separation calculated for incident angles from 0° to 90° for both P1 and P2 is given in Figure 5.17. The angle of separation gives the lower bound of the incident angle. Any incident angle smaller than the lower bound results in total internal reflection on surface 2 of the prism (also evident from Figure 5.19 and 5.20, discussed in constraint 3) and, thus, omitted.



(a) *(b)*

Figure 5.17: Angle of separation with respect to optic axis (HeNe) for (a) P1 (apex angle = 60°) and (b) P2 (apex angle = 45°).



(a) *(b)*

Figure 5.18: Transmitted beam width for (a) P1 (apex angle = 60°) and (b) P2 (apex angle = 45°).

– **Constraint 2 : Transmitted beam width**

From Figure 5.16(b) we deduced the relation between input beam width, $w1$ and the transmitted beam width, $w3$ as follows:

$$w3 = w1 \times \frac{\cos\theta_{tr}}{\cos\theta_{in}} \sqrt{\frac{n(\lambda)^2 - \sin^2\theta_{in}}{n(\lambda)^2 - \sin^2\theta_{tr}}}, \tag{5.9}$$

where,

$$\theta_{tr} = \sin^{-1}(\sin A \sqrt{n^2(\lambda) - \sin^2\theta_{in}} - \sin\theta_{in}\cos A). \tag{5.10}$$

θ_{tr} is a transmission angle, θ_{in} is the incident angle, A is the apex angle of the prism (see Figure 5.16(a)). $n(\lambda)$ is calculated from Equation 5.8. In Figure 5.18, the transmitted beam widths are plotted for both P1 and P2. We require the transmitted beam width to be smaller than the clear aperture of L3 ($< 18mm$). So, any $w1$ that produced $w3 > 18mm$ are omitted and, thus, the transmitted beam width provided the upper bound of θ_{in}.

– **Constraint 3: Fresnel transmittance**

Fresnel transmittance, T_p for different incident angles are calculated using Equation 5.10 to 5.15. Notice, the input beam is p-polarized. For surface 1 of any prism (P1 or P2), t_{p1} is the transmission co-efficient and T_{p1} is the transmittance. The equations to calculate t_{p1} and T_{p1} are:

$$t_{p1}(\lambda) = \frac{2\cos\theta_{in}}{\sqrt{1 - \left(\frac{\sin\theta_{in}}{n(\lambda)}\right)^2} + n(\lambda)\cos\theta_{in}}, \tag{5.11}$$

$$T_{p1}(\lambda) = \frac{n(\lambda)\sqrt{1 - \left(\frac{\sin\theta_{in}}{n(\lambda)}\right)^2}}{\cos\theta_{in}} |t_{p1}|^2. \tag{5.12}$$

Similarly, for surface 2 of the same prism, t_{p2} is the transmission co-efficient and T_{p2} is the transmittance. The equations to calculate t_{p2} and T_{p2} are:

$$t_{p2}(\lambda) = \frac{2n(\lambda)\cos\left(A - \sin^{-1}\left(\frac{\sin\theta_{in}}{n(\lambda)}\right)\right)}{\cos\left(A - \sin^{-1}\left(\frac{\sin\theta_{in}}{n(\lambda)}\right)\right) + n(\lambda)\cos\theta_{tr}}, \tag{5.13}$$

$$T_{p2}(\lambda) = \frac{\cos\theta_{tr}}{n(\lambda)\cos\left(A - \sin^{-1}\left(\frac{\sin\theta_{in}}{n(\lambda)}\right)\right)}|t_{p2}|^2. \tag{5.14}$$

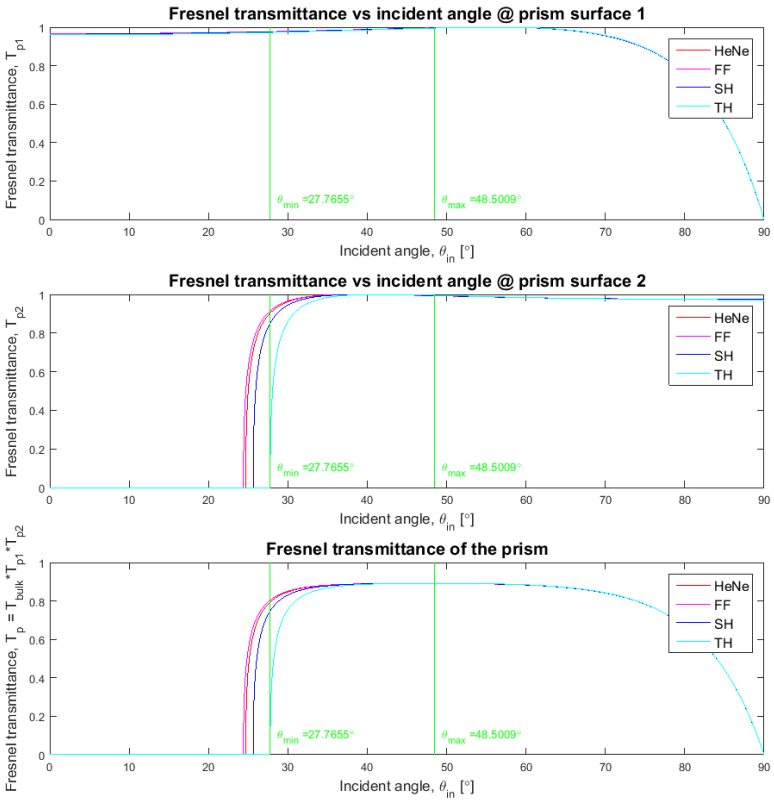


Figure 5.19: Fresnel transmittance for P1 (apex angle = 60°).

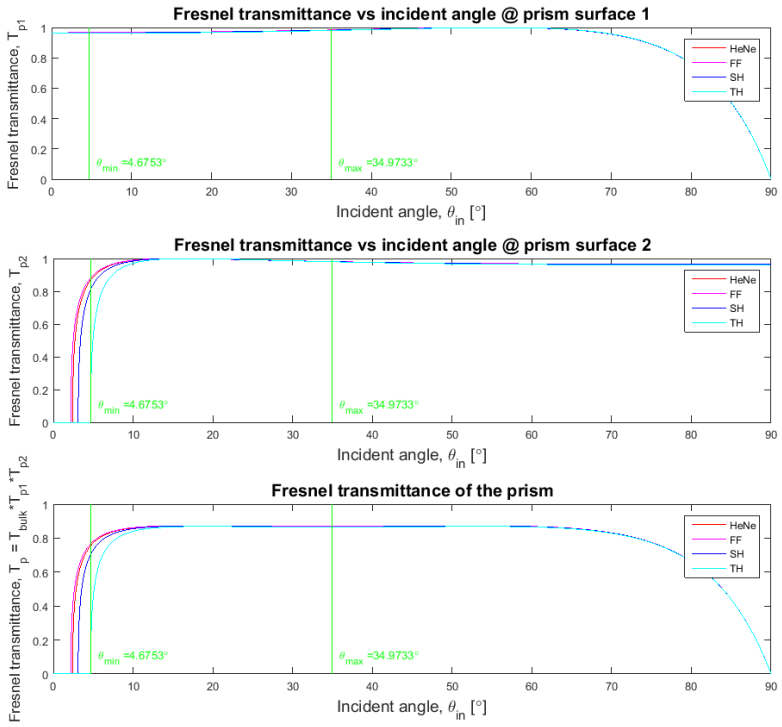


Figure 5.20: Fresnel transmittance for P2 (apex angle = $45°$).

Finally, the total transmittance, T_p is calculated as:

$$T_p(\lambda) = T_{p1}(\lambda) \times T_{p2}(\lambda) \times T_{bulk}(\lambda), \tag{5.15}$$

where, T_{bulk} is picked up from Figure 5.15. The T_{p1}, T_{p2} and T_{bulk} were calculated and plotted for all wavelengths relevant to our experiment.

In summary, the angle of separation and Fresnel transmittance fixed the lower limit of θ_{in}, whereas, the transmitted beam width requirement fixed the upper limit for each prism. We chose the following angles: The reason behind the choice of θ_{max} as the

Table 5.5: Choice of angle of incidence for SH and TH for P1 and P2.

	P1 (Apex angle = 45^o)	P2 (Apex angle = 60^o
Design objective	Incident angle	Incident angle
θ_{SH}	21.0619^0	35.4775^0
θ_{TH}	$\theta_{max} = 34.9733^0$	$\theta_{max} = 48.5009^0$

angle of incidence for TH signal on both P1 and P2 will become evident when we discuss the prism alignment procedure in Section 5.3.3.

- **Prism frequency resolution:**

  Above we considered the angle of incidence; but what about the position of beam profile with respect to the vertex of the prism? Figure 5.16(b) definitely suggests that placing the beam near the edges will chip off the tail of the beam profile. Moreover, choosing to be close to the vertex would be same as choosing a prism with smaller base. The Rayleigh resolving power [29], ρ of a prism is proportional to the length of its base, B, i.e.

$$\rho \propto B. \tag{5.16}$$

  The closer the beam is to vertex, the lower the spatial resolution of the prism. Lower spatial resolution causes the tail of the FF beam profile to leak into the SH/TH space which defeats the purpose of the prism. On the other hand, the further the beam is from the vertex, the greater the lateral shift of the signal, as the prisms are rotated from SH configuration to TH configuration. So, we faced yet another trade off, and decided to keep the incident beam near the middle part of the prism arms.

5.2.5 Output

To choose the right detector, we needed to calculate the expected power of the signal SH or TH. The noise floor of our detector should be sufficiently below the expected power. From literature [30], we collected the SH efficiency to be, $\eta(70^\circ, 6[GW/cm^2]) = 1.5 \times 10^{-11} GW/cm^2$.

In Section 5.2.1 we calculated the peak intensity of our laser to be, $I_{peak,FF} = 6.1574 GW/cm^2$

Now, $\eta(70^\circ, I_{peak,FF}) = \eta(70^\circ, 6[GW/cm^2]) \frac{I_{peak,FF}}{6} = 1.5393 \times 10^{-11}$

Therefore, the expected levels of SH is,

$I_{peak,SH} = \eta(70^\circ, I_{peak,FF}) \times I_{peak,FF} = 0.0947 W/cm^2$
$P_{peak,SH} = I_{peak,SH} \times \frac{\pi}{2} w_{0x} w_{0y} = 889.9 nW$
$P_{avg,SH} = P_{peak,SH} \times t_p \times f_p =$ **5.695 pW**

To successfully detect power in pW range, we needed to use a sensitive detector with low NEP and a lock-in amplifier.

Lock-in amplifier

Since the signal was expected to be submerged into noise, we used a lock-in amplifier (SR830) along with a chopper (SR540). The chopping frequency was set around $310 Hz$. The chopper had a duty cycle of 50%. Since chopping modulates the signal into square waves, lock-in can only pick-up the amplitude of the first frequency of the Fourier spectrum of a square wave. For a voltage $V_{s,rms}$ read from lock-in amplifier, the actual amplitude of the square potential from the photodiode is,

$$V_{pd} = \frac{\pi\sqrt{2}}{4} V_{s,rms} = 1.1107 V_{s,rms} \tag{5.17}$$

The integration time of the lock-in amplifier was $t_{int} = 100 ms$ with a $24 dB$ roll-off filter.

Photo detector

For signal detection, we used a UV-enhanced TE-cooled Si photodiode with H-series dual gain FET input transimpedance amplifier. We operated the photodetector in low-gain configuration ($R_f = 0.6 G\Omega$) at temperature $-40^\circ C$. The power source was also noise free. We measured

the responsivity of the detector at FF, HeNe, SH wavelength (Table 5.6) and compared it with the value given in the manual. Since they match very closely (Figure 5.21), we assumed the responsivity at TH given in the manual is also correct.

Table 5.6: Responsivity measurement of the detector at relevant wavelengths.

	Power measured by a Powermeter, P [μW]	ND filters [to prevent detector saturation], α [dB]	Voltage detected on Lock-in amplifier, , $V_{s,rms}$ [mV]	Voltage from PD, $V_{pd} = 1.1107\, V_{s,rms}$ [mV]	**Responsivity,** $\boldsymbol{R = \frac{V_{pd}}{P \times 10^{-(\alpha/10)} \times R_f}}$ [A/W]
FF	28.5	-66	0.046	0.0511	0.4746
HeNe	8	-50	1.9	2.11	0.4393
SH	2	-50	1.95	2.166	0.1805
TH	-	-	-	-	0.15

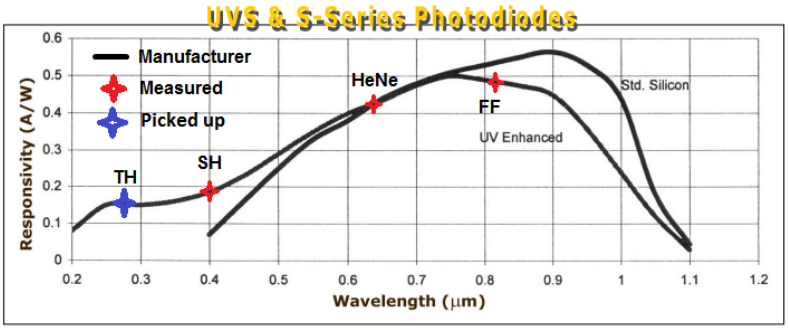


Figure 5.21: Comparison between measured and manufacturer-given responsivity of the detector.

The theoretical maximum that an amplifier can achieve is the DC voltage of the power supply ($\pm 15V$, for our case). A more practical upper limit before saturation occurs would be $\pm 13V$. The maximum power of signal that would saturate the transimpedance amplifiers (assuming the responsivity of the detector remains the same) is given in Table 5.7. On the other hand, the minimum

power that would submerge the signal is calculated from its NEP value ($< 1.5 \times 10^{-14} W/Hz^{1/2}$) [31]. The minimum detectable power (noise floor) is also given in Table 5.7.

Table 5.7: Noise floor and saturation power of the detector at relevant wavelengths.

	$NEP(\lambda) = NEP(810\,nm) \times \frac{R(810\,nm)}{R(\lambda)}$ $[W/Hz^{1/2}]$	$Noise\ Floor = NEP(\lambda) \times \sqrt{1/(2 \times t_{int})}$ $[fW]$	$Saturation\ power = \frac{13}{R_f \times R_{SH}}$ $[nW]$
SH	3.944×10^{-14}	2.789	120.036
TH	4.746×10^{-14}	3.356	144.44

5.3 Alignment procedure

Until now, we have characterized all the elements of our system and discussed all the issues that were considered while selecting them. Next, we will discuss the alignment procedure to put the system together. The sketch of the whole experimental setup is drawn in Figure 5.22. This setup follows the block diagram description given in Figure 5.4. As we complete the alignment of each node (element) on the setup, it will gradually start to take shape and, finally, match the version given in Figure 5.4. Each time new samples are put into the system, all the following alignment procedure needs to be repeated.

