The dissertation of Corey J. Cochrane has been reviewed* and approved by the following:

Patrick M. Lenahan
Distinguished Professor of Engineering Science and Mechanics
Dissertation Adviser
Chair of Committee

S. Ashok
Professor of Engineering Science

Mark Horn
Professor of Engineering Science and Mechanics

Jerzy Ruzyillo
Professor of Electrical Engineering and Materials Science and Engineering

Judith A. Todd
P.B. Breneman Department Head
Head of the Department of Engineering Science and Mechanics

*Signatures are on file in the Graduate School.
Abstract

This work focuses on the development of new techniques for the study of spin dependent transport and trapping centers in fully processed micro and nanoelectronics. The first, and most interesting, technique offers a very low cost means to study spin dependent transport in microelectronics as an alternative to electrically detected magnetic resonance (EDMR). EDMR measurements generally require strong static magnetic fields, typically 3 kG or greater, and high frequency oscillating electromagnetic fields, typically 9 GHz or higher. In this work, it is demonstrated that large spin dependent recombination and tunneling signals can be detected in the absence of the oscillating electromagnetic field at zero magnetic field. The physics behind this technique is based upon the mixing of singlet and triplet energy states of the electron spin pairs involved in the spin dependent processes. In this study, we show that this technique can be applied to Si and SiC based devices. Theoretically, it can be applicable to devices of all material systems in which defects play a role in spin dependent transport, some of which include CdTe and GaN. Although the resolution of the $g$ value is sacrificed in this new measurement, the technique can detect electron-nuclear hyperfine interactions and possibly dipolar and exchange interactions. The technique also has great promise in microelectronic device reliability studies as it is directly applicable to time dependent dielectric breakdown in thin film dielectrics and bias temperature instabilities in transistors. Other applications of this new physics include self-calibrating magnetometers, spin based memories, quantum computation, and miniature EDMR spectrometers for wafer probing stations. The second technique involves the utilization of passage effects that arise when performing magnetic field modulation in EDMR. When certain conditions are met, the higher order harmonics of the spin dependent signal can contain much useful information; one of them being the fast passage signal. In this work, we designed a multiband virtual lock-in amplifier that can simultaneously demodulate any of these higher order harmonics. This tool has allowed for the identification of a very important recombination center in 4H SiC MOSFETs; the silicon vacancy. To the best of our knowledge, this was the first study that utilized passage effects for defect identification in EDMR. And finally, this work involves the development of an adaptive signal averaging technique that is capable of reducing the noise variance of a single scan by a factor of 10 or more which reduces the time of acquisition by the same amount. This technique is applicable to all methods in which signal averaging is utilized, some of which include medical imaging, electrocardiography, or electroencephalography.
# Table of Contents

List of Figures .............................................................................................................................................. vi  
List of Tables .............................................................................................................................................. xii  
Acknowledgements ..................................................................................................................................... xiii  

## CHAPTER 1 – INTRODUCTION ............................................................................................................... 1  
1.1 – Atomic Scale Defect Detection .......................................................................................... 1  
1.2 – Electrically Detected Magnetic Resonance ........................................................................ 2  
1.3 – Zero- & Low-Field Spin Dependent Transport .................................................................. 4  

## CHAPTER 2 – ELECTRON PARAMAGNETIC RESONANCE ............................................................... 5  
2.1 – Angular Momentum and Magnetic Moments due to Spin ................................................. 6  
2.2 – Magnetic Resonance of Electrons and Nuclei .................................................................... 11  
2.3 – Continuous Wave EPR Practice ....................................................................................... 22  
2.4 – The Bloch Equations ........................................................................................................ 25  
2.5 – The Rotating Reference Frame ........................................................................................ 31  
2.6 – Spin Orbit Coupling ......................................................................................................... 36  
2.7 – Hyperfine Interactions ...................................................................................................... 38  

## CHAPTER 3 – SPIN DEPENDENT TRANSPORT & ITS ELECTRICAL DETECTION ...................... 46  
3.1 – Spin Dependent Transport ............................................................................................... 46  
3.2 – Pair of Electrons: Singlet and Triplets ............................................................................. 49  
3.3 – Kaplan, Solomon, and Mott Model .................................................................................. 58  
3.4 – Zero-Field Spin Dependent Transport ............................................................................. 65  

## CHAPTER 4 – SIGNAL PROCESSING OF SPIN DEPENDENT SIGNALS ......................................... 81  
4.1 – Noise Observed in EDMR ............................................................................................... 81  
4.2 – Lock-in Amplification ...................................................................................................... 86  
4.3 – Signal Averaging ......................................................................................................... 91  
4.4 – Adaptive Signal Averaging ............................................................................................ 93  
4.4.1 – ASA Algorithm ................................................................................................... 95  
4.4.2 – ASA Performance Enhancement ....................................................................... 100  
4.4.3 – ASA Simulation and Results ............................................................................. 106  
4.5 – Higher Order Harmonic Detection ................................................................................. 111  

## CHAPTER 5 - EXPERIMENTAL ........................................................................................................... 116  
5.1 – Devices Used ............................................................................................................. 116  
5.2 – Low- / Zero-Field Spectrometer Development .............................................................. 118  
5.2.1 – Magnetic Field Control ..................................................................................... 120  
5.2.2 – RF Resonant Circuit .......................................................................................... 124  
5.2.3 – Lock-In Amplifier ............................................................................................. 131  

## CHAPTER 6 – HIGHER ORDER HARMONIC DETECTION OF EDMR ............................................ 135  
6.1 – Passage Effects .......................................................................................................... 135  
6.2 – 4H SiC BJT Response: Defect Aggregate ...................................................................... 140  
6.3 – 4H SiC MOSFETs: The Silicon Vacancy Defect .......................................................... 144
List of Figures

Figure 2.1: Classically analogy of the relationship between orbital angular momentum and a magnetic moment for a charged particle orbiting an axis.............................................................. 7

Figure 2.2: Illustration of the dipole-like fields generated by a bar magnet, an electron, and a magnetic nucleus. Though not drawn to scale, the relative size of the field lines and particle are intended to illustrate the differences. For example, electrons are much lighter smaller particles that nuclei and will therefore generate a significantly larger field due to its larger magnetic moment relative to its intrinsic angular momentum. ......................................................... 10

Figure 2.3: Comparison of orbital, electron spin, and nuclear spin angular momentum relationships. The nuclear spin in this case is that of aluminum which is a spin $I = 3/2$. The magnetic field defines the axis of quantization................. 15

Figure 2.4: Comparison of the Zeeman energies of an electron and a magnetic sodium nucleus in the presence of a magnetic field. A single electron will only have two energy states, whereas a single nucleus may contain many energy states. Here, the sodium Na atom has a nuclear spin $I = 3/2$ with 100% abundance. .......................................................... 20

Figure 2.5: $\theta$ and $\phi$ define the angles that the magnetic field form with the principle axes................................................. 22

Figure 2.6: Bruker EMX EPR spectrometer.............................................................................................................................. 23

Figure 2.7: Illustration of (top) electromagnetic wave, (left) a TE_{102} microwave cavity demonstrating a maximum linear polarized magnetic field at the center of the cavity and (right) the magnitude of this field as a function of time. As the figure illustrates, a linearly polarized field is equivalent to two opposite rotating circularly polarized magnetic fields. This concept is important when understanding the rotating reference frame. It is imperative that the two magnetic field components $B_i$ and $B_j$ are perpendicular to each other because of the torque required to observe the EPR effect................. 23

Figure 2.8: Energy splitting of the two electron spin states in the presence of a magnetic field. When the resonance condition of an electron is met, it is able to transition between states.......................................................... 25

Figure 2.9: Illustration of electron magnetic moments precessing about the axis of the applied magnetic field due to spin in the up and down states ....................................................................................................................................... 28

Figure 2.10: Illustrating the orientations of individual magnetic moments due to spin of unpaired electrons in the (top) absence and (bottom) presence of an externally applied magnetic field. In the absence of a magnetic field, the angular momentum quantization axis is undefined so the magnetic moments are pointing in random directions. In the presence of a magnetic field, the electron spins precess about or against the applied magnetic field. Summing these spins results in a magnetization in the direction of the applied field because the transverse components cancel and there are a net positive spins in the lower energy state. .................................................................................................................. 30

Figure 2.11: Illustration of the laboratory frame and rotating reference frame. In the laboratory reference frame, the $i$ and $j$ axes are fixed and the field rotates. In the rotating reference frame, the $i$ and $j$ axes rotate, the $i'$ and $j'$ axes are fixed, and the field remains constant along $i'$............................................. 32

Figure 2.12: Illustration of the effective field and projection of the magnetization in the $i'$ directions................................. 34

Figure 2.13: Illustration of magnetization components in the laboratory reference frame ................................................................. 34

Figure 2.14: (left) precession of the magnetization vector in a cone and (right) plot of the magnetization components through resonance................................................................. 36

Figure 2.15: In the Bohr model of an atom, an electron orbits the nucleus. (left) Intrinsic angular momentum of electron and nuclei give rise to electron and nuclear magnetic moments. (right) To the electron, the nucleus appears to be orbiting around it. The orbital angular momentum of the electron due to its motion about the nucleus give rise to an orbital magnetic moment................................................................. 37

Figure 2.16: Dipole field generated from a magnetic nucleus with an unpaired electron in an s and p (two orientations) orbitals........................................................................................................ 42

Figure 2.17: Imaginary defect system to illustrate hyperfine interactions................................................................. 44

Figure 2.18 illustrates the energy level diagram for the imaginary system illustrated in figure 2.17 and the spectrum observed................................................................. 45

Figure 3.1: Illustration of the three main types of recombination in a semiconductor................................................................. 47
Figure 3.2: Illustration of precursor spin pair formation prior to recombination. In this model for SDR, only singlet pairs that form will responsible for recombination. ................................. 48

Figure 3.3: State transition diagram for raising and lowering operators for a single electron .................................................. 52

Figure 3.4: Illustrating the state transition diagram for effect of raising and lowering operators on triplet and singlet states. The up arrows represent the paths of the raising operator and the down arrows represent the paths of the lowering operator. .................................................................................. 55

Figure 3.5: Comparison of electron up and down spin vectors and their vector quantum averages ................................. 59

So, equation (3.3.2) can be rewritten as, ............................................................................................................... 59

Figure 3.6: Probability of singlet and triplet pairings as a function of the angle between the spin vectors. Note that pure singlets are formed 50% of the time when $\theta = 180^\circ$ and pure triplets are formed 100% of the time when $\theta = 0^\circ$. .................................................. 60

Figure 3.7: (left) By fixing the electron spin $SA$, one can envision the other electron spin $SB$ pointing in an isotropic distribution of angles with respect to that fixed axis. Note that of all the orientations that $SB$ can be possess, it will have the least probability of aligning exactly with $SA$ but the highest probability of aligning perpendicular to it along the disk illustrated around the sphere. Because it is only the relative angle $\theta$ between the pairs that matters, and not the absolute angles $\theta, \phi$, one can map the (middle) 3-dimensional coordinate system to a (right) 2-dimensional coordinate system which accounts for all orientations of pairings and only requires the use of the angle $\theta$ .................................. 62

Figure 3.8: Singlet triplet mixing allows the triplet states to transition into singlet states because all of the basis states are degenerate at zero magnetic field. As a result, recombination will be enhanced at zero-field where the mixing occurs. ..... 68

Figure 3.9: Singlet triplet mixing allows the triplet states to transition into singlet states because at certain magnetic fields, they are energetically accessible. As a result, recombination will be enhanced at these tiny fields (very close to 0G) at which the mixing occurs. Note that if the exchange interaction is present in a measurement, one should expect to observe a split peak about zero magnetic field .......................................................... 71

Figure 3.10: Singlet triplet mixing due to dipolar coupling between two electrons. As a result, recombination will be enhanced at these tiny fields (very close to 0G) at which the mixing occurs. .......................................................... 76

Figure 3.11: Example illustrating a 66.67% abundant spin $I = 1/2$ magnetic nucleus that is in the vicinity of the electron spin pair. At zero applied magnetic field, a recombination response will be observed due to singlet triplet mixing. However, the magnetic nucleus will exert a magnetic field on the spin pair. Therefore, additional recombination responses can be detected when the externally applied magnetic field cancels the hyperfine field.................................................. 77

Figure 3.12: Reorientation of the electron and nuclear quantization axes at high, moderate, and low magnetic fields. ..... 77

Figure 4.1: Sources of noise in an EDMR spectrometer. Shot, flicker, and thermal noise arise from the DUT, electromagnetic interference (EMI) arises from spectrometer hardware, and additional thermal noise arises from the preamp, and the sampler or digitizer is responsible for quantization noise .................................................................................................................. 81

Figure 4.2: (left) Frequency spectrum of a MOSFET under 3 different biasing conditions. The numbers in red indicate the amount of current flowing through the device. (right) simulated flicker, thermal, shot, and preamp noise for a MOSFET that has 5 uA of current passing through it. The bottom plot compares the summation of the simulated noise sources to the frequency spectrum of a MOSFET that has 5 uA of current passing through it. Note that the amount of flicker noise that is present is dependent on the amount of DC current flowing through the device ................................................. 83

Figure 4.3: Illustration the basic components of a preamplifier that should be used with an EDMR spectrometer ............. 84

Figure 4.4: Individual EDMR scans of identical devices with gate areas of 500 $\mu$m$^2$, 40,000 $\mu$m$^2$, 160,000 $\mu$m$^2$ and SNR of EDMR spectra for each measurement .......................................................................................................................... 85

Figure 4.5: (left) Simplified block diagram of adaptive signal averaging and (right) adaptive linear predictor used in the adaptive signal averaging technique. Note that the desired signal is replaced with the unfiltered, conventional average. .......................... 94

Figure 4.6: (a) This figure illustrates the reduction in noise power for both the unfiltered and filtered averages as a function of scan for different values of the exponential weighting factor. The range of $\lambda$ is $0.55 \leq \lambda \leq 0.95$ with 0.05 increments, (b) This figure illustrates the reduction in noise power for both the unfiltered and filtered averages as a function of scan for different values of the turn on time $L$. The turn on time was varied from $2 \leq L \leq 40$ at increments of 4 scans. ............... 107
Figure 4.7: (a) The top plot illustrates the desired simulated signal and the bottom plot illustrates a representative noise sequence \( v(n) \) that was added to create \( x(n) \). The three smaller signals to the left are enlarged in a few of the plots in the figure to better visualize results of the simulation. (b) The bottom plot illustrates the 100th individual unfiltered scan and the top figure illustrates its filtered version using the average of 100 unfiltered scans as the desired signal in the ALP. Note that only two of six line shapes are apparent in the individual unfiltered trace while five of the six line shapes are revealed in an individual filtered trace. (c) The top plot illustrates the average of 85 filtered scans and the bottom figure illustrates the average of 100 unfiltered scans. Note that all 6 line shapes are almost completely resolved in the filtered average. (d) The top plot illustrates the average of 985 filtered scans and the bottom figure illustrates the average of 1000 unfiltered scans. Note that even now, the conventional average still does not match the SNR achieved when averaging 85 filtered scans.