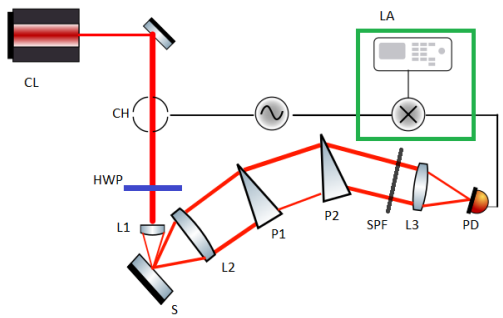


Figure 5.22: Simplified sketch of the experimental setup (CL: collinear lasers).

5.3.1 Alignment lasers

The introduction of alignment lasers in the input path, makes the alignment much easier. Alignment lasers AL1 $(632.8nm)$ and AL2 $(410nm)$ are discussed in Section 5.2.2. The procedure to make them collinear is very trivial and requires one dichroic and one mirror for two of the lasers and two mirrors for the last one. Our setup to make three lasers collinear is given in Figure 5.23:

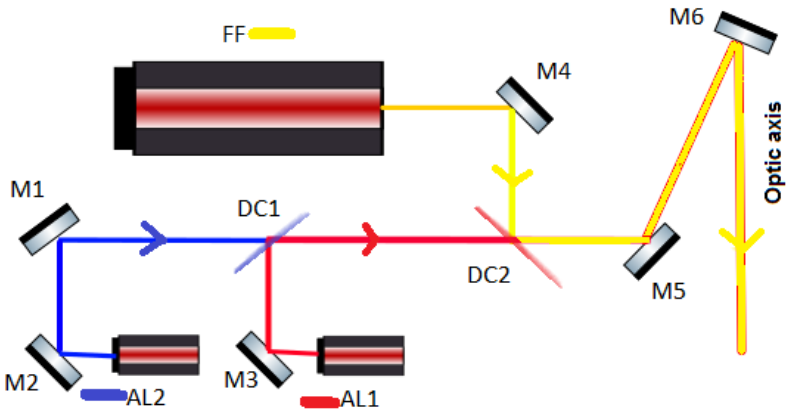


Figure 5.23: CL (Collinear Laser) alignment of AL1, AL2 and FF.

For brevity, the collinear lasers (CL) are shown as one element in Figure 5.22. Once the lasers are collinear, a suitable direction of optic axis, with a suitable height, were chosen. Two irises (IR1, IR2) of same height (placed far from each other along the optic axis) were used to make sure that the optic axis is parallel to the optical table. Also, the irises serves as a memory of our choice of optic axis. In case, any laser slightly drifts, we can always make it come back and hit the pinhole (closed iris) again. Moreover, when closed as far as possible, the iris can be used as an axial point source object (since it is small compared to the relatively large beam width of FF, AL1 or AL2). This becomes very helpful in future alignment. Of course, when experiments were done, these irises were kept completely open.

5.3.2 Sample alignment

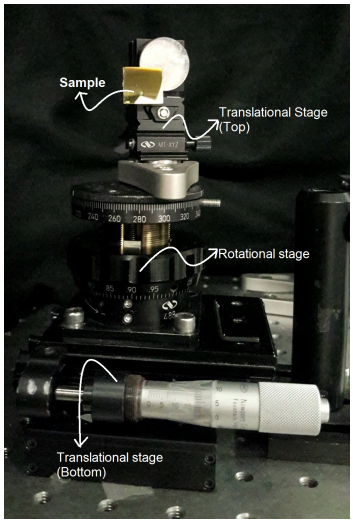


Figure 5.24: The sample stage consists of two translational stage, top and bottom, and a rotational stage.

The sample is placed on a stage that has a rotational stage, a top translational stage and a bottom one. The bottom translational stage moves the rotational axis and confirms that the rotational axis and optic axis are on the same vertical plane. The top translational stage translates the sample and, at the same time, confirms that the rotational axis is almost on the sample surface. This configuration of stages provides any sample with all the degrees of freedom required to fix the reflection angle each time the sample is changed. To achieve that goal, first, the sample is positioned so that the reflected beam back reflects and falls on the same path as the input. This angular position on the rotational stage is assigned the value of $0°$. Once we moved the rotational stage to our desired angular position of $70°$, we immediately fixed our sample and put a UV enhanced Al mirror (M7) to fold the optic axis towards a convenient direction. Four irises were put (IR3, IR4 before M7 and IR5, IR6 after M7) to further fix our optic axis position and height. We made sure that the height of

the optic axis (before and after UV mirror, M7) was the same as we initially chose. The setup after sample alignment and fixing the optic axis in the new geometry looks like Figure 5.25.

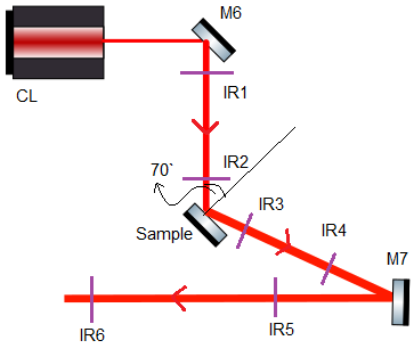


Figure 5.25: Fixing optic axis after the sample with irises (IR: Iris).

5.3.3 Prism P1, P2 alignment

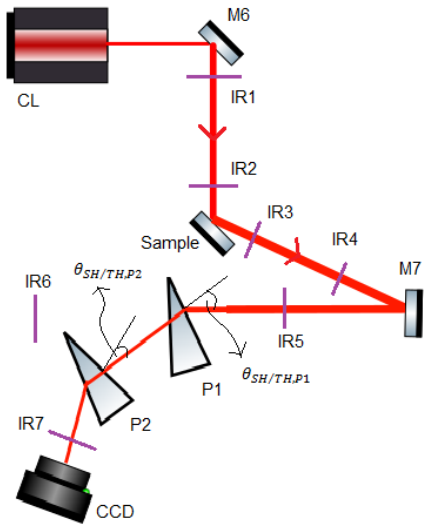
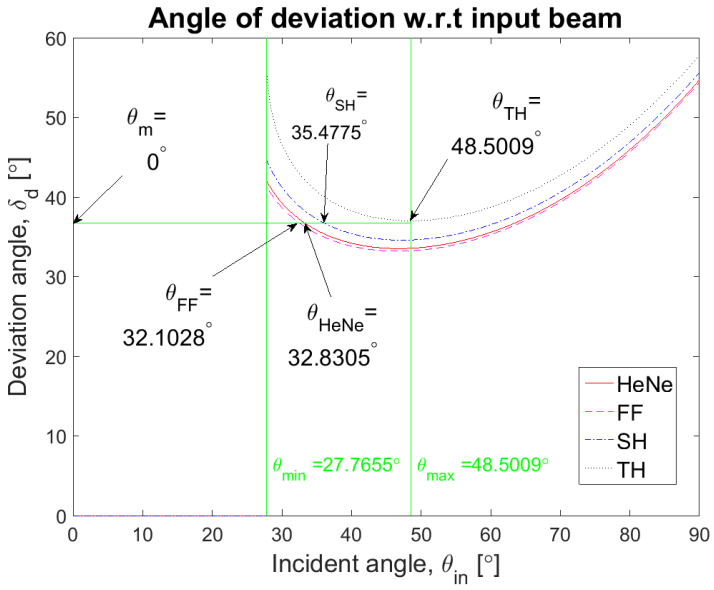
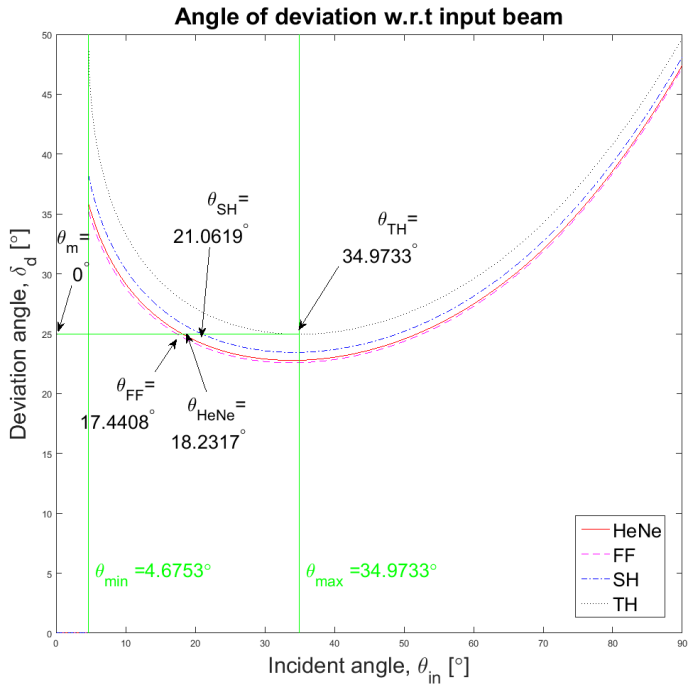


Figure 5.26: Fixing optic axis after the prisms P1, P2 with irises (IR: Iris).

Once the optic axis after the reflection from the sample is fixed with irises, prisms P1 and P2 are placed (see Figure 5.26). Another iris IR7 is put after P2 to confirm that the prisms do not ruin the parallel positioning of the optic axis with respect to the table. As the prisms are rotated to configure for different wavelengths, the iris would tell us whether the wavelength that we intend to detect is collinear to the optic axis. The prisms are sitting on carefully designed stages (similar to the stage designed for the sample) that provides them with all the required degrees of freedom. Introduction of prisms in the system spatially separates the wavelength. Moving the detector to follow the signal would require a stage with very complex dynamics. But, if designed carefully, we can fix our detector to a single position and change the angular position of each prism to confirm that the signal we intend to detect always becomes parallel to the optic axis as it leaves P2. We can achieve that goal by properly choosing a constant angle of deviation for the all the wavelengths. The deviation angle that provides us the maximum allowable incident angle, θ_{max} for TH is a prudent choice. This deviation angle confirms that the incident angle of all the wavelengths are inside the bound region. So, the width of the collimated beam of desirable signals (SH or TH) are equal to or less than the clear aperture of L3. Also, none of the wavelengths will face total internal reflection at surface 2. For example, if we choose the path of HeNe for 32.8305° incident angle on P1 as our the optic axis and fix it by putting irises after P1, once we change the incident angle to 35.4775°, we would notice that SH alignment laser (AL2) hits the iris opening. For the same setup, if we, now, change the incident angle to 32.1028°, the FF beam will hit the iris opening. This coincidence occurs because our choice of incident angles were such that the deviation for all the wavelengths are the same. There could be slight lateral shift of the beams with respect to the optic axis. This will be discussed in Section 5.3.4. Our choice of incident angles and the corresponding rotational axis positions are given in Table 5.8.



(a)



(b)

Figure 5.27: Choice of incident angle on (a) P1 and (b) P2 from its angle of deviation calculation.

Table 5.8: Choice of incident angle on P1 and P2.

	P1 (Apex angle = 45°)		P2 (Apex angle = 60°)	
Design objective	Incident angle	Rotational stage position	Incident angle	Rotational stage position
θ_m	0^0	1.72768	0^0	0 (Home)
θ_{FF}	17.4408^0	0.87664	32.1028^0	32.1028
θ_{HeNe}	18.2317^0	0.83709	32.8305^0	32.8305
θ_{SH}	21.0619^0	0.69558	35.4775^0	35.4775
θ_{TH}	34.9733^0	0 (Home)	48.5009^0	48.5009

All the incident angles can be measured in relation to the incident angle of HeNe (optic axis). Even if there were slight errors in absolute incident angle of the optic axis, as soon as we selected that angle, all the other wavelengths are guaranteed to become parallel to optic axis at relative incident angles. This feature greatly enhances the tolerance of the system with respect to prism position.

5.3.4 Short pass filter SPF, Lens L3 and Photodetector PD alignment

The SPF, L3 and PD are put in tandem with the help of a tube. The tube helps reduce the noise by blocking the light that may fall on the detector from any undesirable angle. It also helps to keep the components aligned to each other. Once an optic axis is selected after P2, the tube is made collinear to the optic axis with help of irises. Inside the tube L3 is fixed at a distance of $85.82mm$ from the detector surface (see Figure 5.28). So, the collimated light focuses on the detector surface. The whole tandem is placed on a translational axis that moves in transverse direction (left-right).

This helps overcome the slight decentering of the signal, once P1 and P2 are reconfigured for a new wavelength detection.

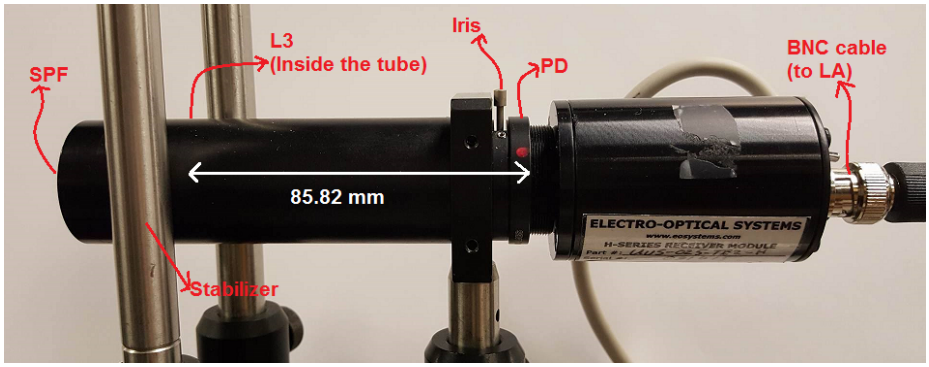


Figure 5.28: Alignment of Short pass filter SPF, Lens L3 and Photodetector PD with respect to Prism P2.