5

Figure 4.8: (a) Histogram of the 85th sequence of prediction errors \( y(n) \) with a superimposed Gaussian fit. This figure justifies the assumption made earlier about the prediction error distribution being Gaussian. (b) Histogram of the averaged (985 scans) prediction errors \( y_{985}(n) \) with superimposed Gaussian fit. This close correspondence is expected as suggested by the central-limit theorem.

Figure 4.9: The top plot illustrates the reduction in noise power for both unfiltered and filtered averages as a function of scan. The bottom plot illustrates the noise reduction factor between these two variance levels as a function of scan. Note that the filtered average maintained a noise power level that was consistently 10 times lower than that of the unfiltered average. This would correspond to a SNR improvement by a factor of more than 3.

Figure 4.10: The effect modulation has on the resultant signal when passing through resonance. (top) Passage through field and (bottom) passage through time. The frequency simulated is extremely low to illustrate the effects of the modulation over the Gaussian line shape.

Figure 4.11: Simulation of the effects of magnetic field modulation over the Gaussian line shape illustrated in figure 4.10.

Figure 5.1: Photos taken of the devices used in this study: 4H SiC BJT, amorphous SiC:H capacitor, older generation 4H SiC MOSFET, newer generation 4H SiC MOSFET, and Si MOSFET.

Figure 5.2: Bipolar opoisolation circuit obtained from Analog Devices application note HCNR200 used to isolate grounds of the controller and the Varian magnet.

Figure 5.3: Block diagram and photo of X-band EDMR spectrometer.

Figure 5.4: Block Diagram and photograph of the low field EDMR spectrometer.

Figure 5.5: PI controller developed in Labview to control the sweeping \( B_0 \) magnetic field.

Figure 5.6: Magnetic field correction algorithm. Because the magnetic field samples read in by the computer will not always align as a function of time, a field correction algorithm is needed.

Figure 5.7: Series resonant circuit and series resonant circuit with additional capacitor for impedance matching.

Figure 5.8: Simulated response of the resonant circuit with a resonant frequency of 210 MHz. The top plot illustrates the real component and the bottom plots illustrates the imaginary component. Note that at resonance, the real (resistance) component is almost matched to 50\( \Omega \) and the imaginary component (reactance) almost goes to zero.

Figure 5.9: (left) Resonant circuit designed for the low-field spectrometer and (right) data gathered to test the response of the circuit. A sniffer coil was used to detect the EMF that was generated by the solenoid.

Figure 5.10: Block diagram of the VLIA developed in Labview. As indicated by the figure, the user has the ability to demodulate up to any 4 harmonics simultaneously.

Figure 5.11: Block diagram of the demodulator used in the VLIA. As indicated by the figure, the user has the ability to control the fundamental frequency, the demodulated harmonic, demodulator phase, low pass filter order and time constant, and decimation/averaging method.

Figure 5.12: Illustration of the average and resampling block. This simplified illustration demonstrates how this block reduces the number of samples from \( N \) to \( K \) using a windowed averaging method. In this particular example, \( N = 10 \) and \( K = 5 \).

Figure 6.1: Illustration of adiabatic passage. The vertical lines indicate the moment at which the spins are in resonance. Note that in adiabatic passage, the spin packets have time to recover before consecutive modulation cycles pass through resonance.
Figure 6.2: EDMR results on a 4H SiC BJT. This plot illustrates (top left) the in-phase, 2\textsuperscript{nd} harmonic detected spectrum as a function of modulation frequency, (top right) the quadrature, 2\textsuperscript{nd} harmonic detected spectrum as a function of modulation frequency, and (bottom) peak-to-peak amplitude of the quadrature, 2\textsuperscript{nd} harmonic detected spectrum as a function of modulation frequency. Note that a 6 kHz, there is a change in line shape for both the in-phase and quadrature demodulated spectra. The in-phase signal flips sign while the quadrature signal changes from a second derivative to an absorption spectrum. This is the transition into fast passage. ..................................................................................................................... 141

Figure 6.3: Data obtained on 4H SiC BJT at X-band frequencies. The BJT signal is composed of a dominant isotropic center line with multiple sets of isotropic (and slightly anisotropic) side peaks. The left column of figures illustrate 3D surface plots constructed from linearly interpolated EDMR scans acquired for different orientations of the device within the applied magnetic field. The top right plot illustrates the conventional 1\textsuperscript{st} harmonic detection of the EDMR signal with extensive signal averaging and bottom right figure illustrates the 2\textsuperscript{nd} harmonic fast passage detection of the EDMR signal. The spectra indicate that a defect aggregate is most likely involved........................................................................................................................... 142

Figure 6.4: Least squares fit of fast passage spectrum illustrated in figure 6.2 using Gaussian functions. ......................... 143

Figure 6.5: 650G wide X-band (left) spectra and (right) its derivative illustrating very small anisotropic hyperfine interactions. These hyperfine interactions are not studied in this work because the low-field spectrometer does not allow for such high fields to be generated to cancel the fields produced by these interactions.................................................. 144

Figure 6.6: Illustration of the silicon vacancy in 4H SiC and the results of the simulation using previously reported hyperfine parameters. .......................................................................................................................... 146

Figure 6.7: This figure illustrates the simulated and measured EDMR response using the 2\textsuperscript{nd} harmonic detection of the fast passage signal. The close correspondence between the simulation and the data indicate that the observed spectrum is most likely dominated by the silicon vacancy................................................................................................................... 147

Figure 7.1: Process of SDR in a pn junction via paramagnetic deep level defects through the formation of spin pairs. (a), (1) conduction electron and defect electron couple to form a triplet pair, (2) pair dissociates because recombination is forbidden. (b), (1) conduction electron and defect electron couple to form a singlet pair, (2) because angular momentum is conserved, recombination of electron and hole is now possible. (c) Top view photograph and (d) simplified illustration of the cross sectional view of the device used in this study. ................................................................................................... 148

Figure 7.2: Low-field detection of SDR via EDMR compared to zero-field detection of SDR in the base collector junction of a 4H SiC BJT. (a) 3D mesh of representative EDMR scans on a SiC diode for a series of $B_1$ amplitudes. (b) Amplitude of the zero-field and low-field resonant signals as a function of $B_1$. The EDMR amplitudes increase monotonically with increasing $B_1$. Note that the increase in the $B_1$ field saturates the resonant SDR response but doesn’t affect the amplitude of the zero-field response. (c) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied. The arrows in the resonant trace indicate the Breit-Rabi shift of the normally (high-field) isotropic hyperfine peaks at low-fields. (d) Comparison of the derivative of a spectrum acquired without any modulation (blue, top) and the spectrum obtained from a 1kHz, 2 Gauss modulation (red, bottom).............................................................................................................. 149

Figure 7.3: (a) First derivative of spectra observed at zero field (blue) and at X band (red) taken at 9.5 GHz, 3394 Gauss. (b) Second derivative of the spectra illustrated in (a). Note that each peak observed at X band is also seen in the zero field spectrum. (c) Zero-field and (d) derivative of spectrum acquired with low modulation amplitude (~ 0.25 Gauss). Note the deviation in slope at precisely zero Gauss illustrated by inset of (c) which shows up as a double peak in (d). ............... 151

Figure 7.4: This set of figures illustrates the behavior of the zero-field response as a function of applied bias. (a) Simulation of the recombination current for a 4H SiC diode at room temperature. (b) Illustration of the peak-to-peak amplitude of the zero-field signal as a function of applied bias plotted against the simulated recombination current displayed in (a) and the DC diode current. The close similarity of the plots demonstrates that this phenomenon is a recombination process. (c) The percent change in the zero-field SDR current acquired by dividing the peak-to-peak amplitude by the dc current as a function of bias voltage. Note that this response peaks at 2.35 volts with a relatively large change of almost 1%. (d) Integrated spectra acquired for the data plotted in Figs. 4(b) and 4(c). Note that as the bias is increased beyond the built-in voltage ($V_i \geq 2.65$ V), the SDR signal begins to decrease. Note also that beyond this voltage, a double peak appears to be forming. The four traces on the inset of this figure illustrate the transition of the single peak into two. .............................................................................................................................................................. 154

Figure 7.5: BJT zero-field spectrum versus temperature. Note that the hyperfine lines are independent of temperature as indicated by the first and second derivatives............................................................................................................................................ 155
Figure 7.6: Low-field detection of SDR via EDMR compared to zero-field detection of SDR. (a) 3D mesh and (b) 2D traces of representative EDMR scans on a SiC MOSFET for a series of $B_L$ amplitudes. (c) Amplitude of the zero-field and low-field resonant signals as a function of $B_L$. The SDR amplitudes increase monotonically with increasing $B_L$. Note that the increase in the $B_L$ field saturates the resonant SDR response and also affects the amplitude of the zero-field response unlike the case for the BJT. (d) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied. 

Figure 7.7: Advantages of low-field SDR. (a) Low- and zero-field trace of SDR in 4H SiC MOSFET and (b) zoomed in plot of the resonant signal. Note that at low fields, the isotropic side peaks maintain their splitting, however, they are shifted toward 0G. Note that the apparent hyperfine peaks that are separated by 68 G are not consistent with the low field shift predicted by the Breit Rabi correction. Because the lower peak is located at a field at half the resonant field, it may actually correspond to a forbidden transition that is sometimes observed at low magnetic fields when dipolar interactions are present [101]. (c) Comparison of zero-, low-, and high-field (shifted from 76G and 3395G respectively) spectra and (d) comparison of all three normalized responses centered about 0G.

Figure 7.8: Comparison of SDR via EDMR and zero-field detection. Here, (a) illustrates the first derivatives of these spectra while (b) and (c) represent the derivative of the spectra. Note that the inflections indicated by the green arrows suggest the presence of exchange or dipolar interactions of the electrons in the spin pair.

Figure 7.9: This figure illustrates the SDCP signal and its derivative observed in the absence of externally applied electromagnetic radiation. The inset of the figure illustrates the biasing scheme used in the measurement.

Figure 7.10: Low-field detection of SDR using the biasing scheme of BAE via EDMR and zero-field detection of SDR. (a) 3D mesh and (b) 2D traces of representative EDMR scans on a SiC MOSFET for a series of $B_L$ amplitudes. (c) Amplitude of the zero-field and low-field resonant signals as a function of $B_L$. The EDMR amplitudes increase monotonically with increasing $B_L$. Note that the increase in the $B_L$ field saturates the resonant SDR response and also affects the amplitude of the zero-field response like the case for the other MOSFET investigated. The inset of this figure illustrates the technique of biasing scheme of BAE. (d) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied.

Figure 7.11: (a) Comparison of zero-field and X-band spectra on a 4H SiC MOSFET using the bipolar amplification effect. The derivative of this spectrum is illustrated in figure (b) and (c).

Figure 7.12: Because tunneling also obeys the law of conservation of angular momentum, the formation of singlet pairs in neighboring defects can be used describe spin dependent tunneling of electrons through a dielectric. (a) An electron will not tunnel if the two coupled electrons do not form a singlet pair. (b) With the application of electromagnetic radiation, the electrons are able to flip their spin state which enhances the chance of a singlet pairing and allows the conduction electron to tunnel into a trap. (c) The trapped electron will only be allowed to tunnel to neighboring traps if the tunneling process continues to conserve angular momentum.

Figure 7.13: Low-field and zero-field detection of SDT. (a) 3D mesh and (b) 2D traces of representative EDMR scans on a SiC MOSFET for a series of $B_L$ amplitudes. (c) Amplitude of the zero-field and low-field resonant signals as a function of $B_L$. The EDMR amplitudes increase monotonically with increasing $B_L$. Note that the increase in the $B_L$ field saturates the resonant SDT response but hardly affects the amplitude of the zero-field response. (d) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied.

Figure 7.14: Comparison of the SDT results obtained at zero-field and X-band. (a) Zero-field signal plotted against the X-band spectrum centered at zero Gauss and (b) its derivative. Note that the large inflection point at zero-field appears as a double peak when the derivative is taken indicates that there is very strong electron-electron coupling. These interactions are most likely due to the exchange or dipolar interactions of electrons in neighboring defects.

Figure 7.15: Illustration of the size of the resonant and zero-field phenomena. (a) Amplitude of zero-field SDT $\Delta IZF$ and DC leakage current $I$ as a function of applied bias (b) Amplitude of resonant field SDT $\Delta IRF$ and DC leakage current $I$ as a function of applied bias. (c) Change in current ($\Delta IZF/I$) over zero-field as a function bias. (d) Change in current over a resonant field ($\Delta IRF/I$) as a function of bias. Note that these curves are inconsistent with recombination which suggests that the zero- and low-field responses are due tunneling.

Figure 7.16: (a) Resonant (79G, 220MHz ) spectrum obtained using a high stressing bias of 4.5 V and a modulation of 6G. Figure (b) illustrates the derivative and (c) illustrates the zoomed in version of the derivative. The two symmetrically spaced peaks indicated by the arrows are most likely attributed to the EPR half-field transitions that arise in the presence of dipolar interactions of neighboring defects.
Figure 7.17: Results of high bias stress on an amorphous SiC capacitor while performing in-situ EDMR and zero-field detection. This figure illustrates (a) final spectrum obtained after experiment, (b) amplitude of zero-field SDT signal as a function of time, (c) amplitude of -80G resonant signal as a function of time, and (d) amplitude of +80G resonant signal as a function of time. (Note the running average in plots (b)-(c) is simply the average from the 1st scan to the last scan whereas the windowed average represents the average of the five most recent scans as a function of time.)

Figure 7.18: Illustrations of $P^b_0$ and $P^b_1$ defects. These defects are created at the Si/SiO$_2$ interface as a result of elevated bias and temperature stressing of a MOSFET.

Figure 7.19: (a) Comparison of the DCIV measurements (bottom, red) pre and (top, blue) post stress. The mild stressing conditions involved applying a -25V to the gate of the MOSFET for 10,000 seconds at 150°C. The increase in the amplitudes of the two peaks indicates that a significant number of defects were created at the Si/SiO$_2$ interface. (b) Comparison of low-field measurements made (bottom, red) pre and (top, blue) post stress. (The pre stress spectrum is offset by 4pA for clarity.) As illustrated, the spectrum obtained from the stressed device reveals a zero-field and two resonant signals which are present due to the generation of defects at the Si/SiO$_2$ interface as a result of the negative bias temperature stress.

Figure 7.20: (a) Comparison of measurements taken (top, blue) with and (bottom, red) without electromagnetic radiation when biased with +1.50V. (b) Comparison of measurements taken (top, blue) with and (bottom, red) without electromagnetic radiation when biased with -0.30V. The significant inflection point observed precisely at zero Gauss is likely due to the exchange interaction the electrons experience when an appropriate electric field bias is applied to the gate of the MOSFET.

Figure 7.21: Comparison of the zero-field and low-field ($B_0 = 75G, \nu = 209MHz$) amplitude responses as a function of applied oscillating $B_1$ magnetic field when the MOSFET was biased with $V_G = +1.5V$ and $VDS = +0.32V$. As expected, the resonant SDR signal saturates when maximum power is applied and ceases to exist when no power is applied. The zero-field SDR signal is completely unaffected by this perturbing field.

Figure 7.22: (a) Illustration of the zero-field and low-field amplitude responses as a function of applied bias plotted against the current measured in the DCIV method. Note that the zero-field response is highly correlated with the low-field and DCIV responses. This is direct evidence that the zero-field signal is due to spin dependent recombination at the Si/SiO$_2$ interface of the MOSFET. (b) Three dimensional mesh plot of the integrated zero-field spectra used to plot the amplitude response in (a). Note that as the voltage is decreased from $V_G = +1.5V$, a double peak begins to form which is most likely attributed to electron exchange coupling.

Figure 7.23: Comparison of (top) high-field ($B_0 \approx 3287 G, \nu = 9.228 GHz$), (middle), low-field ($B_0 \approx 110 G, \nu = 305 MHz$), and (bottom) zero-field signal (no applied radiation). Note that the two signals observed at high-field are unresolvable in the zero- and low-field measurements because of the reduction in resolution that occurs when making measurements at low magnetic fields.

Figure 8.1: Illustration of the technology used as a magnetic field sensor.

Figure 8.2: Illustration of the response of the sensor in the presence and absence of a magnetic field. The magnetic field is easily measured by calculating the shift in the response from zero magnetic field, the ideal response. The measured field is the $B_{offset}$. In this schematic illustration, $B_{offset}$ is 15 Gauss. The measured field illustrated was chosen merely to make the underlying principle clear. However, much smaller local fields could and generally would be utilized in the measurement.

Figure 8.3: Illustration as the technology used as a miniature spectrometer for wafer probing stations.

Figure 8.4: Photograph of a desktop zero- and low-field spectrometer.
List of Tables

Table 2.1: Number of nucleons within a nucleus will determine whether it can be detected with NMR. ..................................... 10
Table 2.2: Summary of orbital, electron spin, and nuclear spin angular momentum relationships. ........................................ 14
Table 2.3: Summary of operators. ............................................................................................................................................. 18
Table 3.1: Summary of total and pair states for two electrons. .................................................................................................. 54
Table 3.2: Probability of pure singlet and pure triplet states given the individual electron spin states. .................................... 57
Table 6.1: Amplitudes of the simulated Gaussian functions illustrated in figure 6.3 ................................................................. 143
Table 6.2: Table illustrating the more probable combinations of magnetic nuclei at the site of a silicon vacancy .......... 145
Acknowledgements

I would like to first thank my advisor Dr. Patrick Lenahan for his help and support for providing me with the background necessary to understand the topics of magnetic resonance and semiconductor physics. He has been an outstanding mentor and friend throughout my time at Penn State. Under his supervision and guidance, he has taught me to become a well-rounded and an independent researcher. I would like to thank him specifically for providing me with the freedom to apply my background of digital signal processing to investigate new and undiscovered physics. It was this work in particular that made my graduate career at Penn State most enjoyable.

Next, I would like to thank the members of my thesis committee: Professors of engineering science and mechanics S. Ashok and Mark Horn and Professor of electrical engineering Jerzy Ruzyllo. I would also like to thank the faculty and staff of the Engineering Science Department for their support and guidance throughout my time at Penn State University. I want to acknowledge my labmates for their support and friendship: Jacob Follman, Thomas Aichinger, Thomas Pomorski, Brad Bittel, Michael Mutch, Mark Anders, Michael Pigott, Duane McCrory, and Tom Roher. Of all the fun times had, I will especially miss the groups weekly wing nights.

I would like to thank my parents Diane and Ed, for it was them who financially supported my undergraduate years which provided me with the skills and foundation for this work. I really appreciate the day trips they made to PSU for our monthly lunch visits. I want to thank my brothers EJ and Clayton for their support and friendship. Their occasional visits to PSU have left me with many fun and interesting memories. I also want to thank my Aunt Marty who has always been a second mother to me. Although outside their area of expertise, I want to thank the entire family for their continued willingness to help me over the years understand the difficult topics of signal processing, semiconductor physics, and quantum mechanics.

Finally, and most importantly, I would like to thank my wife Andrea for her continuous love, understanding, and patients during my time as a student. She has always provided me with unconditional support and encouragement. My more innovative ideas have always come about in her presence. This work would not be possible without her. I also want to give a special thanks to the Lesh family: Judy, Geoff, and Eric. I want to thank them for welcoming me into their family and providing me with an environment to perform my research while away from Penn State.

I would also like to acknowledge our lab sponsors. This work has been primarily supported by the US Army Research Laboratory, with additional support from Intel Corporation and General Electric under the U.S. Department of Commerce under award no. NIST60NANB10D109. I want give a special thanks to Aivars Lelis from the US Army Research Lab for his support, collaboration, and friendship.
CHAPTER 1 – INTRODUCTION

The semiconductor industry is currently pressing its technological limits in attempts to keep pace with Moore’s Law. Despite much skepticism, corporations still manage to mass produce micro- and nano-electronic circuits which enables the continuing trend of doubling the number of transistors that can be placed on a chip every 18 months. Only recently has it been realized by scientists and engineers that, even though the goal they are trying to meet is steadily increasing, the requirements for advances in the downscaling of technology are ever more challenging. Today, devices are being fabricated on an extremely small scale (oxide thicknesses of about 10 atoms across) and are being made with new and more complex material systems. At such extremely small scales, individual atomic scale defects can be responsible for the improper functioning of such devices. As a result, new and more innovative ways are needed to detect these atomic scale defects in fully processed devices. Also, many times the device fabrication process is essentially optimized through trial and error in order to determine which material systems and processing steps are necessary to achieve the best performing and most reliable devices. Despite being a relatively successful approach to design a working device over the period of many years, it is most definitely not the best practice to design a working device in a few years. In order to expedite this process, one needs to understand why certain processing steps don’t work as well as other ones. With a physical understanding of why some processing conditions work better than others, it is likely that process optimization would be more rapid and less expensive. In order to answer this question of why, once again, defect detection at the atomic scale is of the utmost importance for a complete understanding of processing. An understanding of the effects processing variations have on defect structure and density will surely save device manufacturers much time and money. The problem that remains is that there are a very limited number of options for defect detection at the atomic scale.

1.1 – Atomic Scale Defect Detection

There are many techniques which allow scientists and engineers to understand the defects in fully processed devices that are responsible for performance limitations. However, only a handful can actually provide insight at the atomic level in a fully functioning device. This is extremely important when one wants to understand the implications of downscaling devices. The imaging techniques of tunneling electron microscopy (SEM) and scanning tunneling microscopy (STM) are capable of atomic scale defect detection but are quite limited to a visual imaging. These techniques cannot be readily applied to fully processed devices and can provide little information on individual point defect structure and chemistry. These techniques are also of limited value in the detection of low density defects such as vacancies or
other extrinsic point defects. In addition, they are also invasive techniques which destroy the fully processed device when looking at interface layers. Therefore, defects actually observed in such measurements may not be the result of device processing but rather due to mechanical damage caused by the sample preparation. There is a non-invasive way to detect defects in fully processed transistors which can actually detect defects on a much smaller scale; it is based upon magnetic resonance (MR). In addition to being non-invasive, these measurements offer the capability of detecting only those defects that are directly responsible for the performance limitations and proper functionality of fully processed devices.

1.2 – Electrically Detected Magnetic Resonance

Magnetic resonance (MR) was discovered by a Russian physicist by the name of Yevgeny Zavoisky in 1944 [1]. Since then, MR has been productively applied in many areas. The most commonly utilized MR techniques are electron paramagnetic resonance (EPR) (also commonly referred to as electron spin resonance, or ESR), nuclear magnetic resonance (NMR), magnetic resonance imaging (MRI), optically detected magnetic resonance (ODMR), electrically detected magnetic resonance (EDMR), and a handful of other approaches. In all of its forms, the heart of magnetic resonance is the principle that particles such as electrons, protons, and neutrons have an intrinsic quantum mechanical property called spin. MR can be observed by inducing transitions between electron spin states via EPR or nuclear spin states via NMR. Although MR detection typically involves an effective change in the electronic properties of resonant circuit or cavity (EPR, NMR, MRI), MR can also be observed by monitoring spin dependent transport, such as recombination, tunneling, scattering, hopping, photoconductivity, or photoemission, within fully processed electronic device (EDMR, ODMR). In all of its forms, MR is extremely useful in understanding the structure of matter, understanding electronic transport, defect detection, and medical imaging. It is also useful in any area in which atomic scale detection or manipulation of paramagnetic entities is required.

EDMR measurements are typically performed at X band frequencies (9 GHz) for the same reasons that X-band is most commonly utilized in conventional EPR. This frequency band is typically used because of the tradeoffs between increased resolution and cavity dimensions. (High frequency measurements allow for better resolved spectra but require the utilization of smaller microwave cavities) EDMR allows for the electrical detection of atomic scale defects such as dangling bonds or spin related transport phenomenon within a fully processed micro- or nano-electronic device. Much work with EDMR has already been performed on devices that include metal oxide semiconducting field effect transistors (MOSFETs) [2] [3] [4] [5] [6], bipolar junction transistors (BJTs) [7] [8], solar cells [9] [10], and capacitors [11]. Defects such as dangling bonds are a major cause of reliability problems in modern solid
state material systems. MR detection of these defects can allow the determination of the physical and chemical nature of the defect structure, thus permitting one to understand the material physics behind the limitations in performance which they cause. Unlike the measurement of reflected microwaves in EPR, EDMR signals are obtained by monitoring a change in spin dependent current (recombination, tunneling, hopping, or scattering) from an appropriately biased device placed inside an EPR spectrometer. It is a very powerful technique which has a sensitivity many orders of magnitude greater than that of conventional EPR. The physical mechanisms responsible for this increased sensitivity are not completely understood; however, many have attempted to explain it [12] [13] [14] [15] [16] [17]. The collective agreement at this time is based on a model proposed by Kaplan, Solomon, and Mott (KSM) [13]. The key idea in the KSM model is that a very large change in device current can be accounted for by realizing that spin pair states are formed prior to a microscopic transition event and exist in this state for a finite amount of time [15]. In the case of recombination, a conduction electron (or a valence band hole) and defect (with unpaired electrons) will couple to form an intermediate spin pair state if they are close enough in proximity. Because the conduction electron (or the valence hole) and the defect electron have two discrete spin states in the presence of an externally applied magnetic field, the coupled spins can be represented by the symmetric triplet states $T_+ = ↑↑$, $T_0 = (↑↓ + ↓↑)/\sqrt{2}$, $T_- = ↓↓$ each having total spin $s = 1$ and the antisymmetric singlet state $S_0 = (↑↓ − ↓↑)/\sqrt{2}$ which has a total spin $s = 0$. These pair states can then be destroyed either by dissociation (electrons released back to the conduction band or holes back to the valence band) or recombination (the annihilation of the electron hole pair). What distinguishes whether these pairs will recombine or dissociate is the pair’s initial total angular momentum. If the total angular momentum is conserved between the pair, then recombination becomes possible. On the other hand, if angular momentum is not conserved for this transition, then the pair will dissociate and the charge carrier will return to its respective energy band. When spin orbit coupling is negligible, the total angular momentum will be solely defined by the spin pair and can be modeled solely by singlet recombination [13]. However, when spin orbit coupling is not negligible, orbital angular momentum needs to be taken into account which gives a non-negligible probability that the pairs in the triplet state will recombine [14]. In either situation, the difference in recombination for the resonant and non-resonant cases can simply be understood from a statistical point of view. When the pairs experience a resonant field which saturates the spin system (destroys the magnetization), the spin distribution of pairs becomes randomized and hence, the probability of a recombination or tunneling event is modified. This can be measured as a change in current when linearly sweeping over the magnetic resonant field. The resonance spectrum is typically observed as an absorption spectrum or as the derivative of an absorption spectrum when a lock-in technique with magnetic field modulation is utilized.
1.3 – Zero- & Low-Field Spin Dependent Transport

In some cases, performing EDMR at lower magnetic fields (radio frequencies) can be useful [18] [19] [20] [21] [22]. For example, EDMR has been used to explore a transport mechanism for electrons tunneling through a series of quantum dots via a Pauli blockade for quantum computing applications [20] [19] [22]. In a number of recent studies, changes in device current have been observed in the presence of low sweeping magnetic fields in the absence of electromagnetic radiation. In some of these studies, this low magnetic field phenomenon is attributed to nuclear spin dynamics and their interactions with electrons [23] [24] [25]. Another low field phenomenon has been observed in organic semiconductors and is commonly referred to as magnetoresistance [26] [27] [28] [29] [30]. This phenomenon has typically been associated (without definitive evidence) with the changing of singlet and triplet ratios as a function of magnetic field. One particular study suggested a model to describe this phenomenon which is very similar to the KSM model mentioned earlier [28]. Note that the models and proposed mechanisms to describe the experiments in the organic semiconductors are only speculative.