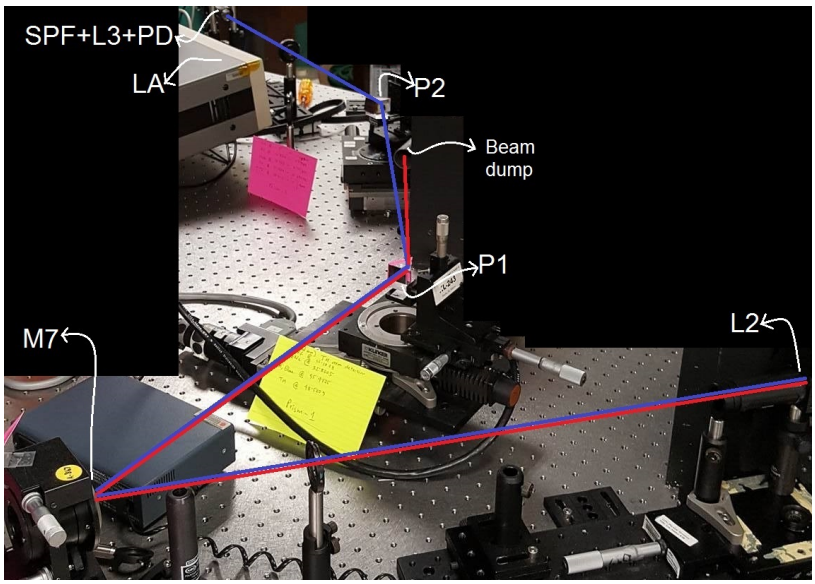
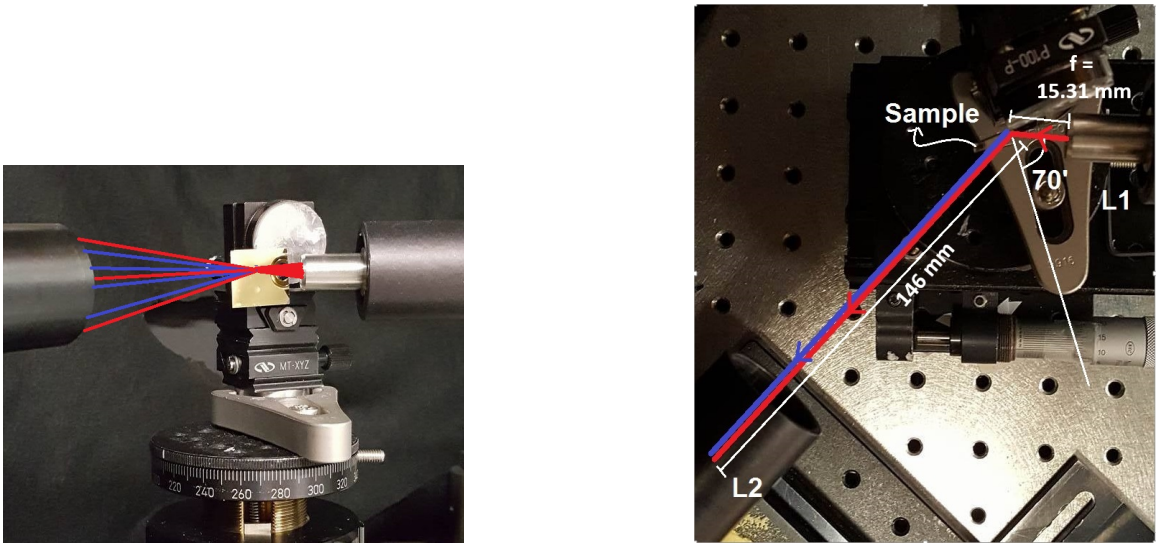


Figure 5.29: Output path after the alignment of Prism P1, P2, Short pass filter SPF, Lens L3 and Photodetector PD.

Any light that gets modulated by the chopper, can potentially get reflected from objects that scatter stray light and fall on the detector introducing noise. So, it is a good practice to catch any unwanted chopped signal in a beam dump. Figure 5.29 shows the real image of the setup after the alignment of the output path was finished.

5.3.5 Lenses L1 and L2 alignment



(a) *(b)*

Figure 5.30: (a) The front view and (b) the top view after L1 and L2 alignment. Red and blue rays represents the pump and the signal respectively.

Now, we can put L1 and L2 in the system. With the help of iris IR3 (see Figure 5.25), we fixed the height of the L2 so that the optic axis goes through the center of it. Since L2 was positioned inside a lens tube which was screwed to the holder, the tip/tilt of L2 can be ignored. Then, we fixed the height and tip/tilt of L1 with the help of IR2 and IR3. This was done by completely closing both IR2 and IR3. This creats an axial object that was image onto small opening of IR3. The irises were kept completely open after the alignment is done. The axial position of L2 is fixed to $146mm$. This

length was determined using ZEMAX by optimizing a merit function for spot size at the detector (see Figure 5.13). The position of L1 is changed until the SH/TH signal reaches a maximum level.

5.3.6 Sanity check with $LiNbO_3$ crystal

After the alignment procedures are complete, a $LiNbO_3$ crystal was placed in place of the sample. As we changed the L1 position and focused on the $LiNbO3$ crystal, we generated SH from it. The SH signal was strong enough to image with a CCD. Figure 5.31(b) shows the blue spot generated from $LiNbO3$ crystal.

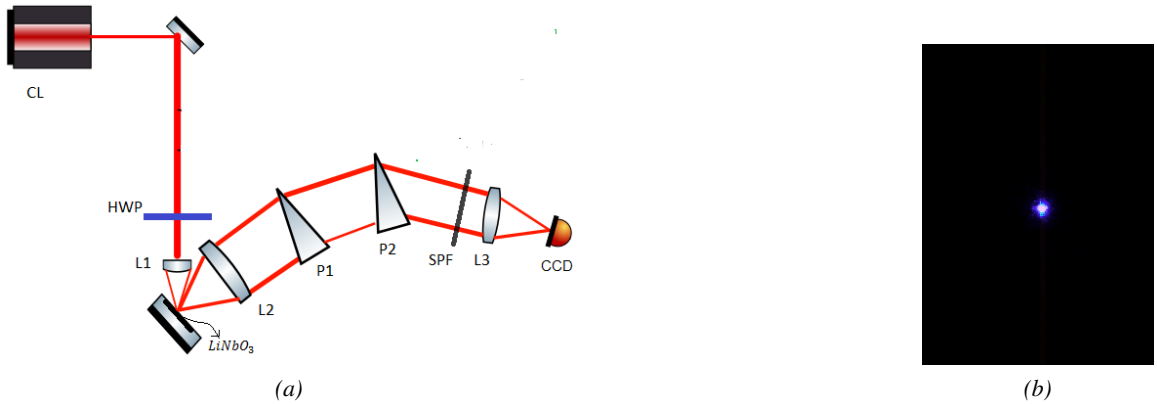


Figure 5.31: (a) Alignment sanity check with $LiNbO_3$ crystal; (b) Image of SH generated from $LiNbO3$ crystal captured with CCD (Lumera).

5.4 System loss measurement

The system performance at respective wavelengths would help us determine the percentage of signal that was lost in the system (system loss). We measured the transmittance of each system component in the system with a power meter and checked whether the data is in agreement with the component's manufacturer data.

Table 5.9: Loss measurement at different nodes of the experimental setup and the total system loss measurement.

	SH = 405 nm			TH = 270 nm
Experimental setup node points	Detected power [μW]	Calculated Transmittance [normalized to 1]	Transmittance from manual	Transmittance from manual
Sample	4.30			>0.35569 (SPF-FF01-276)
Mirror (UV enhanced Al)	4.10	0.953488	> 0.90	> 0.90
P1+P2 (fused Silica)	3.45	0.841463	> 0.80	> 0.80
L3 (Achromatic)	3.08	0.892753	> 0.80	> 0.80
L2 (Achromatic)	2.00	0.649350	> 0.80	> 0.80
Total loss of output path		0.4651	> 0.4608	> 0.1639

As seen on Table 5.9, the calculated transmittance matches the actual measured transmittance of the system very well. Finally, we incorporated this loss to the power of the detected signal to calculate the actual amount of signal generated on the sample.

5.5 Experiments on signal confirmation

Once the signals were detected, two sets of experiment were be done to confirm the presence of signals.

5.5.1 Polarization dependence of signal

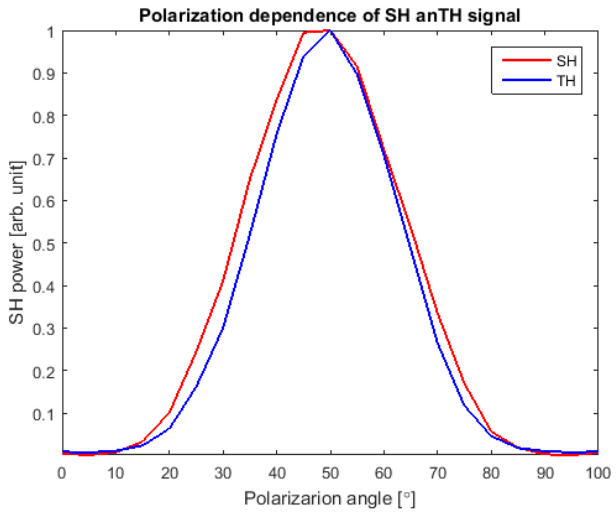


Figure 5.32: Dependence of SH (from plain Au surface) on FF beam polarization.

The FF pump laser is vertically polarized. At 45° angle from the fast axis of the crystal, the polarization of the FF beam turned 90° and became p-polarized light with respect to the sample. Figure 5.32 suggests that the SH/TH signal reached maximum for p-polarized light. Only p-polarization can interact with metal surface and generate SH/TH. So, detected signal was definitely SH/TH generated on the sample.

5.5.2 Pump intensity dependence of signal

Signal from MI samples

The ALD thickness values are approximate values.

ALD thickness : $0nm$

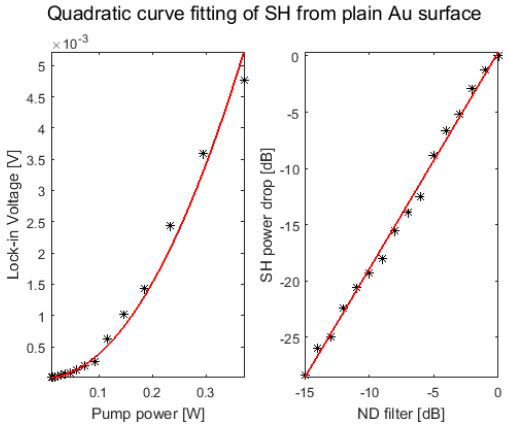


Figure 5.33: Quadrature dependence of SH on FF intensity. Sample is a plain Au surface.

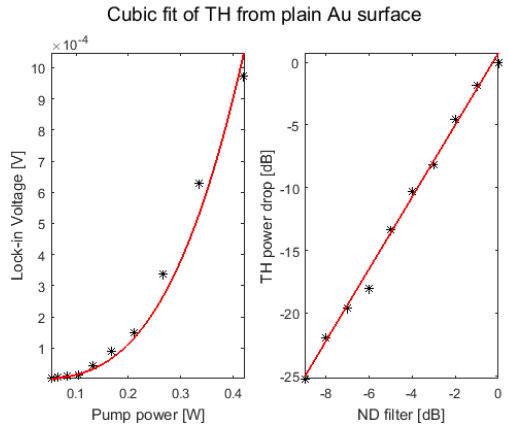


Figure 5.34: Cubic dependence of TH on FF intensity. Sample is a plain Au surface.

ALD thickness : $2nm$

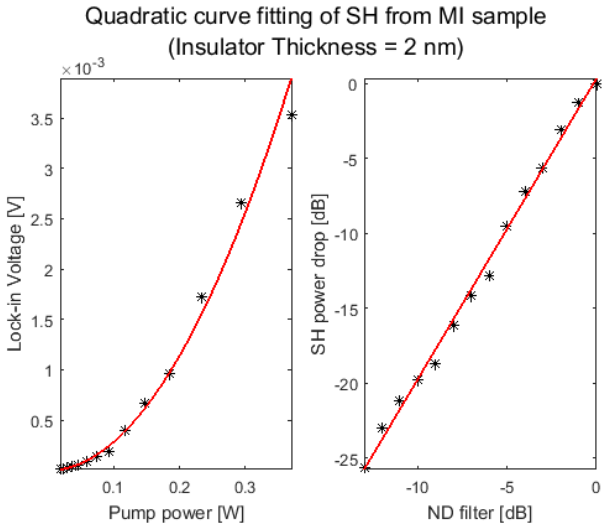


Figure 5.35: Quadrature dependence of SH on FF intensity. Sample is a $2nm$ ALD coated Au surface.

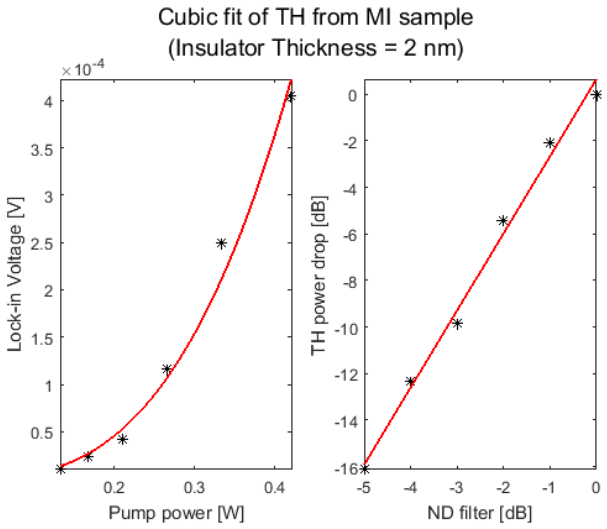


Figure 5.36: Cubic dependence of SH on FF intensity. Sample is a $2nm$ ALD coated Au surface.

ALD thickness : $4nm$

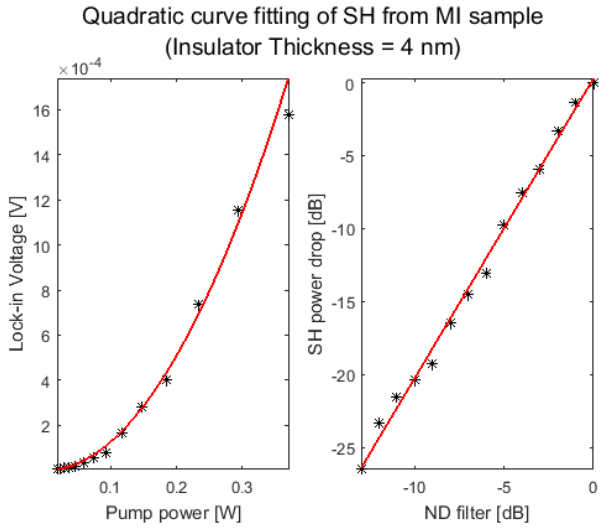


Figure 5.37: Quadrature dependence of SH on FF intensity. Sample is a $4nm$ ALD coated Au surface.

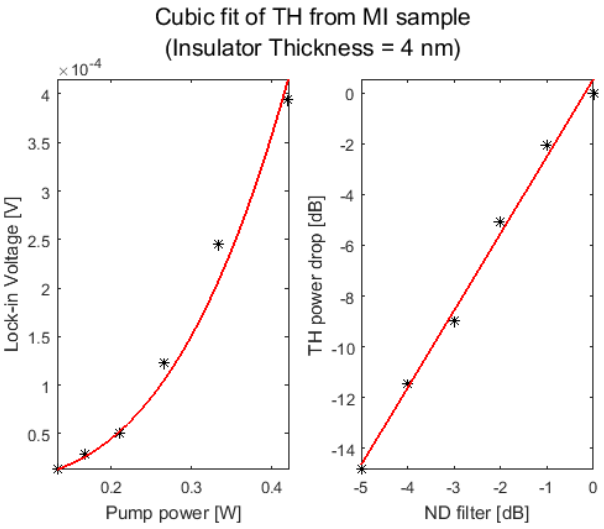


Figure 5.38: Cubic dependence of SH on FF intensity. Sample is a $4nm$ ALD coated Au surface.

ALD thickness : $10nm$

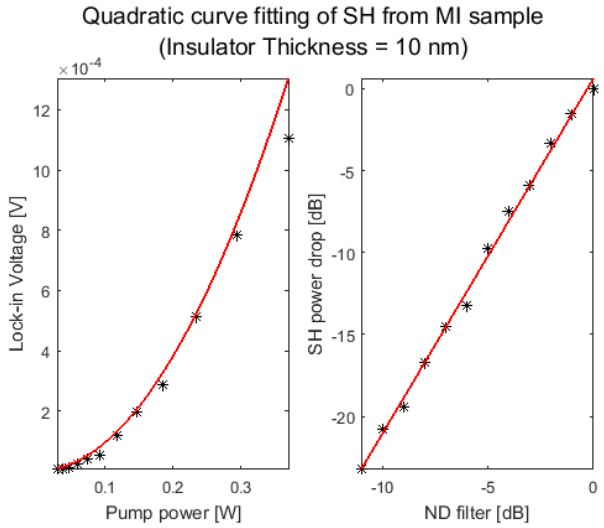


Figure 5.39: Quadrature dependence of SH on FF intensity. Sample is a $10nm$ *ALD coated Au surface.*

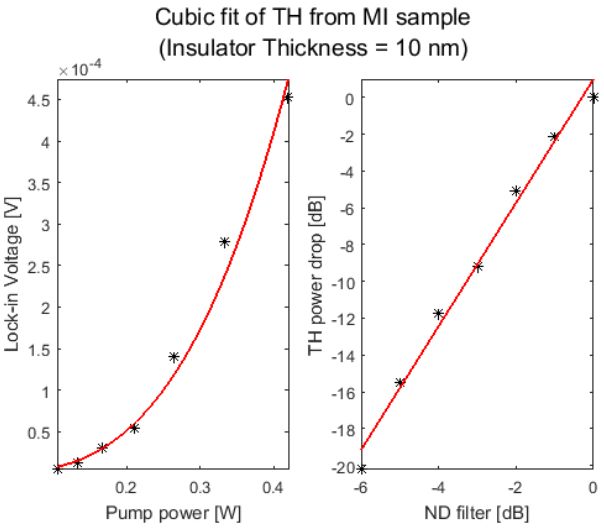


Figure 5.40: Cubic dependence of SH on FF intensity. Sample is a $10nm$ *ALD coated Au surface.*

ALD thickness : $20nm$

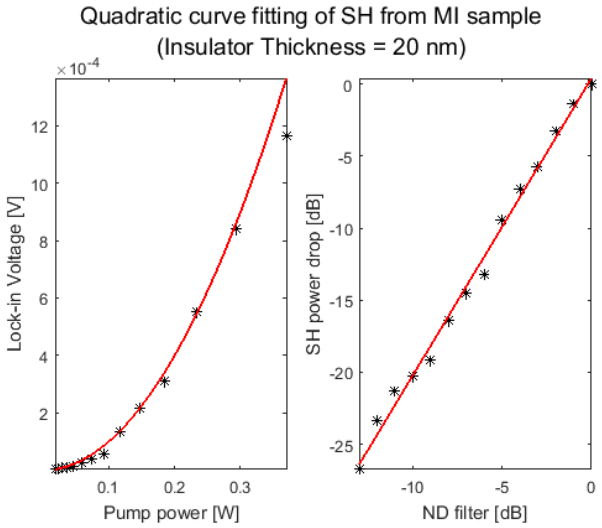


Figure 5.41: Quadrature dependence of SH on FF intensity. Sample is a $20nm$ *ALD coated Au surface.*

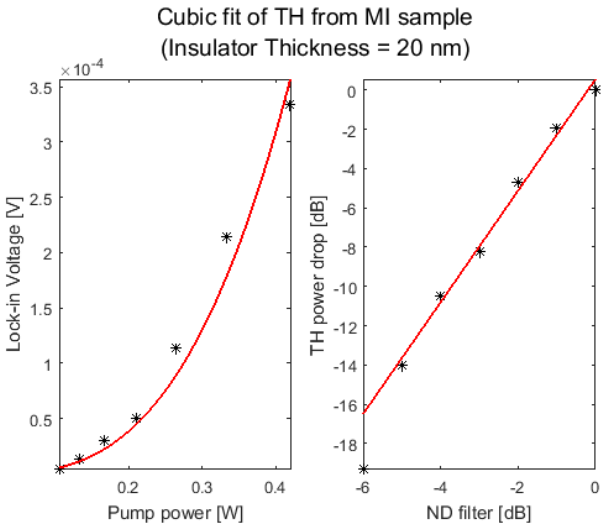


Figure 5.42: Cubic dependence of SH on FF intensity. Sample is a $20nm$ *ALD coated Au surface.*

Signal from MIM samples

The ALD thickness values are approximate values.

ALD thickness : $0nm$

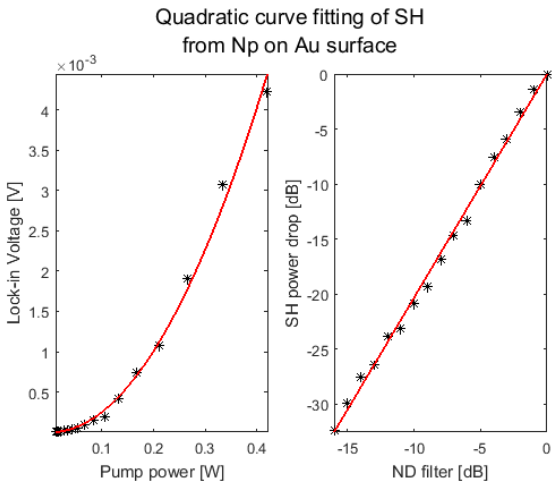


Figure 5.43: Quadrature dependence of SH on FF intensity. Sample is a plain Au surface with NPs.

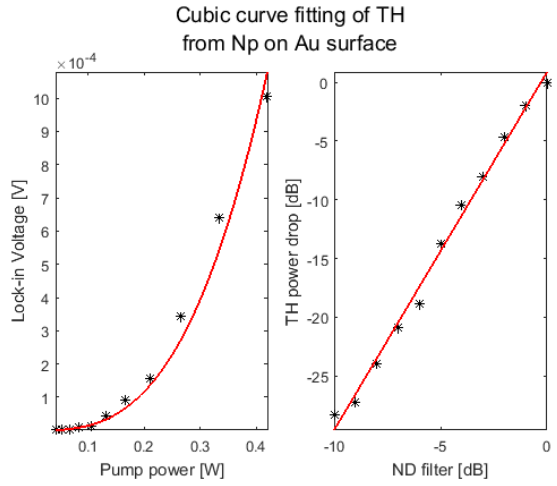


Figure 5.44: Cubic dependence of TH on FF intensity. Sample is a plain Au surface with NPs.

ALD thickness : $2nm$

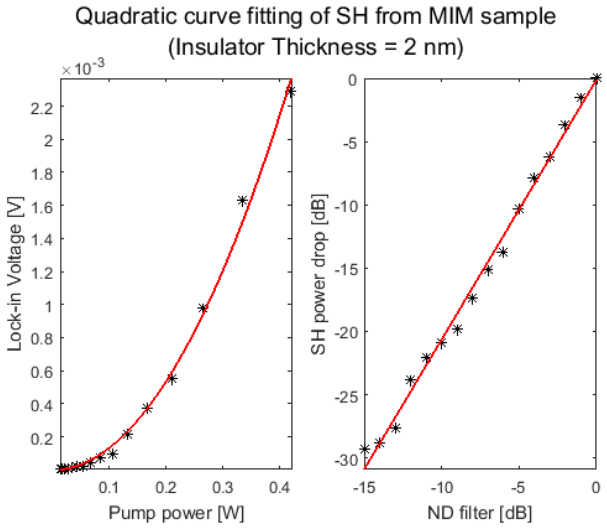


Figure 5.45: Quadrature dependence of SH on FF intensity. Sample is a $2nm$ *ALD coated Au surface with NPs.*

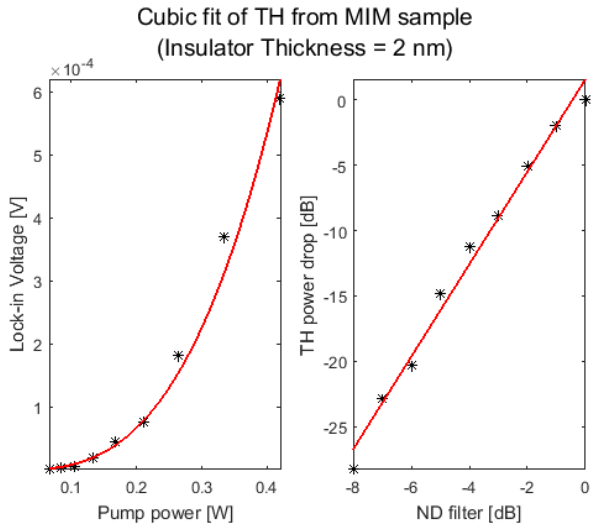


Figure 5.46: Cubic dependence of TH on FF intensity. Sample is a $2nm$ *ALD coated Au surface with NPs.*

ALD thickness : $4nm$

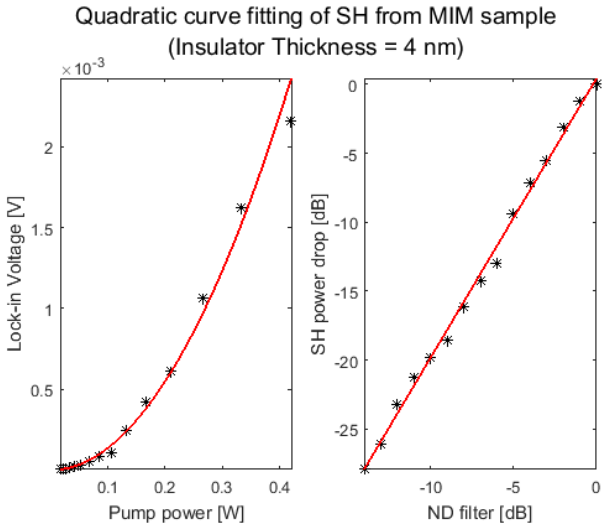


Figure 5.47: Quadrature dependence of SH on FF intensity. Sample is a $4nm$ ALD coated Au surface with NPs.

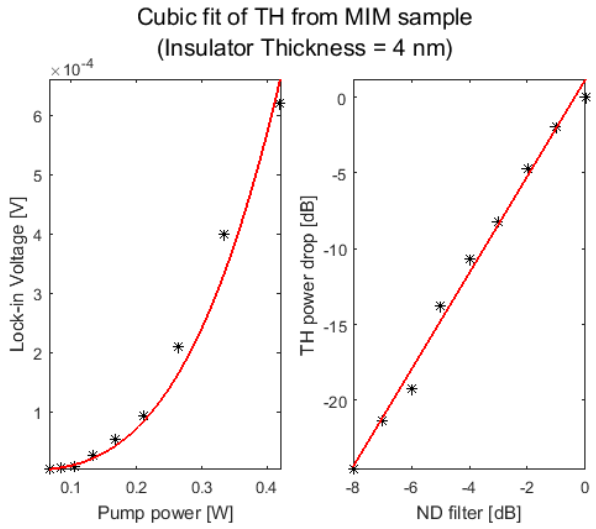


Figure 5.48: Cubic dependence of TH on FF intensity. Sample is a $4nm$ ALD coated Au surface with NPs.

ALD thickness : $10nm$

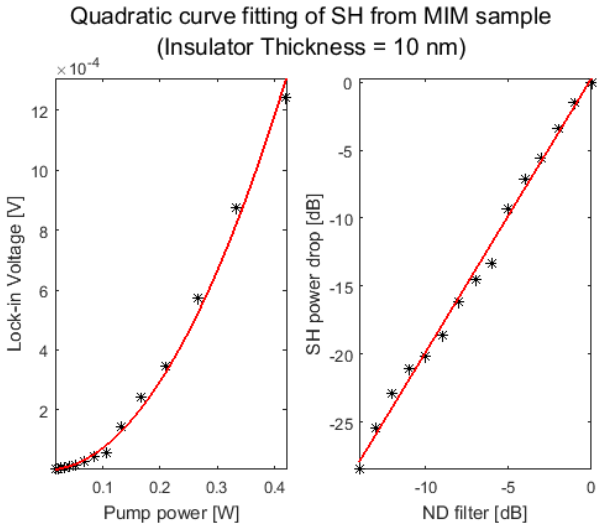


Figure 5.49: Quadrature dependence of SH on FF intensity. Sample is a $10nm$ *ALD coated Au surface with NPs.*

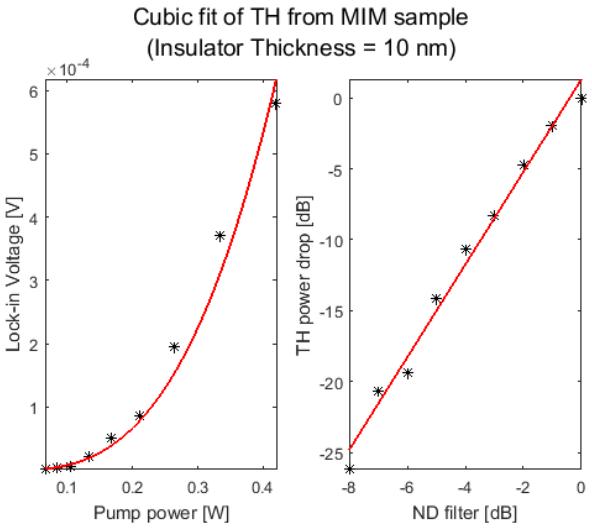


Figure 5.50: Cubic dependence of TH on FF intensity. Sample is a $10nm$ *ALD coated Au surface with NPs.*

ALD thickness : $20nm$

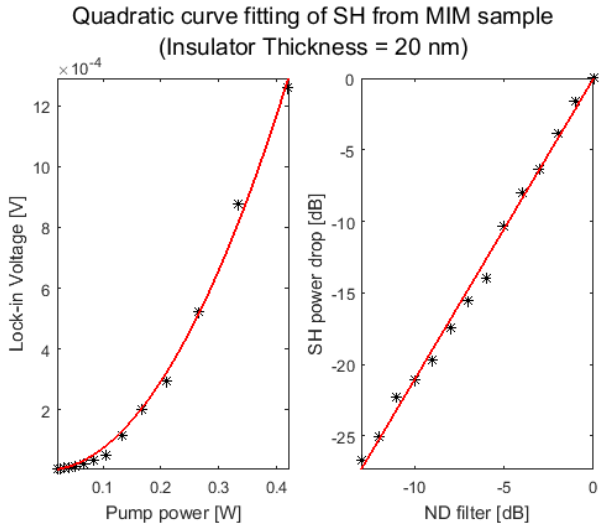


Figure 5.51: Quadrature dependence of SH on FF intensity. Sample is a $20nm$ *ALD coated Au surface with NPs.*

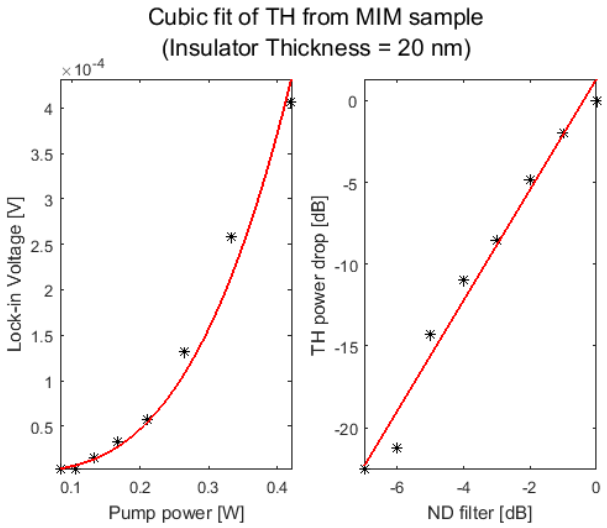


Figure 5.52: Cubic dependence of TH on FF intensity. Sample is a $20nm$ *ALD coated Au surface with NPs.*

The quadratic and cubic dependence of signal on FF intensity (see Figure 5.33 to 5.52) confirms that they are SH and TH signal respectively generated on the sample.