We have also observed magnetic field induced current change phenomenon in our low field measurements in the absence of electromagnetic radiation. We demonstrate this low field effect to be associated with spin dependent recombination (SDR) in Si MOSFETs, SiC diodes, and SiC MOSFETs [18] and spin dependent tunneling (SDT) in a-SiC:H capacitors. In EDMR, the microscopic transitions of recombination, tunneling, hopping, and trapping may be hard to discern from each other because they may all produce similar signals. However, we have the ability to utilize devices in our measurements in which we know certain mechanisms (such as recombination and tunneling) will dominate the observed currents. Also, unlike any of the previous studies at zero-field, we can to clearly resolve the electron-nuclear hyperfine interactions and electron-electron dipolar and exchange interactions using magnetic field modulation because of the crystalline devices that we use. To the best of our knowledge, no one has previously been able to measure these interactions using the zero field phenomenon. Not only is this newly discovered physics important for spectroscopists interested in less complex and far cheaper means for semiconductor defect detection, but this phenomenon is also directly applicable to the fabrication of new and innovative devices. The primary goal of this research will be to better understand the zero-field SDR and SDT phenomena, how they relate to EDMR, and how they can be utilized as a means for defect detection in fully processed micro- and nano-electronics. Also, we wish to be able to harness its potential in the design of new magnetic field sensors, magneto-isolation circuits, and inexpensive low-/zero-field spectrometers.
CHAPTER 2 – ELECTRON PARAMAGNETIC RESONANCE

All electrons and some nuclei have intrinsic angular momentum that arises because of the quantum mechanical property of spin. The observation that certain particles contain intrinsic angular momentum was first performed by Otto Stern and Walther Gerlach in 1922 [31] [32] [33]. They showed that a beam of electrically neutral sodium atoms directed through an inhomogeneous magnetic field could be deflected into separate, but discrete, regions. This result suggested that many elementary particles contain intrinsic angular momentum because of the observed interaction with the externally applied magnetic field. Electrically neutral particles were first used because electrically charged particles would be directly influenced by the force on a charged particle due to the electric field \( F = qE \). Here, \( q \) is the charge of the particle and \( E \) is the applied electric field. This experiment could be most simply extended to electrically charged particles (for example, electrons and protons) by applying an \( E \) field in a direction that was perpendicular to both the velocity of the charged particle and magnetic field such that force due to each field component cancels [31]. As a result, charged particles such as electrons and protons could be studied using slightly modified apparatuses. For an electron beam directed through such an apparatus, the electrons will be deflected in one of two possible directions; up or down. Therefore, electrons could occupy one of two states, the spin up state or spin down state. (Particles with two spin orientations are referred to as spin \( s = \frac{1}{2} \) particles, or fermions.) The phenomenon was first termed “spin” because it was initially thought that the intrinsic angular momentum originated from particles “spinning” about an axis. Spin gave the electron and other particles an extra degree of freedom in quantum theory. The idea of an additional degree of freedom was first introduced by Wolfgang Pauli with his principle of exclusion, now known as the Pauli Exclusion Principle (PEP) [34]. The principle was based on the idea that no two electrons within the same orbital could occupy the same quantum state. (More specifically, the principle states that identical fermions must have total wave functions that are antisymmetric to one another. More on this in the upcoming sections.) He was lead to this conclusion based on earlier observations from narrow splitting of hydrogen spectral lines, called fine structure. He essentially introduced a new degree of freedom for the electron but did not really know what to attribute it to. It was actually Ralph Kronig that first attributed it to the electron rotating about an axis [35]. This idea was not published initially because Pauli speculated that the particle would have to be spinning faster than would be allowed by Einstein’s theory of relativity [35]. George Uhlenbeck and Samuel Goudsmit came up with similar ideas and published their results not too long after [32] [35] [36] [37]. This idea, although not precisely correct, will be briefly reviewed because it actually provides a very close approximation to what is actually measured. It is based upon the idea of discretized orbital angular momentum.
2.1 – Angular Momentum and Magnetic Moments due to Spin

In classical mechanics, the orbital angular momentum $L$ of a charged particle of mass $m$ and charge $q$ traveling in a circular path of radius $r$, is given by

$$L = r \times p = r \times mv \hat{k} \quad (2.1.1)$$

where $p$ is the linear momentum and $v$ is the tangential velocity. Therefore, if the orbit is circular, the magnitude of the velocity as a function of angular momentum can be written as,

$$|v| = \frac{|L|}{m|r|} \quad (2.1.2)$$

This circling charge acts as a current loop that also creates a magnetic moment in the direction perpendicular to the plane of orbit. The magnetic moment that is generated is given by,

$$\mu = i \cdot A \hat{k} = q \frac{v}{2\pi r} \pi r^2 \hat{k} \quad (2.1.3)$$

where $i$ is the effective current created by the charged particle and $A$ is the area of the loop in which the particle travels clockwise around the $z$-axis. (Note that if this particle is not charged, a magnetic moment will not be generated because it will appear as if there is no current.) Equations (2.1.1) and (2.1.3) are pictorially illustrated in figure 2.1.
Figure 2.1: Classically analogy of the relationship between orbital angular momentum and a magnetic moment for a charged particle orbiting an axis.

From equation (2.1.3), the magnitude of the velocity as a function of the magnetic moment is given by,

$$|v| = \frac{2|\mu|}{q|r|} \quad (2.1.4)$$

Setting the two velocity quantities of (2.1.2) and (2.1.4) equal to each other allows one to find the ratio of the magnetic moment and the angular momentum, otherwise known as the gyromagnetic ratio $\gamma$,

$$\gamma = \frac{|\mu|}{|L|} = \frac{q}{2m} \quad (2.1.5)$$

In quantum mechanics, the angular momentum is quantized in units of $\hbar$. This was first proposed by Stefan Procopiu and then independently by Niels Bohr shortly afterward. This can be pictured quasi-classically by assuming that the orbital period $T$ of the charged particle must be an integer multiple $n$ of the wavelength $\lambda$,

$$T = 2\pi|r| = n\lambda \quad (2.1.6)$$

Using the DeBroglie wavelength,

$$\lambda = \frac{\hbar}{|p|} = \frac{\hbar}{m|v|} \quad (2.1.7)$$

and combining (2.1.6) and (2.1.7) demonstrates the quantization of angular momentum in units of $\hbar$,

$$T = 2\pi|r| = n\frac{\hbar}{m|v|} \to m|v||r| = |L| = n\hbar \quad (2.1.8)$$

Substituting this equation into the relation relating the magnetic moment and the angular momentum given by (2.1.5), yields an equation that relates the magnetic moment of a particle to the charge and mass of that particle.

$$\frac{|\mu|}{|L|} = \gamma \to |\mu| = \gamma|L| = \frac{q}{2m} n\hbar \quad (2.1.9)$$
This relationship is extremely important as will be shown later because electron and nuclear angular momentum due to spin are postulated to behave similarly as this orbital angular momentum analogy. This simple derived relation illustrates that a particle will have a greater magnetic moment the higher the charge and the lower the mass. Using the lowest integer multiple \( n = 1 \), the Bohr magneton \( \mu_B \) and nuclear \( \mu_N \) magneton are defined as,

\[
\mu_B = \frac{q_e \hbar}{2m_e} = 9.274 \times 10^{-24} \left( \frac{\text{J}}{\text{T}} \right) \\
\mu_N = \frac{q_p \hbar}{2m_p} = 5.051 \times 10^{-27} \left( \frac{\text{J}}{\text{T}} \right)
\]

(2.1.10)  (2.1.11)

The Bohr magneton was named after Niels Bohr. The constant naturally evolved from his model of the atom. The nuclear magneton is named so because it is the base unit of angular momentum observed from a nucleus, the location where protons reside. These simple relationships illustrate why electrons generate a significantly larger magnetic moment than do protons. Note that the magnet moment is dependent upon the charge of the particle. Therefore, the electrons will have a negative magnetic moment. It has been shown by Paul Dirac’s relativistic treatment that the gyromagnetic ratio was off by a factor of 2 for the electron. Subsequent theoretical and experimental work actually showed that this number is more precisely given by 2.0023193 [38] [39] [40]. This constant is known as the free electron spin g-factor denoted by \( g_e \). Nuclei have varying \( g \) values as well which will be denoted by \( g_N \). Therefore, the gyromagnetic ratios are actually equal to,

\[
\gamma_e = g_e \frac{\mu_B}{\hbar} = g_e \frac{q_e}{2m_e} \\
\gamma_N = g_N \frac{\mu_N}{\hbar} = g_N \frac{q_p}{2m_p}
\]

(2.1.12)  (2.1.13)

Note that the \( g \) value gets absorbed into the gyromagnetic ratio. Because these particles possess magnetic moments, they can be thought of as microscopic bar magnets which generate dipole fields. For the electron, this radial dependent dipole field is given by,

\[
\mathbf{B}_e(r) = \frac{\mu_0}{4\pi} \left\{ \frac{3(\mathbf{\mu}_e \cdot \mathbf{r})\mathbf{r} - \mathbf{\mu}_e r^2}{r^5} \right\}
\]

(2.1.14)
where \( \mu_0 \) is the permeability, \( \mu_e \) is the electron magnetic moment vector, and \( r \) is the radial vector from the origin of the electron. The direction of the field lines will depend on whether the electron has a spin up or spin down orientation. When electrons pair in a molecular orbital, they must have opposite oriented spins as postulated by the PEP. Because electron spin is additive, two electrons that occupy the same orbital will have a net spin of zero which will cause their magnetic moments to cancel. A particle, or group of particles, must have a net spin in order to be detected via magnetic resonance because it is the magnetic moments that are detected in the technique. As a result, only unpaired electrons will be rendered EPR active. In many material systems, atomic scale defects involve broken bonds or vacancies. Many of times, these defects will have unpaired electrons, thereby making them EPR active so that atomic scale defect identification is possible.

It was just shown with a quasi-classical analogy that electrons and protons exhibit intrinsic angular momentum and possess a magnetic moment due to their spin property. (Protons are also Fermions which have a spin of \( s = \frac{1}{2} \).) However, protons cannot always be directly measured via magnetic resonance measurements. It is precisely how protons and neutrons pair up within a nucleus which determines whether or not they can be detected via NMR. Like protons and electrons, neutrons are fermions and have a spin of \( s = \frac{1}{2} \). They do not possess charge so the classical analogy to derive a magnetic moment due to orbital motion cannot be applied for these particles. The reason for this phenomenon is because neutrons are composed of subatomic charged particles called quarks \[37\] \[41\].

Similar to how electron spin is additive in an orbital, proton and neutron spins are additive within a nucleus \[41\]. This provides the nucleus a net spin denoted by \( I \) and is typically referred to as the nuclear spin angular momentum quantum number. Also, similar to the way electrons pair up in orbitals, protons and neutrons pair up within a nucleus because all fermions obey the Pauli Exclusion Principle \[42\]. Therefore, if there are any unpaired protons or neutrons, the net spin of the nucleus will not be zero and therefore possess a net nuclear magnetic moment. This would render the atomic site as NMR active because it is precisely the net magnetic moment of the nucleus that is being detected in the NMR measurement. Table 2.1 illustrates the different pairings of nucleons (protons and neutrons) within a nucleus and the type of spin expected \[41\].

Because the combination of protons and neutrons provide the nucleus with a net spin, the nucleus will possess intrinsic angular momentum and a magnetic moment. As a result, a nucleus can also be thought as a microscopic bar magnet similar to the electron. The nuclear field can be described by a dipole field,

\[ \text{Quarks themselves are thought to be made of even smaller particles called strings according to String theory which is a contender for the Theory of Everything (TOE) \[37\]. This topic is far beyond the scope of this study so will not be explained any further.} \]
Table 2.1: Number of nucleons within a nucleus will determine whether it can be detected with NMR.

\[
B_N(r) = \frac{\mu_0}{4\pi} \left\{ \frac{3(\mu_N \cdot r)r - \mu_N r^2}{r^5} \right\}
\]

(2.1.15)

where \( \mu_0 \) is the permeability, \( \mu_N \) is the nuclear magnetic moment vector, and \( r \) is the radial vector from the origin of the electron. As will be discussed later, these nuclear magnetic fields can actually be detected indirectly via EPR from what are known as electron-nuclear hyperfine interactions. Figure 2.2 illustrates the dipole-like fields generated by a bar magnet, an electron, and a magnetic nucleus. Even though this treatment of spin is inaccurate, it provides a result that doesn’t deviate too far from experimental observation. A more precise treatment of spin was developed by Dirac in 1928 using relativistic quantum mechanics [38] [39].

Figure 2.2: Illustration of the dipole-like fields generated by a bar magnet, an electron, and a magnetic nucleus. Though not drawn to scale, the relative size of the field lines and particle are intended to illustrate the differences. For example, electrons are much lighter smaller particles that nuclei and will therefore generate a significantly larger field due to its larger magnetic moment relative to its intrinsic angular momentum.
This treatment is necessary because an electron (or magnetic nucleus) spinning about an axis to create the observed magnetic moment would need to have the particle spinning faster than relativity theory would allow. It was therefore accepted that these particles are not literally spinning but rather exhibit an intrinsic property that provides them with this additional angular momentum. From then on, these phenomena were commonly referred to as electron spin angular momentum denoted by $s$ and nuclear spin angular momentum denoted by $I$.

### 2.2 – Magnetic Resonance of Electrons and Nuclei

The discussion in the previous section provided a quasi-classical orbital angular momentum analogy to think about a particle that contains intrinsic angular momentum and a magnetic moment due to spin. To this day, spin is not completely understood.\(^2\) However, there are many concepts that are widely accepted that need to be understood in order to understand EPR and NMR. The details of the following sections well documented in various text books [41] [43] [44] [45] [40] [46] [47].

The key concept of magnetic resonance is that spin angular momentum is quantized, as originally postulated by Bohr. The question remains: what defines the axis to which quantization is measured? In the absence of an externally applied magnetic field, the axis of quantization (of angular momentum) is randomly distributed for all particles that have spin. However, when a magnetic field is applied, the electrons and magnetic nuclei will tend to reorient themselves such that their quantization axis will be oriented in the direction of the applied magnetic field. But how exactly does this affect the resonance of the electrons and magnetic nuclei? It is postulated that the electron spin and magnetic nuclear spin angular momenta behave similarly to that of orbital angular momentum.

Orbital angular momentum refers to the momentum an electron possesses from its motion about a nucleus. Here, $l$ represents the orbital angular momentum quantum number and $m_l$ represents the orbital projection angular momentum quantum number. The magnitude of the orbital angular momentum is defined by,

$$L = (L_x \hat{i} + L_y \hat{j} + L_z \hat{k})$$  \hspace{1cm} (2.2.1)\strut

where,

\(^2\) Although much is known about spin today, the concept can be very difficult to understand. Feynman wrote that, “It appears to be one of those few places in physics where there is a rule which can be stated very simply, but for which no one has found a simple and easy explanation. The explanation is down deep in relativistic quantum mechanics. This probably means that we do not have a complete understanding of the fundamental principle involved.” [120]
\[ |L| = \sqrt{l(l + 1)}\hbar \]  

(2.2.2)

Here \( l \) is discretized and can take on the values 0 to \( n - 1 \), where \( n \) is the principal quantum number, or energy of the system. The magnitude of orbital angular momentum projected along the axis of quantization (the axis to which the field is applied, usually indicated by the \( \hat{k} \) axis, or \( z \) direction), is quantized in units of \( \hbar \), and is defined by

\[ L_z = m_t \hbar \]  

(2.2.3)

where \(-l \leq m_t \leq l\) and can take only integer values. The components of \( L_x \) and \( L_y \) are random and cannot be predicted because the Heisenberg uncertainty principle states that only one orientation can simultaneously be known while knowing the length [41] [40]. (As a result, when used as quantum mechanical operators, \( |L| \) and \( L_z \) commute, i.e., measurement of one state does not affect the measurement of the other state)

The electron and magnetic nuclei obey the same angular momentum principles. The electron spin angular momentum quantum number is denoted by \( s \) and the electron spin projection angular momentum quantum number is denoted by \( m_s \). As mentioned earlier, an electron has spin \( s = 1/2 \) and \( m_s = \pm 1/2 \). Therefore, the magnitude of the electron spin angular momentum is defined by,

\[ S = (S_x \hat{i} + S_y \hat{j} + S_z \hat{k}) \]  

(2.2.4)

where,

\[ |S| = \sqrt{s(s + 1)}\hbar \bigg|_{s=1/2} = \frac{\sqrt{3}}{2} \hbar \]  

(2.2.5)

and the magnitude of the electron spin angular momentum projected along the axis of quantization is,

\[ S_z = m_s \hbar = \pm \frac{1}{2} \hbar \]  

(2.2.6)

where the number of states is defined by,

\[ \text{# of Electron Spin States} = 2s + 1 = 2 \]  

(2.2.7)
(Note that the electron possesses both orbital $L$ and spin $S$ angular momentum. Angular momentum is additive and is typically denoted by $J = L + S$. We are only concerned with spin angular momentum at this time however.) Similarly, a magnetic nucleus has a nuclear spin angular momentum quantum number denoted by $I$ and a nuclear spin projection angular momentum quantum number denoted $m_I$. The magnitude of the nuclear spin angular momentum is defined by,

$$I = (L_x \hat{i} + L_y \hat{j} + L_z \hat{k})$$ \hspace{1cm} (2.2.8)

where,

$$|I| = \sqrt{I(I + 1)} \hbar$$ \hspace{1cm} (2.2.9)

and the component directed along the axis of quantization is,

$$I_z = m_I \hbar$$ \hspace{1cm} (2.2.10)

where the number of states is also defined by,

$$\text{# of Nuclear Spin States} = 2I + 1$$ \hspace{1cm} (2.2.11)

These relationships are summarized in table 2.2. The visual relationships for all three angular momentum phenomena are compared in figure 2.3. Cone like pictures illustrate that the angular momenta along the quantization axis are known precisely ($L_z, S_z, I_z$), whereas the angular momenta along the other axes ($L_x, L_y, S_x, S_y, I_x, I_y$) are random.

It was shown earlier with equation (2.1.9) that the magnetic moment of a particle due to orbital motion is equal to the gyromagnetic ratio of that particle multiplied by its angular momentum, i.e. $\mu = \gamma L$. It can be assumed that the electron spin and nuclear spin angular momentum behave similarly. Therefore, the magnetic moments generated by an electron and a magnetic nucleus become,

$$\mu_e = -\gamma_e S = g_e \frac{\mu_B}{\hbar} S$$ \hspace{1cm} (2.2.12)

$$\mu_N = \gamma_N I = g_N \frac{\mu_N}{\hbar} I$$ \hspace{1cm} (2.2.13)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
</table>
| Energy | $n$ | Principle QN $n = 0, 1, 2, ...$

<table>
<thead>
<tr>
<th>Orbital Angular Momentum</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
</table>
| $l$ | Orbital Angular Momentum QN | $0, 1, 2, ..., n - 1$
| $m_l$ | Orbital Projection Angular Momentum QN | $-l \leq m_l \leq +l$
| $L$ | Orbital Angular Momentum Vector | $L = (L_x \hat{i} + L_y \hat{j} + L_z \hat{k})$
| $|L|$ | Magnitude of Orbital Angular Momentum | $|L| = \sqrt{l(l + 1)}\hbar$ |
| $L_x$ | Projected Orbital Angular Momentum | $L_x = m_l \hbar$ |

<table>
<thead>
<tr>
<th>Electron Spin Angular Momentum</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
</table>
| $s$ | Electron Spin Angular Momentum QN | $s = \frac{1}{2}$
| $m_s$ | Electron Spin Projection Angular Momentum QN | $m_s = -\frac{1}{2}, \frac{1}{2}$
| $S$ | Electron Spin Angular Momentum Vector | $S = (S_x \hat{i} + S_y \hat{j} + S_z \hat{k})$
| $|S|$ | Magnitude of Electron Spin Angular Momentum | $|S| = \sqrt{s(s + 1)}\hbar$ |
| $S_x$ | Projected Electron Spin Angular Momentum | $S_x = m_s \hbar = \pm \frac{1}{2}\hbar$ |

<table>
<thead>
<tr>
<th>Nuclear Spin Angular Momentum</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
</table>
| $I$ | Nuclear Spin Angular Momentum QN | $I = 0, \frac{1}{2}, \frac{3}{2}, 1, 2, ...$
| $m_I$ | Nuclear Spin Projection Angular Momentum QN | $-I \leq m_I \leq +I$
| $I$ | Nuclear Spin Angular Momentum Vector | $I = (I_x \hat{i} + I_y \hat{j} + I_z \hat{k})$
| $|I|$ | Magnitude of Nuclear Spin Angular Momentum | $|I| = \sqrt{I(I + 1)}\hbar$ |
| $I_x$ | Projected Nuclear Spin Angular Momentum | $I_x = m_I \hbar$ |

Table 2.2: Summary of orbital, electron spin, and nuclear spin angular momentum relationships.

For an isolated electron or magnetic nucleus in a static magnetic field $B$, the potential energy arising from their magnetic moment $\mu$ is defined by

$$E = -\mu \cdot B$$ (2.2.14)

Incorporating the electron and nuclear magnetic moments yields,

$$E_e = -\mu_e \cdot B = \gamma_e S \cdot B = g_e \frac{\mu_e}{\hbar} S \cdot B$$ (2.2.15)

$$E_N = -\mu_N \cdot B = -\gamma_N I \cdot B = -g_N \frac{\mu_N}{\hbar} I \cdot B$$ (2.2.16)
Figure 2.3: Comparison of orbital, electron spin, and nuclear spin angular momentum relationships. The nuclear spin in this case is that of aluminum which is a spin \( I = \frac{3}{2} \). The magnetic field defines the axis of quantization.

If it is assumed that the applied magnetic field is in the \( z \) direction, (i.e., \( B_z \mathbf{k} = B_0 \)), the energies of the electron and magnetic nuclei reduce to,

\[
E_e = g_e \frac{\mu_B}{\hbar} (S_x \hat{i} + S_y \hat{j} + S_z \mathbf{k}) \cdot (B_x \hat{i} + B_y \hat{j} + B_z \mathbf{k}) = g_e \frac{\mu_B}{\hbar} S_z B_0
\]

\[
E_N = -g_N \frac{\mu_N}{\hbar} (l_x \hat{i} + l_y \hat{j} + l_z \mathbf{k}) \cdot (B_x \hat{i} + B_y \hat{j} + B_z \mathbf{k}) = -g_N \frac{\mu_N}{\hbar} l_z B_0
\]

Plugging equations (2.2.6) and (2.2.10) into the above results in,

\[
E_e = g_e \frac{\mu_B}{\hbar} (m_s \hbar)B_0 = g_e \mu_B m_s B_0
\]

\[
E_N = -g_N \frac{\mu_N}{\hbar} (m_l \hbar)B_0 = -g_N \mu_N m_l B_0
\]

It is common in quantum mechanics to use operators to perform specific functions on electron wave functions to extract certain information. One of the more common operators is known as the Hamiltonian denoted by \( \hat{H} \). The purpose of this operator is to extract the energy levels from an electron wave function in a particular state. More commonly, a Hamiltonian operator is applied to a spatially dependent electron wave function \( |\psi(x,y,z,t)\rangle \) to extract energy levels of the electron. (The \( |\cdot\rangle \) notation used is referred to as Bra-Ket which was introduced by Paul Dirac in 1939 to represent orthonormal quantum states [48]. Orthonormal in this case implies that the quantum states, or basis functions, are orthogonal.
and of unit length. $|\cdot\rangle$ is referred to as the ket which represents the wavefunction and $\langle \cdot |$ is referred to as the bra which represents the complex conjugate of the wavefunction. The notation was used because it provided a convenient way for representing quantum states as a linear combination of orthogonal basis functions. The idea is very similar to the idea behind the Fourier series representation of a signal. For example, any periodic signal can be represented by a series of sine and cosine functions, each multiplied by a constant referred to as the Fourier coefficients. Sine and cosine functions are orthogonal to each other and therefore constitute valid basis functions. In the case of quantum mechanics, the electron state may be represented by a linearly combination of basis states. This notation will be important in the discussion of EDMR in the next section.) Applying the operator to the wavefunction,

$$\hat{H}|\psi(x, y, z, t)\rangle = E|\psi(x, y, z, t)\rangle$$

(2.2.21)

where $E$ are the allowed energies of the electron. In other words, the energies are the eigenvalues of the Hamiltonian of an electron in a spatial dependent eigenstate $|\psi(x, y, z, t)\rangle$. This is essentially the Schrödinger equation in its most general form. This energy measure does not extract the energy of the system due to spin however. The energies obtained from the spin wave function are independent to those obtained from spatial electron wave function. Therefore, the total electron wave function is separable for spatial and spin components,

$$|\psi_e\rangle = |\psi(x, y, z, t)\rangle \cdot |s, m_s\rangle$$

(2.2.22)

where $|s, m_s\rangle$ represents the spin state of the electron, which can be in the up or down state,

$$|s, m_s\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle = |\alpha\rangle = |\uparrow\rangle \quad \text{spin up state}$$

(2.2.23)

$$|s, m_s\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = |\beta\rangle = |\downarrow\rangle \quad \text{spin down state}$$

(2.2.24)

The electron spin operators are defined by,

$$\hat{S} = (\hat{\mathbf{s}}_x \hat{i} + \hat{\mathbf{s}}_y \hat{j} + \hat{\mathbf{s}}_z \hat{k})$$

(2.2.25)

and,
\[ \hat{S}_z |s, m_s \rangle = m_z |s, m_s \rangle \]  \hspace{1cm} (2.2.26)

Note the difference between this spin operator and the definition of the spin vector given in equation (2.2.6) that relates the projected spin in the \( z \) direction. For the case of the operator, \( \hbar \) is factored out into the electron spin Hamiltonian \( \hat{H}_e \),

\[ \hat{H}_e = -\bm{\mu} \cdot \bm{B} = g_e \mu_B \vec{S} \cdot \bm{B} \]  \hspace{1cm} (2.2.27)

Applying the Hamiltonian to the electron spin wave function with \( \omega = m \) yields,

\[ \hat{H}_e |s, m_s \rangle = g_e \mu_B \vec{S} \cdot \bm{B} |s, m_s \rangle = g_e \mu_B B_0 \hat{S}_z |s, m_s \rangle = g_e \mu_B B_0 m_s |s, m_s \rangle \]  \hspace{1cm} (2.2.28)

which give the energies of the electron due to spin. Therefore,

\[ \hat{H}_e |\alpha \rangle = E |\alpha \rangle \rightarrow E_\alpha = \frac{1}{2} g_e \mu_B B_0 \]  \hspace{1cm} (2.2.29)

\[ \hat{H}_e |\beta \rangle = E |\beta \rangle \rightarrow E_\beta = -\frac{1}{2} g_e \mu_B B_0 \]  \hspace{1cm} (2.2.30)

These are referred to the electron Zeeman energies, where the difference in energy of the spin up and down states is given by,

\[ \Delta E_e = E_\alpha - E_\beta = g_e \mu_B B_0 \]  \hspace{1cm} (2.2.31)

Similarly, the nuclear spin operators are defined by,

\[ \hat{I} = (\hat{I}_x \hat{I} + \hat{I}_y \hat{I} + \hat{I}_z \hat{I}) \]  \hspace{1cm} (2.2.32)

where,

\[ \hat{I}_z |I, m_I \rangle = m_I |I, m_I \rangle \]  \hspace{1cm} (2.2.33)
Notice again the difference between this operator and equation (2.2.10) that relates the projected nuclear spin in the z direction. For the case of the operator, \( \hbar \) is factored out into the nuclear spin Hamiltonian \( \mathcal{H}_N \),

\[
\mathcal{H}_N = -\mathbf{\mu} \cdot \mathbf{B} = -g_N \mu_N \mathbf{I} \cdot \mathbf{B}
\]  

such that,

\[
\mathcal{H}_N |l, m_i\rangle = -g_N \mu_N (B_x \hat{I}_x + B_y \hat{I}_y + B_z \hat{I}_z) \cdot (\hat{I}_x \hat{I} + \hat{I}_y \hat{I} + \hat{I}_z \hat{I}) |l, m_i\rangle
\]

\[
= -g_N \mu_N B_0 \hat{I}_z |l, m_i\rangle
\]

\[
= -g_N \mu_N B_0 m_i |l, m_i\rangle
\]  

(2.2.35)

The operators and their eigenfunction values are summarized in table 2.3.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbital Operators</td>
<td>$L \psi(x, y, z, t) = (\hat{L}_x \hat{I} + \hat{L}_y \hat{I} + \hat{L}_z \hat{I}) \psi(x, y, z, t)$</td>
</tr>
<tr>
<td>$\hat{L}_z \psi(x, y, z, t) = m_i \psi(x, y, z, t)$</td>
<td></td>
</tr>
<tr>
<td>$\mathbf{S} \psi(x, y, z, t) = (\hat{S}_x \hat{I} + \hat{S}_y \hat{I} + \hat{S}_z \hat{I})</td>
<td>s, m_s\rangle$</td>
</tr>
<tr>
<td>$\mathbf{S}^2</td>
<td>s, m_s\rangle = s(s + 1)</td>
</tr>
<tr>
<td>$\mathbf{S}_z</td>
<td>s, m_s\rangle = m_s</td>
</tr>
<tr>
<td>$\mathcal{H}_e</td>
<td>s, m_s\rangle = g_s \mu_B \mathbf{S} \cdot \mathbf{B}</td>
</tr>
<tr>
<td>$\mathcal{L}</td>
<td>l, m_i\rangle = (\hat{L}_x \hat{I} + \hat{L}_y \hat{I} + \hat{L}_z \hat{I})</td>
</tr>
<tr>
<td>$\mathcal{L}^2</td>
<td>l, m_i\rangle = l(l + 1)</td>
</tr>
<tr>
<td>$\hat{L}_z</td>
<td>l, m_i\rangle = m_i</td>
</tr>
<tr>
<td>$\mathcal{H}_N</td>
<td>l, m_i\rangle = -g_N H_N \mathbf{I} \cdot \mathbf{B}</td>
</tr>
</tbody>
</table>

Table 2.3: Summary of operators.