5.6 Result analysis:

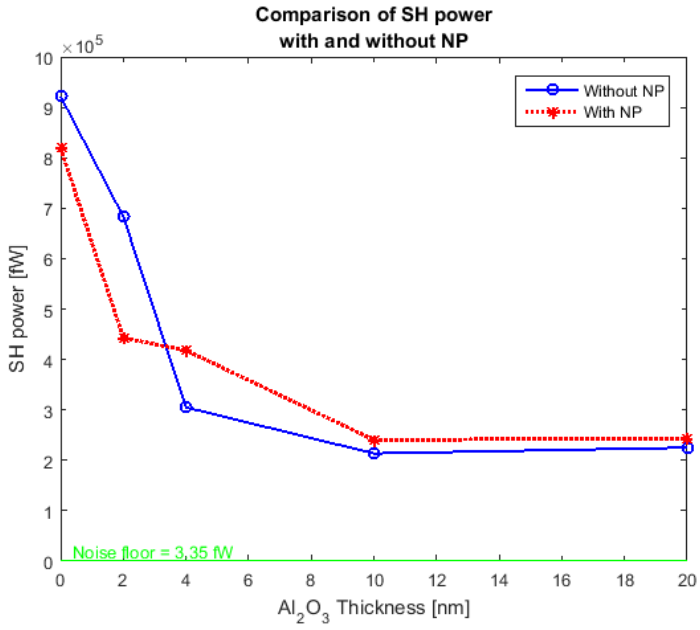


Figure 5.53: Comparison of SH from MI and MIM interface of varying ALD thicknesses.

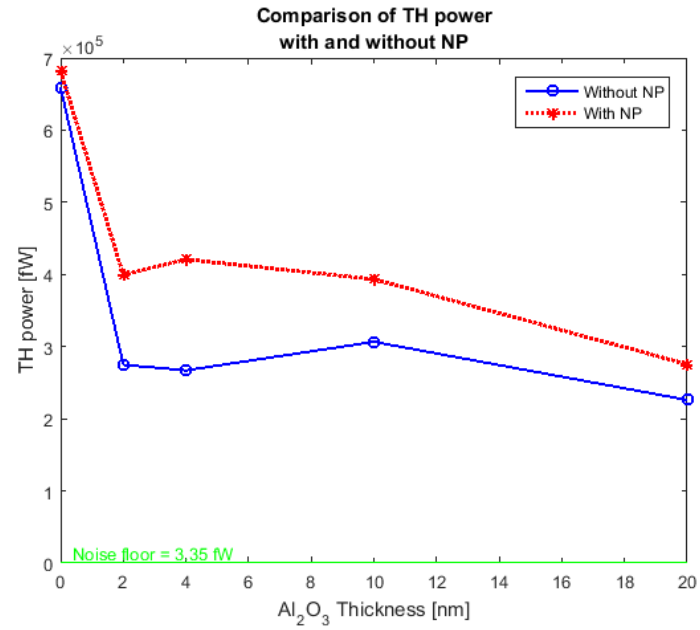


Figure 5.54: Comparison of TH from MI and MIM interface of varying ALD thicknesses.

The SH and TH signal strength with respect to Al_2O_3 film thickness is presented in Figure 5.53 and 5.54 respectively. The important conclusions derived from the above figures will discussed in the following.

- **Higher levels of SH/TH signal:**

  The SH signal from plain Au is 3 order of magnitude larger then what was expected (expected levels are discussed in Section 5.2.4).

- **A novel SHG technique to measure surface coverage of MI samples:**

  The signal strength reduces gradually for the first few nanometers ($\approx 4nm$) of Al_2O_3 thickness. For even higher thicknesses the signal remains unchanged. The amount of reduction in signal strength for higher thicknesses of ALD layer is around 40% (see Figure 5.53) compared to signal from bare Au surface. From preliminary simulation of SHG [†] from bare Au surface and Al_2O_3 deposited Au surface predicts that the SH signal strength reduction for $Al_2O_3 - Au$ interface and the percentage of drop with respect to plain Au is similar to the experimentally detected value. But, interestingly, the amount of drop in signal strength in the simulation model is not dependent on the thickness of Al_2O_3 film. As long as there is a uniform film of Al_2O_3 (it could be as thin as possible), the signal strength drops to a constant value and remains there. Since we saw a gradual drop for first few nanometers of thickness, our initial supposition was that the Al_2O_3 failed to completely cover the Au surface for first few ALD cycles. This happens because it is hard to start nucleation on metal surfaces, due to lack of hydroxyl group [27, 26].

[†]The simulation was done by Dr. Domenico de Ceglia using FEM and Dr. Michael Scalora using FDTD. Results were shared through private communication.

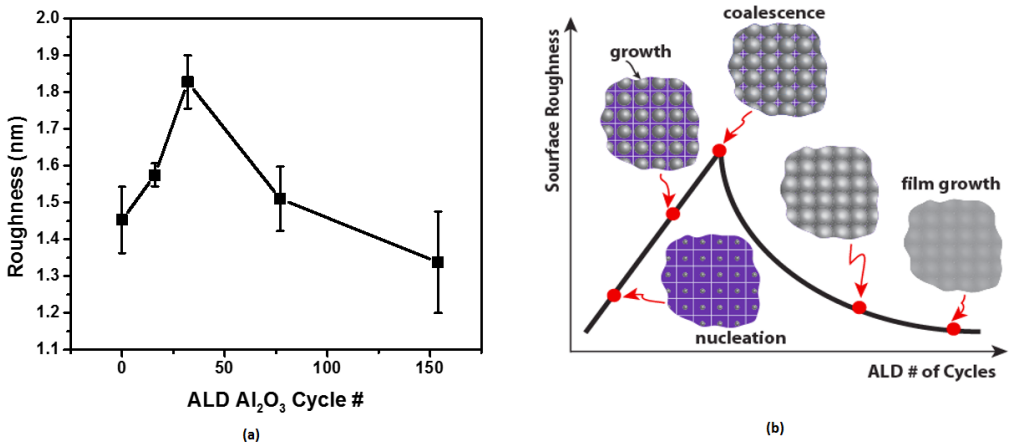


Figure 5.55: Nucleation-growth model of ALD layers: (a) The experimental measurement of surface roughness (using AFM) vs no. of ALD cycle,‡ (b) The theoretical model showing the transition between nucleation, coalescence and film formation.

So, few atomic forced microscopy (AFM) experiments ‡ were done to confirm the nucleation-growth model of oxides on metal surface using ALD. Figure 5.55(a) suggests an increase in surface roughness as the nucleation starts. At this point, the surface coverage of metal oxide is patchy. So, the SH from the pure metal surface and from the MI interface averages out to a lower value. As the number ALD cycles are increased, the patchy nucleation sites start to coalesce, and the SH from the MI interface dominates. Further increase in ALD cycles would completely cover the metal surface and form a film of oxide layer. At this point, the SH is generated only at the MI interface. Increasing the number of ALD cycles even further would have no effect on SHG from the MI interface. Figure 5.56 shows a strong correlation between the signal strength of SH from MI samples and the surface coverage of MI samples based on nucleation growth model and actual roughness measurement using AFM.

‡AFM experiments were done by Zhengning gao from Washington University in St. Louis and shared through private communication

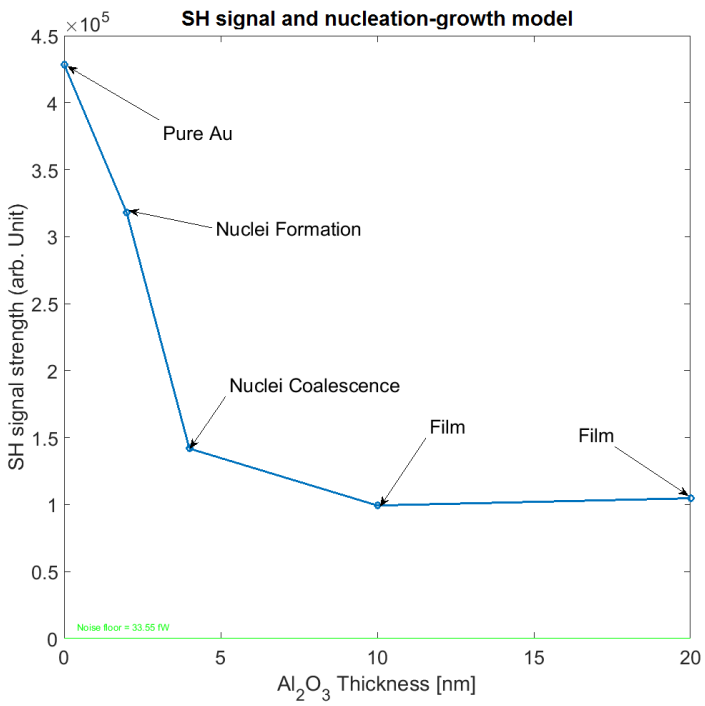


Figure 5.56: Dependence of SHG on surface coverage of MI samples.

- **Increase of TH signal from MIM sample:**

  There was a definite increase in TH signal after the samples were spin-coated with Au nanoparticles (see Figure 5.54). Part of the reason behind this increment could be quantum tunneling. But the insulator thickness should be subnanometer for the quantum tunneling effect to dominate. To reach to a firm conclusion, we need to increase the surface coverage and thickness uniformity, even for first few cycles of ALD.

5.7 Future work

Our future work can progress in multiple directions. Based upon our initial findings, we are interested to look through the following possibilities:

- The Au samples that we used were not ultra-flat surface. In future, we intend to use ultra-flat Au surfaces to cancel any dependency of our experimental results on Au surface roughness.

- We discovered a correlation between the SHG and surface coverage. This phenomenon can be further developed into a novel technique to measure surface coverage non-invasively. The existing technique is to measure the surface roughness through AFM. But using SH to measure the surface coverage would be much cheaper.

- We would like to improve our simulation to take into account the effects of surface coverage ratio.

- Techniques to improve surface coverage is also of great interest to us. A suggestion to achieve that goal would be to treat the sample with UV ray and O_3 gas before ALD. It has been observed that plain Au is actually hydrophilic [32] and even a small amount of carbonaceous deposit on Au surface kills the hydrophillicity. So treating the sample with O_3 would cause the H_2O and the precursors to spread out on the sample and deposit a more uniform layer of oxide. Our final goal is to design MIM interface with subnanometer thickness to enhance the tunneling and produce tunneling induced nonlinear radiation.

CHAPTER VI

CONCLUSION

MIM junctions (when junction size is smaller than photon wavelength) reveal Einstein's quantum description of photon as well as Maxwell's classical description of EM radiation. The juxtaposition of these two worlds makes MIM junctions key to many novel applications in the fields of energy harvesting, high speed communication, sensors and biomedical engineering. For example, a biased MIM can rectify electro-magnetic field [Equation 4.33]. Active devices, such as rectennas (i.e. rectifying antennas) [33], can be designed, as a new form of high efficiency solar energy harvesting device using MIM [34]. It also shows promise in the design of passive devices, such as detectors and modulators. For example, MIM junction can be used as long wavelength photon detectors [23, 35]. On the other end of the spectrum i.e. for short wavelength, high frequency modulators can be designed with MIM junctions [6]. Also, the field enhancement inside the junction may be used in sensing application. With all the prospective applications in mind, experiments on MIM seem very crucial. Experimental validation of QCT, also have a theoretical significance. Validation of QCT would mean the substantiation of transfer Hamiltonian approach for tunnel junctions consisting of regular metals. Our research will, eventually, answer the following critical questions:

- Can we unequivocally attribute the observations of the scattered nonlinear EM radiation to the quantum tunneling effect?

- What is the effect of limited field enhancement (or quenching) on other nonlinear effects, such as Raman scattering? The quantum effect will place ultimate limits on the optimal design of molecular sensing by methods, such as Raman scattering.

- What are the operational properties of MIM diode for energy harvesting? We have an exciting opportunity to do experiments on new MIM devices fabricated at WUStL. The first experiments will be done on nanorod MIM devices. The experiments to measure the efficiency of the devices will be of intense interest in the field.

- Can SH from metal-insulator interface be used as a handle to measure surface coverage and uniformity for the first few layers of insulation? This could be a breakthrough experiment on surface characterization of MI interfaces.

BIBLIOGRAPHY

[1] J. Bardeen, "Tunnelling from a many-particle point of view," *Physical Review Letters*, vol. 6, no. 2, p. 57, 1961.

[2] M. H. Cohen, L. M. Falicov, and J. C. Phillips, "Superconductive tunneling," *Physical Review Letters*, vol. 8, no. 8, p. 316, 1962.

[3] J. Bardeen, "Tunneling into superconductors," *Physical Review Letters*, vol. 9, no. 4, p. 147, 1962.

[4] M. Heiblum, S. Wang, J. Whinnery, and T. Gustafson, "Characteristics of integrated mom junctions at dc and at optical frequencies," *IEEE Journal of Quantum Electronics*, vol. 14, no. 3, pp. 159–169, 1978.

[5] P. Tien and J. Gordon, "Multiphoton process observed in the interaction of microwave fields with the tunneling between superconductor films," *Physical Review*, vol. 129, no. 2, p. 647, 1963.

[6] J. Tucker, "Quantum limited detection in tunnel junction mixers," *IEEE Journal of Quantum Electronics*, vol. 15, no. 11, pp. 1234–1258, 1979.

[7] A. Dayem and R. Martin, "Quantum interaction of microwave radiation with tunneling between superconductors," *Physical Review Letters*, vol. 8, no. 6, p. 246, 1962.

[8] E. O. Kane, *Basic Concepts of Tunneling*. Boston, MA: Springer US, 1969, pp. 1–11.

[9] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics*, ser. Quantum Mechanics. Wiley, 1991, no. v. 1. [Online]. Available: https://books.google.com/books?id=iHcpAQAAMAAJ

[10] J. W. Haus, D. de Ceglia, M. A. Vincenti, and M. Scalora, "Quantum conductivity for metal–insulator–metal nanostructures," *JOSA B*, vol. 31, no. 2, pp. 259–269, 2014.