As a result, the magnetic nuclei will have evenly spaced energy levels, where the nuclear spin \( 2I + 1 \) defines the number of states, or angular momenta configurations there will be for the nucleus. Consider a sodium Na atom with 100% magnetically abundant nuclear spin \( I = 3/2 \). It will have 4 discrete states. Using the notation similar to that of the electron, the nuclear state \( |\text{state}\rangle = |l, m_i\rangle \), can be represented using the four basis states below,
\[ |\eta_1\rangle = \left| \frac{3}{2}, +\frac{3}{2} \right| \]
\[ |\eta_2\rangle = \left| \frac{3}{2}, +\frac{1}{2} \right| \]
\[ |\eta_3\rangle = \left| \frac{3}{2}, -\frac{1}{2} \right| \]
\[ |\eta_4\rangle = \left| \frac{3}{2}, -\frac{3}{2} \right| \] (2.2.36) (2.2.37) (2.2.38) (2.2.39)

Applying the Hamiltonian to these states yields,

\[ \mathcal{H}_N |\eta_1\rangle = E |\eta_1\rangle \rightarrow E_{\eta_1} = -\frac{3}{2} g_N \mu_N B_0 \] (2.2.40)
\[ \mathcal{H}_N |\eta_2\rangle = E |\eta_2\rangle \rightarrow E_{\eta_2} = -\frac{1}{2} g_N \mu_N B_0 \] (2.2.41)
\[ \mathcal{H}_N |\eta_3\rangle = E |\eta_3\rangle \rightarrow E_{\eta_3} = \frac{1}{2} g_N \mu_N B_0 \] (2.2.42)
\[ \mathcal{H}_N |\eta_4\rangle = E |\eta_4\rangle \rightarrow E_{\eta_4} = \frac{3}{2} g_N \mu_N B_0 \] (2.2.43)

These energies are referred to as the nuclear Zeeman energies. The difference between each levels is,

\[ \Delta E_N = g_N \mu_N B_0 \] (2.2.44)

In the cases for both electrons and nuclei, the discrete energy levels vary linearly with the applied magnetic field as illustrated by equations (2.2.31) and (2.2.44), respectively. These energy levels are illustrated in figure 2.4. Note that the energies in figure 2.4 are not drawn to scale. The energy splitting of the nuclear states is significantly smaller because of the smaller nuclear magnetic moment. It will be demonstrated how these two components interact in the discussion of hyperfine interactions.

In electron paramagnetic resonance (EPR) and nuclear magnetic resonance (NMR), one enables transitions from state to another by providing the system with enough energy for one state to transition to the other. The way to provide energy to the system to alter the spin states is to irradiate the sample with photons via electromagnetic radiation. The Planck–Einstein equation relates the energy of a photon \( E_{\text{photon}} \) to its frequency \( v \) by,
Figure 2.4: Comparison of the Zeeman energies of an electron and a magnetic sodium nucleus in the presence of a magnetic field. A single electron will only have two energy states, whereas a single nucleus may contain many energy states. Here, the sodium Na atom has a nuclear spin \( I = \frac{3}{2} \) with 100% abundance.

\[
E_{\text{photon}} = h\nu = \frac{hc}{\lambda}
\]  
\[\text{(2.2.45)}\]

where \( h \) is Planck's constant, \( \lambda \) is the wavelength, and \( c \) is the speed of light. If the energy applied to the system matches the energy difference between spin states for electron or magnetic nucleus,

\[
E_{\text{photon}} = h\nu = \Delta E_e = g_e\mu_B B_0
\]  
\[\text{(2.2.46)}\]

\[
E_{\text{photon}} = h\nu = \Delta E_N = g_N\mu_N B_0
\]  
\[\text{(2.2.47)}\]

then resonance is achieved. Electrons and nuclei are able to absorb the energy of the photons and are therefore able to transition to different spin states.

Many times, the local environment of the unpaired electrons will cause a shift in the resonance condition due to spin orbit coupling. (These will be discussed in more detail in the upcoming sections.) This interaction creates an additional local field \( B_L \) which shift or alter the resonance condition of the free electron as a function of orientation with the applied magnetic field. Incorporating this local field into the Hamiltonian equation \( (2.2.27) \), yields

\[
\mathcal{H}_e = g_e\mu_B B_{\text{eff}} \cdot \hat{S} = g_e\mu_B (B_L + B) \cdot \hat{S}
\]  
\[\text{(2.2.48)}\]
where the spin angular momentum is now quantized along an effective field $B_{eff}$. The goal now: to work the local field into an effective orientation $g$ matrix. Because the $B$ vector is not invertible, linear algebra can be used to factor it out of the parentheses,

$$\mathcal{H}_e = \mu_B B^T \cdot [g_e (I \cdot B)^{-1} B_L + g_e I] \cdot \hat{S}$$

(2.2.49)

where $I$ is the 3x3 identity matrix and the superscript $T$ represents the transpose of a vector. Expanding this equation and realizing that inverting a diagonal matrix is simply computed by inverting the diagonal elements,

$$\mathcal{H}_e = \mu_B [B_x B_y B_z] \cdot \left[ g_e \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} \right]^{-1} \cdot \begin{bmatrix} B_{lx} \\ B_{ly} \\ B_{lz} \end{bmatrix} + g_e \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{s}_x \\ \hat{s}_y \\ \hat{s}_z \end{bmatrix}$$

$$= \mu_B [B_x B_y B_z] \cdot \left[ \begin{bmatrix} B_{lx} / g_e B_x \\ 0 \\ 0 \\ \end{bmatrix} \cdot \begin{bmatrix} 0 \\ B_{ly} / g_e B_y \\ 0 \\ \end{bmatrix} + \begin{bmatrix} g_e \\ 0 \\ 0 \end{bmatrix} \right] \cdot \begin{bmatrix} \hat{s}_x \\ \hat{s}_y \\ \hat{s}_z \end{bmatrix}$$

$$= \mu_B [B_x B_y B_z] \cdot \left[ \begin{bmatrix} g_{xx} \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ g_{yy} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} g_{zz} \end{bmatrix} \right] \cdot \begin{bmatrix} \hat{s}_x \\ \hat{s}_y \\ \hat{s}_z \end{bmatrix}$$

(2.2.50)

Writing in matrix form yields,

$$\mathcal{H}_e = \mu_B B^T \cdot g \cdot \hat{S}$$

(2.2.51)

where $g$ is a diagonal matrix which encompasses the local field information $B_L$ arising from the local environment of the electron. The $g$ value for any orientation can be calculated by

$$g(\theta, \varphi) = \sqrt{g_{xx}^2 \sin^2(\theta) \cos^2(\varphi) + g_{yy}^2 \sin^2(\theta) \sin^2(\varphi) + g_{zz}^2 \cos^2(\theta)}$$

(2.2.52)
where $\theta$ and $\varphi$ define the angles that the magnetic field form with the principal axes. (The principal axes are the ones that lie along the axes of molecular symmetry within the defect center system under investigation.) They are illustrated in figure 2.5,

![Figure 2.5](image)

Figure 2.5: $\theta$ and $\varphi$ define the angles that the magnetic field form with the principle axes.

If the principal axes are unknown, then the $\mathbf{g}$ is defined in its most general form as a 2$^{nd}$ rank tensor,

$$
\mathbf{g} = \begin{bmatrix}
g_{xx} & g_{yx} & g_{zx} \\
g_{xy} & g_{yy} & g_{zy} \\
g_{xz} & g_{yz} & g_{zz}
\end{bmatrix}
$$

(2.253)

With the basic understanding of magnetic resonance reviewed, the next step is to understand how it is measured with hardware.

2.3 – Continuous Wave EPR Practice

The resonance conditions given by (2.2.46) and (2.2.47) imply that there are two parameters that can be adjusted to satisfy the resonance condition of electrons and magnetic nuclei; the externally applied magnetic field and frequency of the electromagnetic radiation. Typically, the frequency is held constant while the magnetic field is swept, and hence the name continuous wave EPR. The experiments are typically performed this way because, simply put, it is much easier to design and implement. In EPR, typical frequencies that are used are in the X-band (9.5 GHz) microwave region. These frequencies are used because, as the frequency is increased, resolution of spectra also increases. However, because microwave cavities are typically utilized, the dimensions decrease as the frequency increases. X-band frequencies are the perfect compromise. At 9.5GHz, the corresponding magnetic field is approximately 3395 G for resonance of electrons. If this field was used for NMR, significantly lower frequencies would be needed to satisfy the resonance condition because of the lower magnetic moment of nuclei. Therefore, frequencies in the radio region are needed, thereby allowing the electromagnetic field to be generated from a resonant circuit which drives a solenoid or surface coil. A simplified continuous wave EPR spectrometer is illustrated in figure 2.6.
The spectrometer is composed of a field controlled electromagnet to provide the quasi static magnetic field $B_0$, and a microwave generator/bridge, waveguide, and cavity to provide the source and path of electromagnetic radiation. As illustrated in figure 2.7, the magnetic field and electric field components of an EM wave are orthogonal to each other and to the direction of propagation. Because these components are perpendicular, the cavities can be designed to maximize either component at the center.

Figure 2.6: Bruker EMX EPR spectrometer.

Figure 2.7: Illustration of (top) electromagnetic wave, (left) a TE$_{102}$ microwave cavity demonstrating a maximum linear polarized magnetic field at the center of the cavity and (right) the magnitude of this field as a function of time. As the figure illustrates, a linearly polarized field is equivalent to two opposite rotating circularly polarized magnetic fields. This concept is important when understanding the rotating reference frame. It is imperative that the two magnetic field components $B_0$ and $B_1$ are perpendicular to each other because of the torque required to observe the EPR effect.
In magnetic resonance, the magnetic field component is of interest because it is the one interacting with
the spin of the electrons. This magnetic field is usually referred to as the $B_1$ oscillating field and is
perpendicular to the quasi static $B_0$ magnetic field. (The reason these fields need to be perpendicular to
each other is because the magnetic moments of the unpaired electrons need to be torqued which is defined
by the differential of angular momentum. This is realized by a continually changing magnetic field in a
direction orthogonal to a static field. This will be further explained and become more apparent in the
upcoming section regarding the Bloch equations.) A transverse electric 102 (TE$_{102}$) microwave cavity,
illustrated in figure 2.7, is typically utilized in the EPR spectrometers because the dimensions of this
cavity allow for a linearly polarized wave to have a maximum magnetic field component at the center
where the sample resides. (The design of TE$_{mnp}$ cavities can be found in various texts [45], [49].) As
illustrated in the figure 2.7, the magnetic field component at the center of the cavity will oscillate at the
microwave source frequency as a function of time. It is shown that this field is equivalent to a left and
right circularly polarized wave. This concept is important when understanding the rotating reference
frame discussed in the next section.

The bridge not only generates the oscillating magnetic field needed to flip the electron spins, but
it is also contains a diode which detects the radiation reflected back from the sample. The current at the
output of the diode is a measure of the reflected microwave power due to absorption of the sample at
resonance. At the beginning of an EPR experiment, a sample is first placed inside the microwave cavity.
Then, the microwave source frequency is tuned to create a standing wave within the loaded microwave
cavity. Next, an iris screw is used to impedance match the loaded cavity with the waveguide. At
resonance, there is a net absorption of EM energy from the unpaired electrons within the sample which
causes an impedance mismatch between the waveguide and cavity. This distorts the standing wave in the
cavity and causes microwaves to be reflected through the waveguide back to the detector diode. Because
the resonance condition is unknown, one usually sweeps the magnetic field to search for resonance.
Therefore, a typically EPR measurements is performed by applying constant microwave radiation while
simultaneously sweeping the magnetic field. Figure 2.8 illustrates the splitting of the two electron energy
states and the net absorption of EM at resonance that occurs for a typical EPR experiment.

The EPR spectrometer also utilizes magnetic field modulation in order to remove flicker noise
from the detector diode in the microwave bridge. This is achieved by sending a sinusoid of audio
frequency through a set of Helmholtz coils mounted on the side of the cavity as illustrated in figure 2.6.
The field that it generates is parallel to the quasi static $B_0$ magnetic field. Therefore, the resonance
condition is modulated such that the information is encoded in a sinusoid of particular frequency and
phase. The signal is then demodulated using lock-in amplification. This technique will be discussed in
more detail in chapter 4.
Figure 2.8: Energy splitting of the two electron spin states in the presence of a magnetic field. When the resonance condition of an electron is met, it is able to transition between states.

### 2.4 – The Bloch Equations

The first demonstration of EPR was performed by Ye. K. Zavoisky in 1944 [1]. Soon after, Felix Bloch developed a set of equations related the microscopic magnetic resonance phenomena into a macroscopic phenomenon [50] [51]. The equations derived below are today known as the Bloch equations and are found it most all EPR and NMR textbooks [41], [43], [44], [45], [52].

In classical mechanics, a magnetic moment experiences a torque when placed in a magnetic field which is defined by,

$$\tau = \frac{dJ}{dt} = \mu \times B$$  \hspace{1cm} (2.4.1)

Using the relationship derived in section 2.1, the magnetic moment $\mu$ and the angular momentum $J$ are related by,

$$J = \frac{\mu}{\gamma}$$  \hspace{1cm} (2.4.2)

the torque becomes,

$$\tau = \frac{d\mu}{dt} = \gamma\mu \times B$$  \hspace{1cm} (2.4.3)
If we assume that the static magnetic field acts solely in the \( \hat{k} \) direction, (i.e. \( \mathbf{B} = B_\gamma \hat{\mathbf{k}} = B_0 \hat{\mathbf{k}} \)), the torque becomes

\[
\tau = \frac{d\mu_x}{dt} \hat{i} + \frac{d\mu_y}{dt} \hat{j} + \frac{d\mu_z}{dt} \hat{k} = \gamma [\mu_x \hat{i} + \mu_y \hat{j} + \mu_z \hat{k}] \times B_0 \hat{k} = \gamma [\mu_y B_0 \hat{i} - \mu_x B_0 \hat{j}] \tag{2.4.4}
\]

Where the components of the torque are equal to

\[
\frac{d\mu_x}{dt} = \gamma \mu_y B_0 \tag{2.4.5}
\]
\[
\frac{d\mu_y}{dt} = -\gamma \mu_x B_0 \tag{2.4.6}
\]
\[
\frac{d\mu_z}{dt} = 0 \tag{2.4.7}
\]

This set of equations are first order ordinary differential equations (ODEs) which are coupled in the \( x \) and \( y \) directions. In order to solve them, take the derivative of equation (2.4.5) and substitute into (2.4.6) which yields,

\[
\frac{d^2\mu_x}{dt^2} = \gamma B_0 \frac{d\mu_y}{dt} = \gamma B_0 (-\gamma \mu_x B_0) = -\gamma^2 B_0^2 \mu_x \tag{2.4.8}
\]

Similarly, take the derivative of (2.4.6) and plug into (2.4.5) to yield,

\[
\frac{d^2\mu_y}{dt^2} = -\gamma B_0 \frac{d\mu_x}{dt} = -\gamma B_0 (\gamma \mu_y B_0) = -\gamma^2 B_0^2 \mu_y \tag{2.4.9}
\]

By doing this, two second order ODEs are obtained and are no longer coupled the \( x \) and \( y \) directions. These equations are easy to solve. For each equation of the form,

\[
\frac{d^2\mu}{dt^2} + \gamma^2 B_0^2 \mu = 0 \tag{2.4.10}
\]

the roots of the characteristic equation can be solved for,
\[ s^2 + \gamma^2 B_0^2 s = 0 \rightarrow s = \sqrt{-\gamma^2 B_0^2} = \pm i\gamma B_0, \]  \hspace{1cm} (2.4.11)

where \( \gamma B_0 \) is known as the Larmor frequency and is denoted by

\[ \omega_0 = \gamma B_0 \]  \hspace{1cm} (2.4.12)

The solution for both differential equations is of the form,

\[ \mu(t) = A_1 e^{i\gamma B_0 t} + A_2 e^{-i\gamma B_0 t} \]  \hspace{1cm} (2.4.13)

Using the Euler identity,

\[ e^{i\theta} = \cos(\theta) + i\sin(\theta) \]
\[ e^{-i\theta} = \cos(\theta) - i\sin(\theta) \]  \hspace{1cm} (2.4.14)

the solution for the magnetic moment is obtained by,

\[ \mu(t) = A_1 (\cos(\gamma B_0 t) + i \sin(\gamma B_0 t)) + A_2 (\cos(\gamma B_0 t) - i \sin(\gamma B_0 t)) \]
\[ = (A_1 + A_2) \cos(\gamma B_0 t) + i (A_1 - A_2) \sin(\gamma B_0 t) \]
\[ = C_1 \cos(\gamma B_0 t) + C_2 \sin(\gamma B_0 t), \]  \hspace{1cm} (2.4.15)

where the magnetic moment components are,

\[ \mu_x(t) = C_1 \cos(\gamma B_0 t) = \mu_x(0) \cos(\gamma B_0 t) \]  \hspace{1cm} (2.4.16)
\[ \mu_y(t) = C_2 \sin(\gamma B_0 t) = \mu_y(0) \sin(\gamma B_0 t) \]  \hspace{1cm} (2.4.17)
\[ \mu_z(t) = \text{constant} \]  \hspace{1cm} (2.4.18)

These responses indicate that the magnetic moments in the transverse \((x,y)\) plane are oscillating. Together, the responses actually represent a precession of the magnetic moments about the static magnetic field \( B_0 \) at the Larmor frequency. This is pictorially demonstrated in figure 2.9.

It was shown earlier that the magnetic moment of an electron was related to the spin angular momentum by, \( \mu_x = g_e \gamma e \mathbf{S} \). It was also shown that \( |\mathbf{S}| = \sqrt{s(s+1)} \hbar \) and \( S_z = m_s \hbar \). Combing these terms allows one to solve for the magnetic moment components of the precessing electron,
\begin{align}
|\boldsymbol{\mu}| &= g_e \gamma_e |\mathbf{S}| = g_e \gamma_e \sqrt{s(s + 1)} \hbar = \frac{\sqrt{3}}{2} g_e \mu_B \\
\mu_z &= g_e \gamma_e S_z = g_e \gamma_e m_s \hbar = \pm \frac{1}{2} g_e \mu_B
\end{align}

(2.4.19) \quad (2.4.20)

where \( s = 1/2 \), \( m_s \pm 1/2 \) for electrons. The angle \( \theta \) defines the average angle of precession is calculated by,

\[
\cos(\theta) |\boldsymbol{\mu}| = \mu_z \rightarrow \theta = \cos^{-1} \left( \frac{\mu_z}{|\boldsymbol{\mu}|} \right) = \cos^{-1} \left( \frac{1}{\sqrt{3}} \right) = 54.7^\circ
\]

(2.4.21)

This angle is typically referred to the magic angle in NMR. The Bloch equations do not hold for individual spins because this is a classical analysis, but they do hold for a collection of magnetic moments due to spin.

![Illustration of electron magnetic moments precessing about the axis of the applied magnetic field due to spin in the up and down states.](image)

Figure 2.9: Illustration of electron magnetic moments precessing about the axis of the applied magnetic field due to spin in the up and down states.

The Boltzmann distribution states that, in thermal equilibrium, the population ratio of lower energy spins \( N_+ \) to higher energy spins \( N_- \) is related by,

\[
\frac{N_+}{N_-} = \exp \left( - \frac{E}{kT_s} \right) = \exp \left( - \frac{\hbar \nu}{kT_s} \right) = \exp \left( - \frac{g_e \mu_B B_0}{kT_s} \right),
\]

(2.4.22)

where \( E \) is the electromagnetic energy applied to the system needed to flip the electron spin, \( k \) is Boltzmann’s constant \( (k = 1.3805 \times 10^{-23} \text{ J/K}) \), and \( T_s \) is the spin temperature. The spin temperature is a way to distinguish the temperature of the spin dependent particles (spin system) from the temperature of
the surrounding lattice. At resonance when the electron spins are flipping, Fermi’s Golden rule states that the number of electrons in the up state will flip to the down state at the same rate as the electrons in the down state flip to the up state [43] [44]. This implies that if the populations of up spins and down spins were the same, there wouldn’t be a net change in absorbed energy, and hence, no signal would be detected. In EPR, one actually detects the difference between the energy of the electrons in the lower energy state absorbing EM energy to transition to the high energy state and that of the electrons in the higher energy state emitting EM energy to transition to the lower energy state. (The latter transition is usually accompanied by the release of phonons and or heat). Because the number of states in the low energy state slightly outnumbers the number of spins in the high energy state according to equation (2.4.22), a net absorption of EM radiation will result. It is desired to make this population difference as large as possible to increase the effect. The two ways to accomplish this is to lower the temperature and or increase the frequency of the EM radiation applied to the system (thereby increasing the field needed for resonance which also increases the polarization of the spins.) Therefore, it is precisely the population difference that limits the sensitivity of conventional EPR. Commercial EPR spectrometers with X-band microwave cavities typically have sensitivities around $10^{10}$ spins. The sensitivity, or minimum detectable spins $N_{\text{min}}$, for an EPR measurement in a microwave cavity is define by [45],

$$N_{\text{min}} \approx \frac{V}{Q_u \eta \omega_0^2 P^{1/2}}$$

(2.4.23)

Where $V$ is the sample volume, $Q_u$ is the unloaded cavity, $\eta$ is the cavity filling factor, $\omega_0$ is the microwave frequency in rad/s, and $P$ is the power of the microwaves. From the equation, it is obvious that the maximum sensitivity can be achieved by reducing the volume size as well as increasing power and frequency of the measurement. However, the implementation of these rules into a physical piece of hardware is not so trivial.

In the absence of an external magnetic field, the electron spins do not have an axis of quantization which implies that their magnetic moments are oriented randomly. However, in the presence of an external magnetic field, the electron magnetic moments are all precessing about the axis of the applied field at different phases in the $+\vec{k}$ (lower energy state) and $-\vec{k}$ (higher energy state) directions. If one were to add the individual magnetic moments, one would get a net magnetic moment in the $+\vec{k}$. The reason for this is because there are more magnetic moments pointed in the $+\vec{k}$ direction than in the $-\vec{k}$ direction. Also, because the moments are precessing at different phases, the transverse components cancel. This net magnetic moment due to collective spins is referred to as the magnetization. This vector is defined by,
Figure 2.10: Illustrating the orientations of individual magnetic moments due to spin of unpaired electrons in the (top) absence and (bottom) presence of an externally applied magnetic field. In the absence of a magnetic field, the angular momentum quantization axis is undefined so the magnetic moments are pointing in random directions. In the presence of a magnetic field, the electron spins precess about or against the applied magnetic field. Summing these spins results in a magnetization in the direction of the applied field because the transverse components cancel and there are a net positive spins in the lower energy state.

\[
B = \sum_i \mu_i = M_0 \hat{k}
\]  

(2.4.24)

where \(M_0\) is the total magnetization. Figure 2.10 illustrates this concept. The magnetization vector also experiences a torque when in the presence of an externally applied magnetic field. Substituting the magnetization vector in for the individual magnetic moment vector in equation (2.4.1) yields,

\[
\frac{dM}{dt} = \gamma (M \times B) \\
= \gamma \left( M_x \hat{i} + M_y \hat{j} + M_z \hat{k} \right) \times \left( B_x \hat{i} + B_y \hat{j} + B_z \hat{k} \right) \\
= \gamma \left( M_y B_z - M_z B_y \right) \hat{i} + \gamma \left( M_z B_x - M_x B_z \right) \hat{j} + \gamma \left( M_x B_y - M_y B_x \right) \hat{k}
\]  

(2.4.25)

However, the spins are constantly interacting with themselves and with the lattice, therefore, the coupled differential equations become,
where, $T_1$ is the longitudinal (spin lattice) relaxation time constant, and $T_2$ is the transverse relaxation time constant. (These time constants are usually easier to think about in a pulsed measurement. $T_1$ is a temperature dependent time that refers to the time constant at which the magnetization recovers in the longitudinal direction after being rotated completely off the quantization axis and $T_2$ corresponds to the time at which the individual spin magnetic moments dephase in the transverse components.) These coupled differential equations are commonly known as the general Bloch equations. These equations are extremely difficult to solve analytically as they appear in equations (2.4.26)-(2.4.28) with the additional of an oscillating magnetic field component. In the next section, these equations will be solved in the rotating reference frame.

**2.5 – The Rotating Reference Frame**

In the previous section, it was shown that the magnetization vector due to collective magnetic moments due to spin could be represented mathematically by a set of three coupled differential equations. However, the magnetization does not experience a torque unless there is a constantly changing magnetic field as indicated by the cross product $\tau = \gamma (\mathbf{M} \times \mathbf{B})$. As discussed in section 2.4 and figure 2.7, the $B_1$ magnetic field component of the linearly polarized electromagnetic wave is perpendicular to the quasi-static sweeping magnetic field $B_0$ component. This will enable the magnetization to be torqued, allowing an EPR effect to be observed. The magnetic field of this linearly polarized wave is represented as a function of time by $B_1(t)$,

$$B_1(t) = 2B_1 \cos(\omega t)\hat{i}$$  

where $\omega = 2\pi v$ and $v$ is the frequency of the microwave radiation and the constant $B_1$ represents the amplitude of this oscillating magnetic field, typically about a few tenths of Gauss or less. It was also noted earlier that this linearly polarized wave can be decomposed into two circularly polarized waves whose magnetic field components are rotating in right and left directions, $B_R(t)$ and $B_L(t)$ respectively (see figure 2.6). These oscillating magnetic field components are mathematically represented by
\[ B_R(t) = B_1 \cos(\omega t) \hat{i} + \sin(\omega t) \hat{j} \]  
\[ B_L(t) = B_1 \cos(\omega t) \hat{i} - \sin(\omega t) \hat{j} \]  

Utilizing one of the rotating components, the total magnetic field becomes,

\[ B = B_0 \hat{k} + B_1 \cos(\omega t) \hat{i} + \sin(\omega t) \hat{j} \]  

and the torque on the magnetization becomes,

\[ \frac{dM}{dt} = \gamma M \times B = \gamma M \times \left[ B_0 \hat{k} + B_1 \cos(\omega t) \hat{i} + \sin(\omega t) \hat{j} \right] \]  

Now assume that we were to rotate with this oscillating field component. The rotation rate of the new axes would be defined by frequency of the oscillating field \( \omega = 2\pi v \). In this coordinate transformation, \( \hat{k} \) axis remains unchanged, however, \( \hat{i} \rightarrow \hat{i}' \) and \( \hat{j} \rightarrow \hat{j}' \). The laboratory and rotating reference frames are illustrated in figure 2.11.