[11] G. Wysin, "Probability current and current operators in quantum mechanics," *Department of Physics, Kansas State University, Manhattan, KS*, vol. 66506, p. 2601.

[12] S. Grover and G. Moddel, "Engineering the current–voltage characteristics of metal–insulator–metal diodes using double-insulator tunnel barriers," *Solid-State Electronics*, vol. 67, no. 1, pp. 94–99, 2012.

[13] J. G. Simmons, "Electric tunnel effect between dissimilar electrodes separated by a thin insulating film," *Journal of applied physics*, vol. 34, no. 9, pp. 2581–2590, 1963.

[14] ——, "Potential barriers and emission-limited current flow between closely spaced parallel metal electrodes," *Journal of Applied Physics*, vol. 35, no. 8, pp. 2472–2481, 1964.

[15] A. Ghatak and S. Lokanathan, *Quantum mechanics*. MacMillan Co. of India, 1975. [Online]. Available: https://books.google.co.in/books?id=CdZUAAAAYAAJ

[16] L. Li, "Study of metal-insulator-metal diodes for photodetection," Ph.D. dissertation, University of Dayton, 2013.

[17] I. Giaever and K. Megerle, "Study of superconductors by electron tunneling," *Physical Review*, vol. 122, no. 4, p. 1101, 1961.

[18] B. Josephson, "Possible new effects in superconductive tunnelling," *Physics Letters*, vol. 1, no. 7, pp. 251 – 253, 1962. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0031916362913690

[19] S. Shapiro, "Josephson currents in superconducting tunneling: The effect of microwaves and other observations," *Physical Review Letters*, vol. 11, no. 2, p. 80, 1963.

[20] P. C. Martin and J. Schwinger, "Theory of many-particle systems. i," *Physical Review*, vol. 115, no. 6, p. 1342, 1959.

[21] L. P. Kadanoff and P. C. Martin, "Theory of many-particle systems. ii. superconductivity," *Physical Review*, vol. 124, no. 3, p. 670, 1961.

[22] D. Rogovin and D. Scalapino, "Fluctuation phenomena in tunnel junctions," *Annals of Physics*, vol. 86, no. 1, pp. 1–90, 1974.

[23] J. Tucker and M. Millea, "Photon detection in nonlinear tunneling devices," *Applied Physics Letters*, vol. 33, no. 7, pp. 611–613, 1978.

[24] V. Ambegaokar and A. Baratoff, "Tunneling between superconductors," *Physical Review Letters*, vol. 10, no. 11, p. 486, 1963.

[25] J. W. Haus, D. de Ceglia, M. A. Vincenti, and M. Scalora, "Nonlinear quantum tunneling effects in nanoplasmonic environments: two-photon absorption and harmonic generation," *JOSA B*, vol. 31, no. 6, pp. A13–A19, 2014.

[26] A. J. Elliot, G. Malek, L. Wille, R. Lu, S. Han, J. Z. Wu, J. Talvacchio, and R. M. Lewis, "Probing the nucleation of al_2o_3 in atomic layer deposition on aluminum for ultrathin tunneling barriers in josephson junctions," *IEEE Transactions on Applied Superconductivity*, vol. 23, no. 3, pp. 1 101 405–1 101 405, 2013.

[27] M. Groner, J. Elam, F. Fabreguette, and S. George, "Electrical characterization of thin al2o3 films grown by atomic layer deposition on silicon and various metal substrates," *Thin Solid Films*, vol. 413, no. 12, pp. 186 – 197, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0040609002004388

[28] I. H. Malitson, "Interspecimen comparison of the refractive index of fused silica∗,," *J. Opt. Soc. Am.*, vol. 55, no. 10, pp. 1205–1209, Oct 1965. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=josa-55-10-1205

[29] B. Ramsay, O. Koppius, and E. Cleveland, "The prism and the theory of optical resolution," *JOSA*, vol. 30, no. 9, pp. 439–444, 1940.

[30] M. C. Larciprete, A. Belardini, M. Cappeddu, D. De Ceglia, M. Centini, E. Fazio, C. Sibilia, M. Bloemer, and M. Scalora, "Second-harmonic generation from metallodielectric multilayer photonic-band-gap structures," *Physical Review A*, vol. 77, no. 1, p. 013809, 2008.

[31] V. Mackowiak, J. Peupelmann, Y. Ma, and A. Gorges, "Nep–noise equivalent power." [Online]. Available: https://www.thorlabs.com/images/TabImages/NoiseEquivalent_Power_White_Paper.pdf

[32] T. Smith, "The hydrophilic nature of a clean gold surface," *Journal of Colloid and Interface Science*, vol. 75, no. 1, pp. 51–55, 1980.

[33] S. Grover, S. Joshi, and G. Moddel, "Quantum theory of operation for rectenna solar cells," *Journal of Physics D: Applied Physics*, vol. 46, no. 13, p. 135106, 2013.

[34] S. Joshi and G. Moddel, "Efficiency limits of rectenna solar cells: Theory of broadband photon-assisted tunneling," *Applied Physics Letters*, vol. 102, no. 8, p. 083901, 2013.

[35] G. Elchinger, A. Sanchez, C. Davis Jr, and A. Javan, "Mechanism of detection of radiation in a high-speed metal-metal oxide-metal junction in the visible region and at longer wavelengths," *Journal of Applied Physics*, vol. 47, no. 2, pp. 591–594, 1976.

[36] "The photoelectric effect without photons," *Center for Theoretical Studies, University of Miami, Coral Gables, Florida*, Feb. 1968.

[37] E. N. Koukaras, "Fermis golden rule," *Physics Department, University of Patras*.

APPENDIX A

FERMI'S GOLDEN RULE AND TRANSITION RATE

Bohr's atomic model can predict the discreteness of atomic spectra by assuming the wave nature of electrons. But it can not explain the relative brightness differences of spectral lines. **Fermi's golden rule** is a form of **time-dependent perturbation theory** that predicts the transition rate of particles from one state to another and explains why the transition between some states is higher than the others, which, in turn, gives the reason behind the relative brightness differences. In fact, photoelectric effect (transition of electrons from conduction band to vacuum) in metals can also be described using Fermi's golden rule [36]. As discussed in Chapter III, Bardeen slightly tweaked [1, 8] the idea behind Fermi's golden rule to account for tunneling current and adequately explained the linear relationship between energy density of states, ρ_e and tunneling current density, $j_{sc}(x)$ in superconducting tunnel junctions. This linear relationship was utilized by Giaever to explain his experiments on SIS tunneling junctions [17]. The usage of Fermi's golden rule can be extended to MIM tunnel junctions, as well. We will now discuss the derivation and implications of Fermi's golden rule.

A.1 Time-dependent perturbation theory

The derivation follows very closely the terms and notation of Koukaras's article [37] on Fermi's golden rule. Let's assume, a system Hamiltonian is perturbed with a time-dependent potential,

$H_T(t)^\dagger$ and the system undergoes a transition from initial state $|i\rangle$ to a final state $|f\rangle$. So, the new Hamiltonian can be written as,

$$\hat{H} = \hat{H}_0 + \hat{H}_T(t), \tag{A.1}$$

where, $\hat{H}_0$ is the unperturbed system Hamiltonian. The eigenvalues and eigenstates of the uncoupled system are E_n and $|\phi_n\rangle$ respectively and n = any positive integer number. According to Schödinger's equation i.e. $\hat{H}\psi = i\hbar\frac{\partial\psi}{\partial t}$, the time-evolution of $|\phi_n\rangle$ in the unperturbed system is just a phase change in the Hilbert space:

$$|\phi_n(t)\rangle = |\phi_n\rangle\, e^{-iE_nt/\hbar}. \tag{A.2}$$

Notice that, addition of any external voltage also changes the phase only. The generalized wavefunction of the system is a linear combination of eigenvectors with constant weights c_n, i.e.

$$|\psi(t)\rangle = \sum_n c_n\, |\phi_n(t)\rangle = \sum_n c_n\, |\phi_n\rangle\, e^{-iE_nt/\hbar}. \tag{A.3}$$

But, when the system is perturbed, the generalized wavefunction does not remain a linear combination of eigenstates anymore, rather, it becomes a 'mixed' state, i.e.

$$|\psi(t)\rangle = \sum_n c_n(t)\, |\phi_n\rangle\, e^{-iE_nt/\hbar}, \tag{A.4}$$

where, $c_n(t)$ are not expansion coefficients, instead, $|c_n(t)|^2$ represents the probabilities of finding the system in state $|\phi_n\rangle$. So, introducing perturbation in a system essentially changes 'pure' states to 'mixed' states. When a mixed state, $|\psi(t)\rangle$ (from Equation A.4) is operated by the new perturbed Hamiltonian, $\hat{H}$ (from Equation A.1), we get,

$$\hat{H}\,|\psi(t)\rangle = [\hat{H}_0 + \hat{H}_T(t)]\,|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle\,,$$

[†]Throughout this thesis, $H_T(t)$ represents the 'transfer Hamiltonian' that accounts for tunneling in a coupled system. To avoid any confusion while transferring the concepts of Fermi's golden rule to the tunneling problem, we used the same notation here.

$$\Rightarrow \hat{H}_0 \sum_m c_m(t) \left|\phi_m\right\rangle e^{-iE_m t/\hbar} + \hat{H}_T(t) \sum_m c_m(t) \left|\phi_m\right\rangle e^{-iE_m t/\hbar} = i\hbar \frac{\partial}{\partial t} \sum_m c_m(t) \left|\phi_m\right\rangle e^{-iE_m t/\hbar},$$

$$\Rightarrow \sum_m c_m(t) E_k \left|\phi_m\right\rangle e^{-iE_m t/\hbar} + \sum_m c_m(t) \hat{H}_T(t) \left|\phi_m\right\rangle e^{-iE_m t/\hbar} =$$

$$i\hbar \sum_m \frac{\partial c_m(t)}{\partial t} \left|\phi_m\right\rangle e^{-iE_m t/\hbar} + i\hbar \sum_m c_m(t) \left|\phi_m\right\rangle \left(-\frac{iE_m}{\hbar}\right) e^{-iE_m t/\hbar}. \tag{A.5}$$

By multiplying Equation A.5 with $\langle\phi_n|$ from the left, we get,

$$E_n c_n(t) e^{-iE_n t/\hbar} + \sum_m c_m(t) T_{nm}(t) e^{-iE_m t/\hbar} = i\hbar \frac{\partial c_n(t)}{\partial t} e^{-iE_n t/\hbar} + E_n c_n(t) e^{-iE_n t/\hbar},$$

$$\Rightarrow \frac{\partial c_n(t)}{\partial t} = \frac{1}{i\hbar} \sum_m c_m(t) T_{nm}(t) e^{-i\omega_{nm} t}, \tag{A.6}$$

where, T_{nm} is the coupling energy between states $|\phi_n\rangle$ and $|\phi_m\rangle$, i.e.

$$T_{nm} = \left\langle\phi_n\right| \hat{H}_T(t) \left|\phi_m\right\rangle,$$

$$= \int_{-\infty}^{\infty} {\phi_n}^*(x) H_T(t) \phi_m(x) dx, \tag{A.7}$$

and, $e^{-i\omega_{nm} t}$ is the phase of coupling energy, when,

$$\omega_{nm} = (E_n - E_m)/\hbar. \tag{A.8}$$

ω_{nm} can also be viewed as the perturbation frequency introduced to the system through external sources, V_{ext}. These external sources can either be a high frequency photon voltage, $V_{ph} = \frac{\hbar\omega_{nm}}{q_e}$ (where, $q_e =$ charge of electron) or a low frequency (almost DC) voltage, V_{dc} or both. The external sources modulate the Fermi level of the system and change the phase of coupling energy, $e^{-i\omega_{nm} t}$ between any two states. For our problem of tunneling current calculation, Bloch states from two separate systems (left, $|l\rangle$ and right, $|r\rangle$) are coupled through tunneling. Equation A.6 becomes very useful to establish the dependence of tunneling current on external voltage in the framework of perturbation theory. This is discussed in Chapter III in details. For now, let's develop the rest of the theory and, finally, derive Fermi's golden rule.

To proceed, we need to assume that, at the beginning the particle was at initial state $|i\rangle$. So, $c_i = 1$

and $c_n = 0$ for all $n \neq i$. The perturbation of the system is very weak and applied for a short period of time, so that, the system reacts adiabatically to the change. From Equation A.6, we get,

$$\frac{\partial c_n(t)}{\partial t} = \frac{1}{i\hbar} c_i(t) T_{ni}(t) e^{-i\omega_{ni}t}. \tag{A.9}$$

For a final state of $|f\rangle$, by integrating Equation A.9 with respect to time, t, we get $c_f(t)$, i.e.

$$\boxed{c_f(t) = \frac{1}{i\hbar} \int_0^t T_{fi}(t') e^{-i\omega_{fi}t'} dt'.} \tag{A.10}$$

A.2 Harmonic perturbation

Now, we will introduce a harmonic perturbation voltage with high (optical) frequency, ω. This could be done physically by illuminating polarized light on the system. So, the perturbation looks like,

$$T(t) = 2T_0 cos(\omega t) = T_0(e^{i\omega t} + e^{-i\omega t}). \tag{A.11}$$

The coupling matrix element due to the harmonic perturbation is, $T_{fi}(t) = \langle \phi_f | \, T(t) \, | \phi_i \rangle = |T_{fi}|(e^{i\omega t} + e^{-i\omega t})$. So, Equation A.10 can be written as,

$$\begin{aligned} c_f(t) &= \frac{|T_{fi}|}{i\hbar} \int_0^t \left(e^{i\omega t'} + e^{-i\omega t'} \right) e^{i\omega_{fi}t'} dt', \\ &= \frac{|T_{fi}|}{i\hbar} \left\{ \frac{e^{(\omega_{fi}+\omega)t} - 1}{i(\omega_{fi} + \omega)} + \frac{e^{(\omega_{fi}-\omega)t} - 1}{i(\omega_{fi} - \omega)} \right\}. \end{aligned} \tag{A.12}$$

The 'Propagator', $\mathcal{P}$ is defined as the inner product of a particles final state $|\phi_f\rangle$ and current state $|\psi(t)\rangle$. Physically, propagator quantifies the probability of finding a particle at $|\phi_f\rangle$ when it started as $|\psi(t)\rangle$. Visually, this represents the overlap component between the states in Hilbert space.

$$\mathcal{P}_f(t) = |c_f(t)|^2 = |\, \langle \phi_f | \psi(t) \rangle \,|^2. \tag{A.13}$$

Since our particle started in $|\phi_i\rangle$, Equation A.13 can be written as:

$$\mathcal{P}_{fi}(t) = |c_f(t)|^2 = \frac{T_{fi}^2}{\hbar^2} \left| \frac{e^{(\omega_{fi}+\omega)t} - 1}{i(\omega_{fi} + \omega)} + \frac{e^{(\omega_{fi}-\omega)t} - 1}{i(\omega_{fi} - \omega)} \right|^2. \tag{A.14}$$

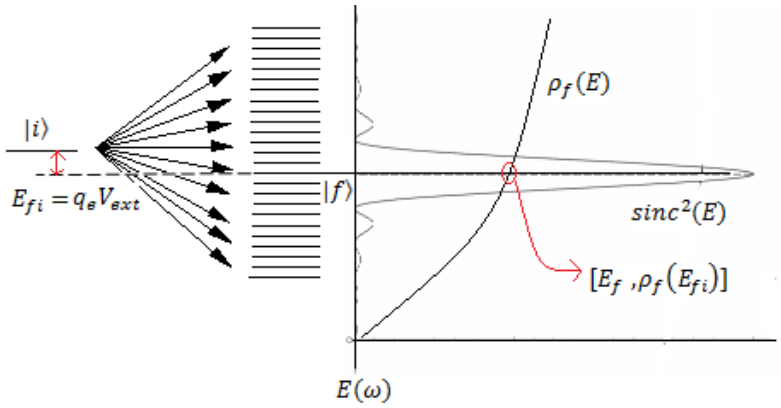


Figure A.1: Transition of particle from $|i\rangle$ to $|f\rangle$

To proceed with the Equation A.14, we made an approximation that, $\omega \approx \omega_{fi}$ and ω is optical frequency, so, $|\omega - \omega_{fi}| << \omega$, and Equation A.14 becomes,

$$\begin{aligned}\mathcal{P}_{fi}(t) &= \frac{|T_{fi}|^2}{\hbar^2}\left|\frac{e^{(\omega_{fi}-\omega)t}-1}{i(\omega_{fi}-\omega)}\right|^2, \\ &= \frac{|T_{fi}|^2 t^2}{\hbar^2}\Big[sinc[(\omega_{fi}-\omega)t/2]\Big]^2. \end{aligned} \tag{A.15}$$

A.3 Harmonic perturbation in a continuum of final states

For a continuum of final states, with density of states $\rho_f(E)$, the propagator can be written as,

$$\begin{aligned}\mathcal{P} &= \int_{-\infty}^{\infty} \mathcal{P}_{fi}(t)\rho_f(E)dE, \\ &= \int_{-\infty}^{\infty} \frac{|T_{fi}|^2 t^2}{\hbar^2}\Big[sinc[(\omega_{fi}-\omega)t/2]\Big]^2 \rho_f(E)dE, \\ &= \int_{-\infty}^{\infty} \frac{|T_{fi}|^2 t^2}{\hbar^2}\rho_f(E)\Big[sinc[(\omega_{fi}-\omega)t/2]\Big]^2 \hbar d\omega, \\ &= \int_{-\infty}^{\infty} \frac{|T_{fi}|^2 t^2}{\hbar^2}\rho_f(E)\hbar\left(\frac{2}{t}\right)\Big[sinc(\omega_{fi}-\omega)\Big]^2 d\omega. \end{aligned} \tag{A.16}$$

Since the $sinc(\omega_{fi} - \omega)$ function has a sharp peak at $\omega = \omega_{fi}$ (see Figure A.1) and has very small bandwidth, $\Delta\omega$, i.e. $(\Delta\omega \approx \frac{4\pi}{t}) \gg \frac{T_{fi}}{\hbar}$, T_{fi} and $\rho_f(E)$ can be taken out of the integration as

a constant, and the rest of the integration i.e. $\int_{-\infty}^{\infty} sinc^2(\omega_{fi} - \omega)d\omega$ function can be written as $2\pi \int_{-\infty}^{\infty} \delta(\omega_{fi} - \omega)d\omega$. So, Equation A.16 can be written as,

$$\mathcal{P} = \frac{2\pi}{\hbar}|T_{fi}|^2 \rho_f(E_f)t \int_{-\infty}^{\infty} \delta(\omega_{fi} - \omega)d\omega. \tag{A.17}$$

The transition rate, Γ is the rate of change of the propagator, $\mathcal{P}$ with respect to time, t, i.e. $\Gamma = \frac{d}{dt}\mathcal{P}$. Therefore,

$$\boxed{\Gamma = \frac{2\pi}{\hbar}|T_{fi}|^2 \rho_f(E_f) \int_{-\infty}^{\infty} \delta(\omega_{fi} - \omega)d\omega.} \tag{A.18}$$

This equation is known as **Fermi's golden rule**.

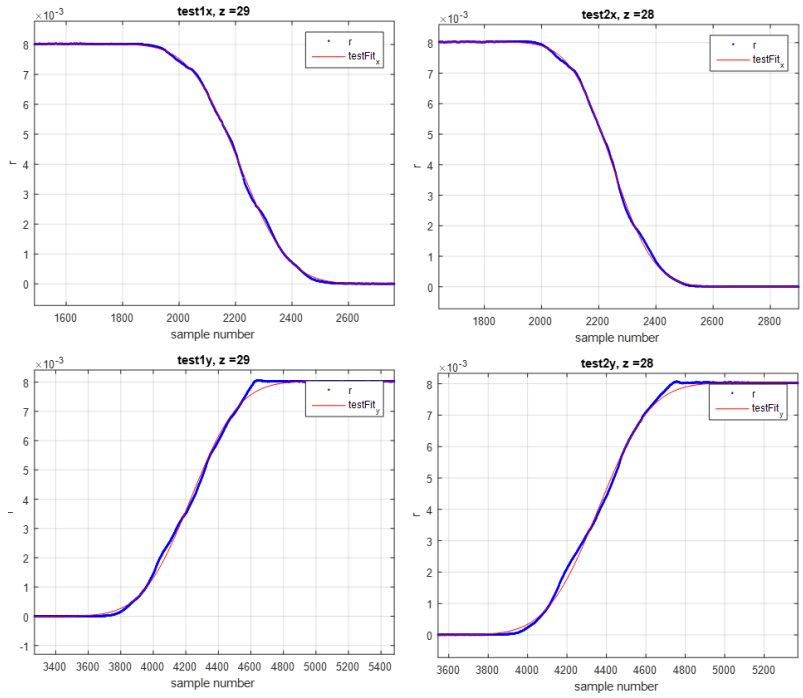
A.4 Summary

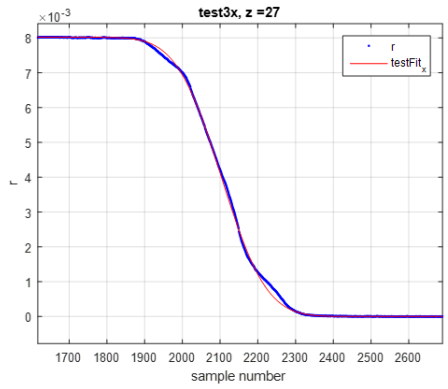
The most important take away in this chapter is that, a Hamiltonian perturbed with high frequency electric fields of photon triggers the transition of state of an electron. The dimension of rate of transition is the same as the dimension of probability current density, i.e. $[T^{-1}]$. In Chapter III we will prove that the transition rate, Γ can, indeed, be treated as the probability current density from each initial state to all the final states. But, the physical meaning of the term T_{fi}, which is needed to calculate Γ [Equation A.18], may still feel like a bit of a mystery. It will become clearer in the Chapter III when we discuss Bardeen's picture of tunneling through MIM (see Section 3.1). Having discussed Fermi's golden rule, it is, now, prudent to grab attention to the fact that, 'harmonic perturbation' is NOT the same as 'harmonic modulation'. Even though, both of them are introduced in the system Hamiltonian as time-dependent potential energy, the treatment of harmonic perturbation (as discussed in this chapter) is quite different from the treatment of harmonic modulation (see Section 4.1). For PAT theory, both of this world are put together (see Chapter V). In this theory, electric field of the photon modulates one system that causes perturbation in another. The results of PAT are then forwarded to QCT for further development.

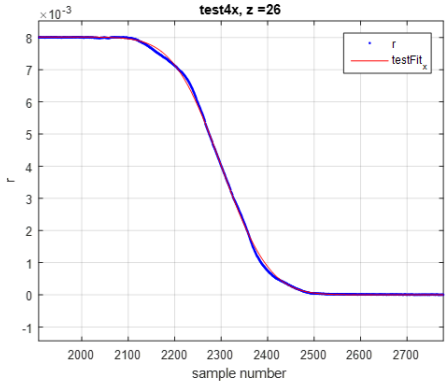
APPENDIX B

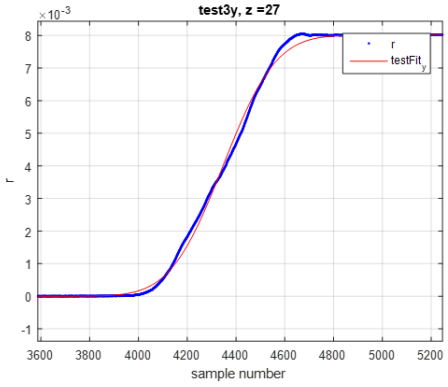
S CURVES FROM KNIFE EDGE EXPERIMENT

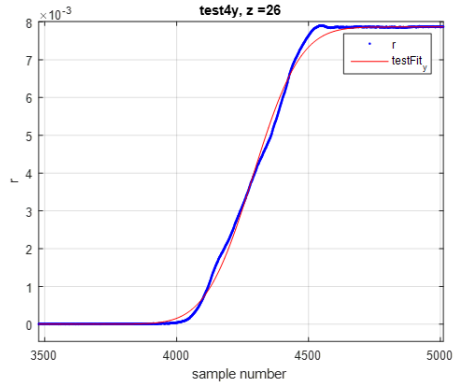
21 samples of knife edge data was taken along the Z axis. $testNx$ represents the N^{th} position on the Z axis for knife edge moving along the X axis. $testNy$ represents the N^{th} position on the Z axis for knife edge moving along the Y axis. S curve was fit for each data set.

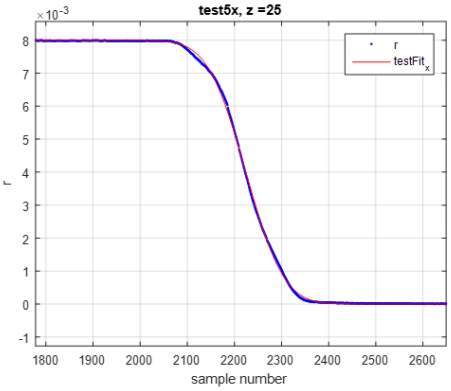


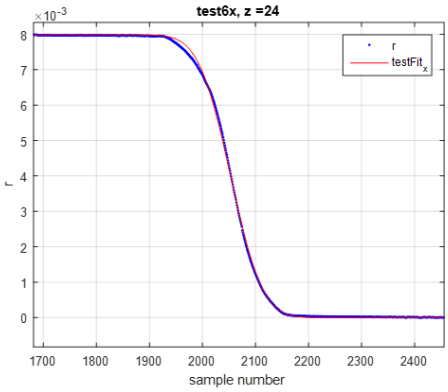

test3x, z =27
r
testFit_x
sample number

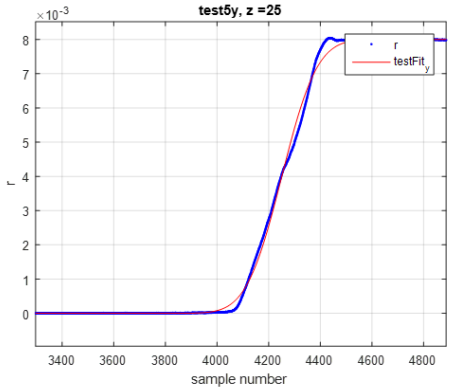

test4x, z =26
r
testFit_x
sample number

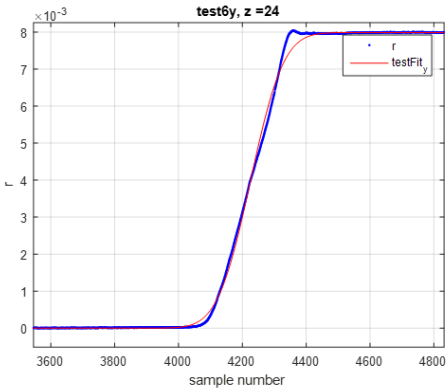

test3y, z =27
r
testFit_y
sample number

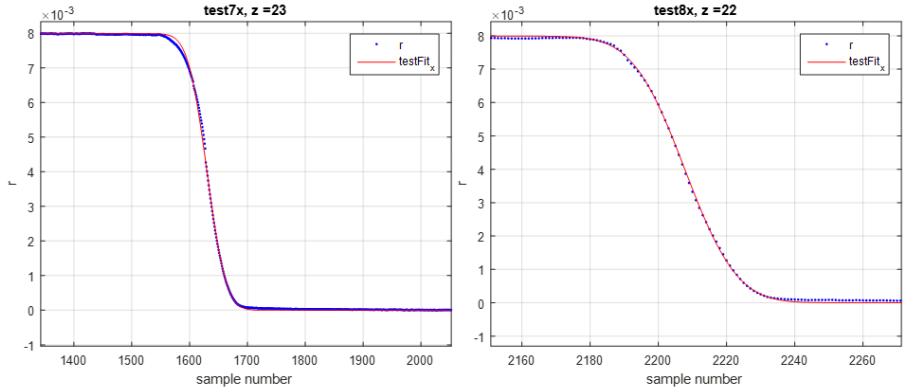

test4y, z =26
r
testFit_y
sample number

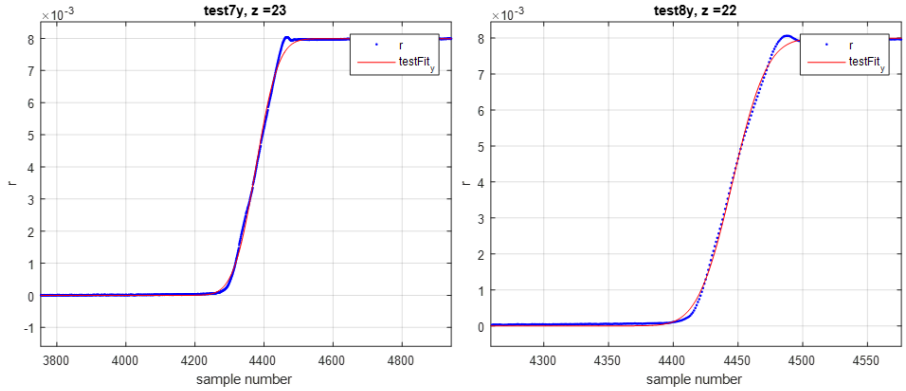

test5x, z =25
r
testFit_x
sample number

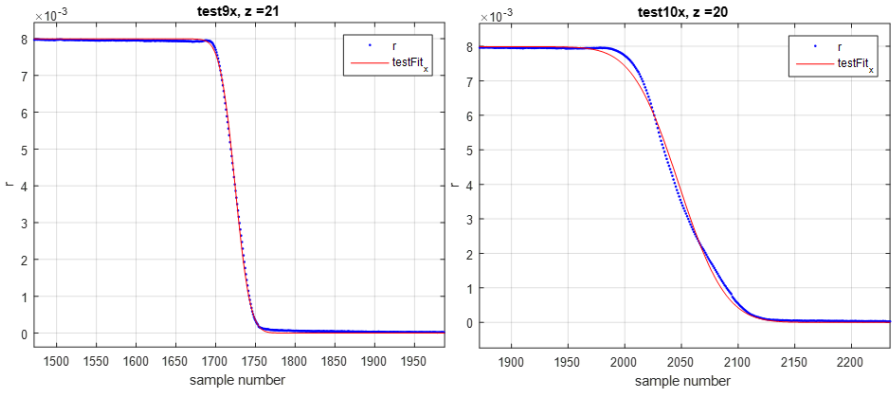

test6x, z =24
r
testFit_x
sample number

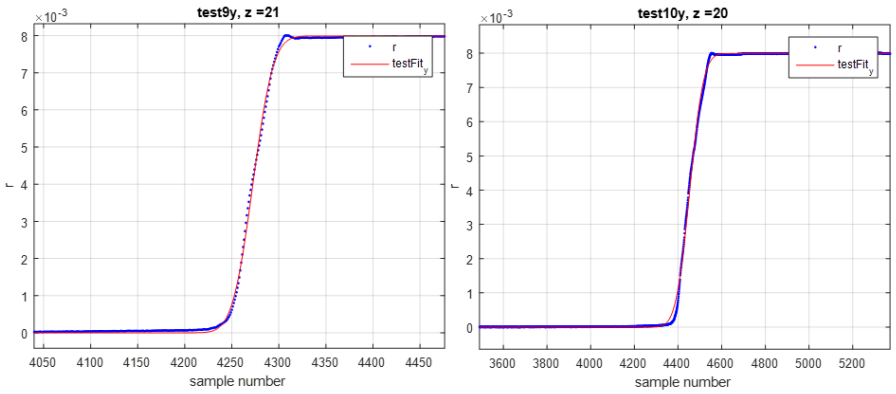

test5y, z =25
r
testFit_y
sample number

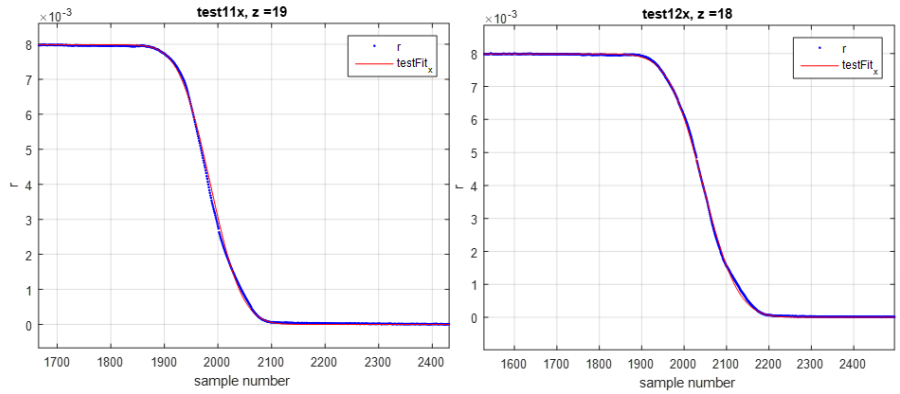

test6y, z =24
r
testFit_y
sample number

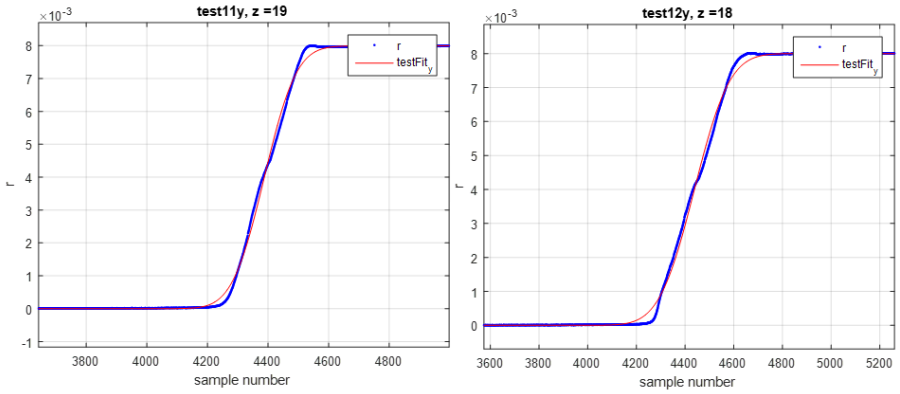

test7x, z =23
test8x, z =22
r
testFit_x
sample number

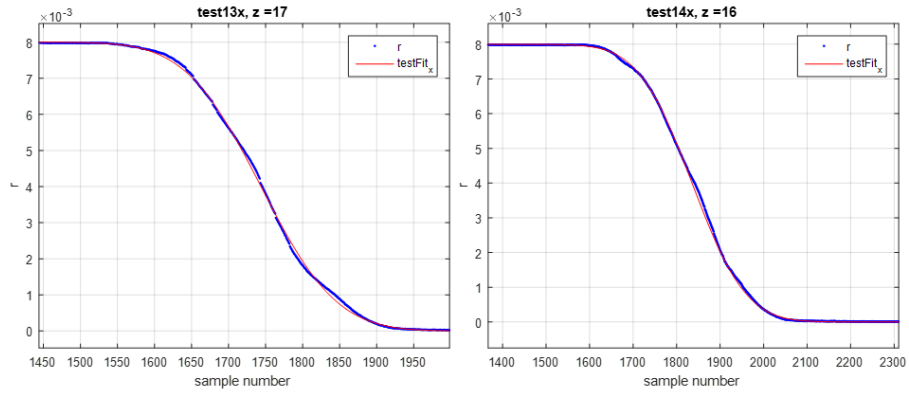

test7y, z =23
test8y, z =22
r
testFit_y
sample number

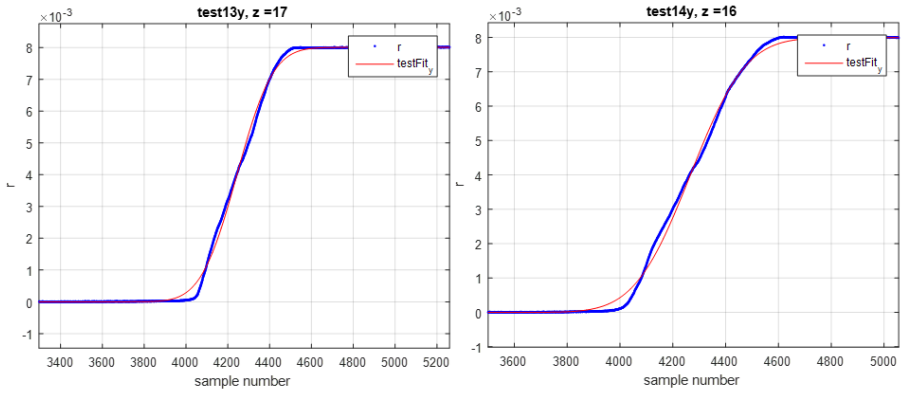

test9x, z =21
test10x, z =20
r
testFit_x
sample number

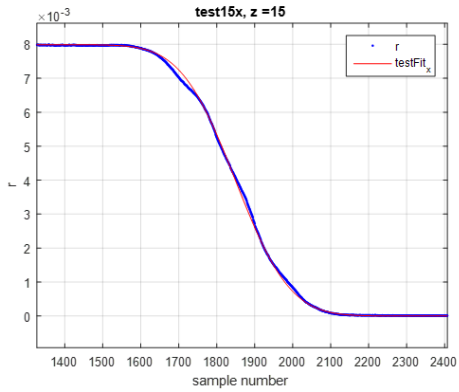

test9y, z =21
test10y, z =20
r
testFit_y
sample number

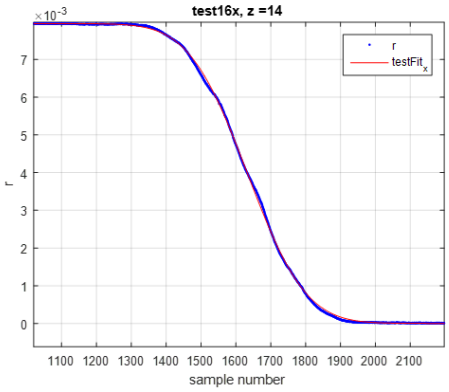

test11x, z =19
test12x, z =18
r
testFit_x
sample number

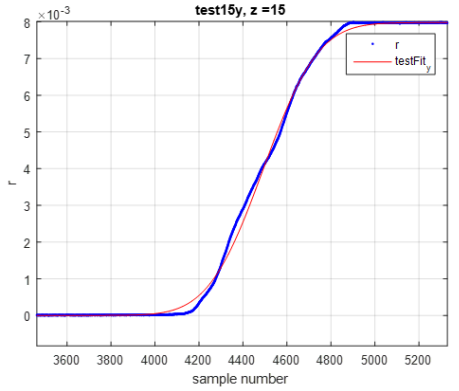

test11y, z =19
test12y, z =18
r
testFit_y
sample number

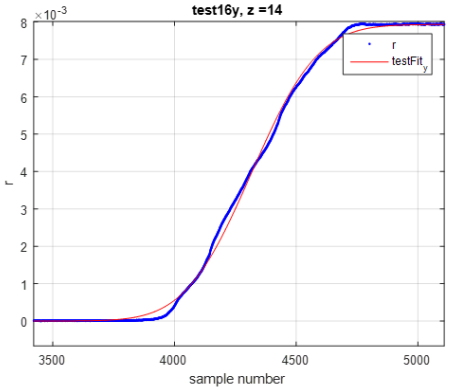

test13x, z =17
test14x, z =16
r
testFit_x
sample number

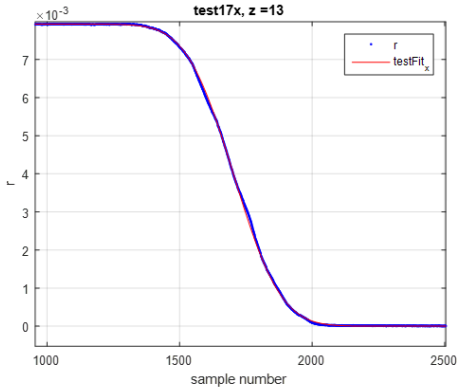

test13y, z =17
test14y, z =16
r
testFit_y
sample number

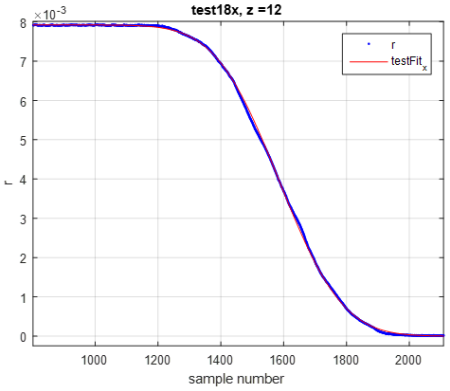

test15x, z =15
×10^-3
r
testFit_x
sample number

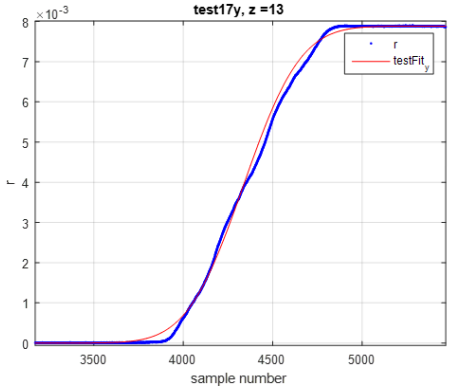

test16x, z =14
×10^-3
r
testFit_x
sample number

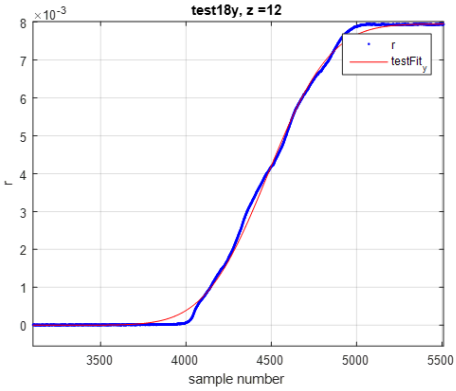

test15y, z =15
×10^-3
r
testFit_y
sample number

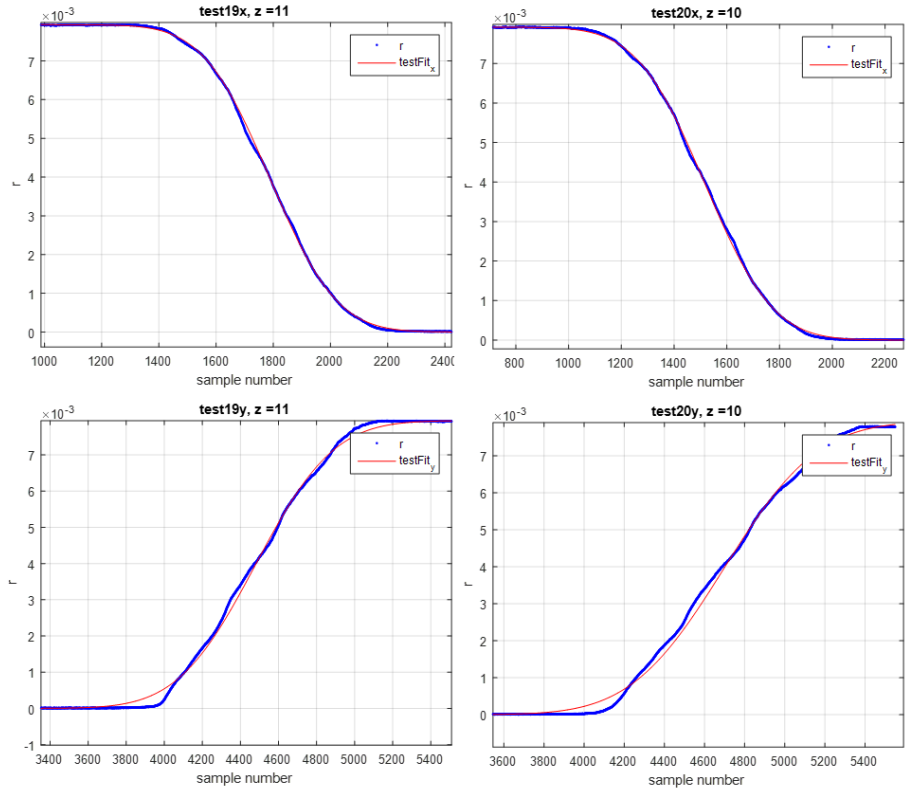

test16y, z =14
×10^-3
r
testFit_y
sample number


test17x, z =13
×10^-3
r
testFit_x
sample number


test18x, z =12
×10^-3
r
testFit_x
sample number


test17y, z =13
×10^-3
r
testFit_y
sample number


test18y, z =12
×10^-3
r
testFit_y
sample number


test19x, z =11
test20x, z =10
test19y, z =11
test20y, z =10
r
testFit_x
testFit_y
sample number
×10^-3