Within the rotating axis, the static magnetic field in the \( \hat{k} \) remains constant as it did in the laboratory frame because the frame is rotating about that axis. Because this frame is rotating, the magnetization components would be fixed to the rotating axes \( M_x \rightarrow M_{x'} \) and \( M_y \rightarrow M_{y'} \), and \( M_z \) remains constant. In order to understand the effects of the oscillating field, we must mathematically represent the field components in the \( \hat{i}, \hat{j} \) plane, one needs to multiply through by \( \gamma \) and subtract the frequency of rotation \( \omega \) from each component account for the rotating reference frame.

Figure 2.11: Illustration of the laboratory frame and rotating reference frame. In the laboratory reference frame, the \( \hat{i} \) and \( \hat{j} \) axes are fixed and the field rotates. In the rotating reference frame, the \( \hat{i} \) and \( \hat{j} \) axes rotate, the \( \hat{i}' \) and \( \hat{j}' \) axes are fixed, and the field remains constant along \( \hat{i}' \).
\[
\frac{dM}{dt} = \mathbf{M} \times \left[ (\gamma B_0 - \omega) \hat{k} + \gamma B_1 \left[ \cos( (\omega - \omega) t') \hat{i}' - \sin( (\omega - \omega) t') \hat{j}' \right] \right]
\]
\[
= \mathbf{M} \times \gamma \left[ (B_0 - \frac{\omega}{\gamma}) \hat{k} + B_1 \hat{i}' \right]
\]
(2.5.6)

This equation shows, an effective field is created,

\[
\mathbf{B}_{\text{eff}} = B_z \hat{k} + B_x \hat{i}' = \left( B_0 - \frac{\omega}{\gamma} \right) \hat{k} + B_1 \hat{i}'
\]
(2.5.7)

Where the angle of the effective field is calculated by,

\[
\theta_{\text{eff}} = \tan^{-1} \left( \frac{B_z}{B_x} \right) = \tan^{-1} \left( \frac{\gamma B_0 - \omega}{\gamma B_1} \right) = \tan^{-1} \left( \frac{\omega_0 - \omega}{\gamma B_1} \right)
\]
(2.5.8)

The effective field will spend most of the time aligned with or against the large magnetic field \(B_0\). Only when the sweep approached resonance does the effective field point in the direction of \(B_1\) in the rotating reference frame. To understand how the resonance signal is detected, first consider the situation when the magnetization follows the effective field. The projection of this magnetization onto the \(\hat{i}', \hat{j}'\) plane is the precisely the ESR signal. In this case, the only component is in the \(\hat{i}'\) direction.

\[
M_{x'} = M_0 \cos(\theta_{\text{eff}}) = M_0 \cos \left( \tan^{-1} \left( \frac{\omega_0 - \omega}{\gamma B_1} \right) \right)
\]
(2.5.9)

The \(\cos(\cdot)\) function is used because the waveguide is only sensitive to detection in the \(\hat{i}', \hat{j}'\) plane due to the way EM waves propagate within the cavity. As the magnetic field is swept at the beginning of a scan, the Larmor frequency \(\omega_0\) is slower than the EM radiation frequency \(\omega\). As a result, the projection of magnetization in the \(\hat{i}', \hat{j}'\) plane is near zero. At resonance, \(\omega_0 = \omega\), and therefore \(M_{x'} = M_0\). The projection of magnetization in the \(\hat{i}', \hat{j}'\) plane is therefore maximized. As the sweeping field passes resonance, the precession frequency of the electrons becomes greater than the frequency of the EM radiation. As a result, the projection of magnetization in the \(\hat{i}', \hat{j}'\) plane is close to zero again. These equations are plotted in figure 2.12 for a resonance field of 3400 G with \(B_1 = 1\)G. If we switch back to the laboratory frame, we see that the magnetization rotating in the \(\hat{i}\) direction appears to be circularly polarized and growing in amplitude. For X-band measurements, this is detected by an impedance mismatch in the cavity and waveguide due to a net absorption of the electrons through resonance.
In low field measurements, this oscillating field can be detected by a receiver coil by the EMF that it produces, through Faraday’s Law of Induction as illustrated by figure 2.13. The detection of the magnetization can also be understood by solving the Bloch equations under steady state conditions. This actually provides a more accurate representation of the magnetization components. Because the field in the \( \hat{i}' \) direction is constant and the field in the \( \hat{j}' \) direction is zero, the analysis becomes significantly easier in the rotating reference frame. Equations (2.4.26)-(2.4.28) can be rewritten in the rotating reference frame as,

\[
B_{eff} = B_0 \hat{k} + B_{y'} \hat{j}' + B_{x'} \hat{i}'
\]

Figure 2.12: Illustration of the effective field and projection of the magnetization in the \( \hat{i}' \) directions.

Figure 2.13: Illustration of magnetization components in the laboratory reference frame.
\[ \frac{dM_{x'}}{dt} = -\gamma M_{y'}B_z - \frac{M_{x'}}{T_2} \quad (2.5.10) \]
\[ \frac{dM_{y'}}{dt} = \gamma (M_xB_z - M_{x'}B_z) - \frac{M_{y'}}{T_2} \quad (2.5.11) \]
\[ \frac{dM_z}{dt} = -\gamma M_{y'}B_z - \frac{M_z - M_0}{T_1} \quad (2.5.12) \]

where,
\[ B_z = B_0 - \frac{\omega}{\gamma} = \frac{\omega_0 - \omega}{\gamma} \quad (2.5.13) \]
\[ B_{x'} = B_1 \quad (2.5.14) \]

Assume steady state conditions (the change in magnetization is small over the differential time), the differential magnetization terms can be approximated with zeros,
\[ \frac{dM_{x'}}{dt} = \frac{dM_{y'}}{dt} = \frac{dM_z}{dt} = 0 \quad (2.5.15) \]

And substituting in \( B_z \) and \( B_{x'} \), equations (2.5.10) - (2.5.12) can be rewritten as,
\[ M_{y'}\Delta\omega = \frac{M_{x'}}{T_2} \quad (2.5.16) \]
\[ M_{x'}\Delta\omega - M_z\omega_1 = -\frac{M_{y'}}{T_2} \quad (2.5.17) \]
\[ M_{y'}\omega_1 = -\frac{M_z - M_0}{T_1} \quad (2.5.18) \]

where,
\[ \Delta\omega = \omega_0 - \omega \quad (2.5.19) \]
\[ \omega_1 = \gamma B_1 \quad (2.5.20) \]

\( \omega_1 \) is the precession frequency of the magnetic moment about the \( B_1 \) field. Note that the magnetic moments are attempting precess about \( B_0 \) and \( B_1 \) simultaneously. This is commonly referred to as a nutation. This results in the magnetization vector that precess in a cone in the rotating reference frame.
about $B_{\text{eff}}$ as illustrated in figure 2.14. Solving these three equations with three unknowns, yields the following magnetization components in the rotating reference frame,

\[ M_{x'} = M_0 \cdot \frac{\omega_1 \Delta \omega T_2^2}{1 + (\Delta \omega T_2)^2 + \omega_1^2 T_1 T_2} \]  
(2.5.21)

\[ M_{y'} = M_0 \cdot \frac{\omega_1 T_2}{1 + (\Delta \omega T_2)^2 + \omega_1^2 T_1 T_2} \]  
(2.5.22)

\[ M_z = M_0 \cdot \frac{1 + (\Delta \omega T_2)^2}{1 + (\Delta \omega T_2)^2 + \omega_1^2 T_1 T_2} \]  
(2.5.23)

These magnetization components are simulated in figure 2.14 using $T_1 = T_2 = 1\mu s$, $B_1 = 0.3 G$, $\nu = 9.5 GHz$, a sweep width of $10 G$, and $g = 2.0023$.

Figure 2.14: (left) precession of the magnetization vector in a cone and (right) plot of the magnetization components through resonance.

### 2.6– Spin Orbit Coupling

Up to this point, only angular momenta due to electron and nuclear spin were taken into account. As a result, the electrons and magnetic nuclei have magnetic moments associated with them, and therefore can be envisioned as microscopic bar magnets as illustrated in figure 2.15. A more precise treatment would take into account the orbital angular momentum arising from the orbital motion of the electron about a nucleus. To further understand this concept, first envision an unpaired electron with quantized orbital angular momentum $L = n\hbar$ orbiting a nucleus. To the electron, it appears as though the nucleus is orbiting around it. This will create an effective field that will couple with the electron spin if large enough. This concept is also illustrated in figure 2.15.
Figure 2.15: In the Bohr model of an atom, an electron orbits the nucleus. (left) Intrinsic angular momentum of electron and nuclei give rise to electron and nuclear magnetic moments. (right) To the electron, the nucleus appears to be orbiting around it. The orbital angular momentum of the electron due to its motion about the nucleus give rise to an orbital magnetic moment.

It was mentioned earlier that a particle of mass $m$, exhibits orbital angular momentum,

$$L = rmv$$  \hspace{1cm} (2.6.1)

if travelling at a constant velocity $v$ on a circular path of radius $r$. If the particle also has charge $q$ then the current it produces is defined by the charge per unit time,

$$i = Zq \frac{v}{2\pi r} = \frac{ZqL}{2\pi mr^2}$$  \hspace{1cm} (2.6.2)

where $r$ is the radius of the effective orbit, and $Zq$ is the nuclear charge of the orbiting particle. The field generated by this current loop at the center can be calculated from the Biot Savart Law,

$$dB = \frac{\mu_0 i \, dl \times r}{4\pi r^3}$$  \hspace{1cm} (2.6.3)

where $dl$ represents the differential length around the current loop. Integrating over the loop using polar coordinates and evaluating at the center, yields,

$$B = \int dB = \frac{\mu_0}{4\pi} i \frac{2\pi r}{r^2} = \frac{\mu_0}{4\pi} \left( \frac{ZqL}{2\pi mr^2} \right) \frac{2\pi r}{r^2} = \frac{\mu_0 q}{4\pi m} \left( \frac{ZL}{r^3} \right)$$  \hspace{1cm} (2.6.4)
This classical derivation of the field generated by the nuclear and felt by the electron illustrates that the effect depends upon the charge of the particle, the distance between the electron and the nucleus, and the angular momentum. To treat the spin orbit coupling in a fully quantitative way, second order perturbation theory must be utilized. This is a very detailed derivation so will not be discussed here. However, it has been shown that the Hamiltonian due to this orbital coupling is defined by [43],

\[ H_L = H_{LS} + H_{LB} = \lambda \mathbf{L} \cdot \mathbf{S} + \mu_B \mathbf{L} \cdot \mathbf{B} \] (2.6.5)

where,

\[ \lambda = \left[ \frac{Z^4 q^2 r^3}{8\pi \epsilon_0 c^2 m_e^2 n^3 \ell (\ell + \frac{1}{2})(\ell + 1)} \right] \] (2.6.6)

\( \lambda \) is a constant which depends upon the charge of the nucleus, the distance \( r \) between the electron and nucleus, and the orbital angular momentum quantum number. The parameter dependencies are not the same, but are similar to those derived with the classical analogy.

It should be noted that spin orbit coupling is also an orientation dependent effect. Therefore, the \( g \) value measured will vary as a function of the orientation of the sample with respect to the external magnetic field. This is why the \( g \) is represented by an orientation dependent matrix. Note that the Hamiltonian is composed of an orbit /spin interaction and an orbit/field interaction. In this study, only the spin interactions will be considered.

\[ H_{LS} = \lambda \mathbf{L} \cdot \mathbf{S} \] (2.6.7)

The measured spin-orbit coupling is usually small. To first order, the interaction is quenched in solids and appears only because the applied field mixes in excited states.

### 2.7 – Hyperfine Interactions

It was shown earlier that nuclei are magnetic if they have an odd number of protons and or neutrons [41]. If these magnetic nuclei are located in the vicinity of the unpaired electron, the resonance condition will be modified because of the magnetic fields they exert on the unpaired electrons. These modifications to the EPR response will almost always aid in the defect identification. The interactions of electrons and nuclei are commonly referred to as electron nuclear hyperfine interactions. An extensive
discussion will be provided to allow for the understanding of this phenomenon because it is precisely these interactions that have been observed for the first time in our work regarding zero-field spin dependent recombination which is discussed in the next section [18]. To understand why hyperfine interactions can be detected via spin dependent recombination, one first needs to understand how they are conventionally detected using EPR. Once again, the details of hyperfine interactions described in the following section is well documented in various text books [41] [43] [44] [45] [40].

It was shown earlier that the energy of an electron due to its magnetic moment in a magnetic field is \( E = -\vec{\mu} \cdot \vec{B} \). The electron will interact with or the applied field and the nuclear field \( \vec{B}_N \),

\[
E_e = -\vec{\mu} \cdot \vec{B}_{eff} = g_e \mu_B \vec{S} \cdot (\vec{B} + \vec{B}_N). \tag{2.7.1}
\]

where \( \vec{B} = B_0 \hat{k} \). There are two main cases in which an electron can experience a nuclear interaction: the nucleus at the atomic site at which the electron primarily resides or the nucleus at a neighboring atomic site. The former will be analyzed here for s and p orbitals.

First consider an unpaired electron residing in an s orbital of a magnetic nucleus as illustrated in figure 2.16. The field generated from this nucleus was given earlier and is restated here for convenience,

\[
\vec{B}_N(r) = \frac{\mu_0}{4\pi} \left\{ \frac{3(\vec{\mu}_N \cdot \vec{r})\vec{r} - \vec{\mu}_N \vec{r}^2}{r^5} \right\} \tag{2.7.2}
\]

The field outside of the nucleus averages to zero because for every field line pointing up, there is a corresponding field line pointing down. However, precisely at the center of the nucleus, all field lines are pointing in one direction. Therefore, because an electron has the highest probability density at the center of the orbital where the nucleus resides, one only needs to consider the average field the experience by the electron at the center of this orbital To find this field, one can use the analogy used in section 2.2 of a charged particle orbiting about an axis. The magnetic moment generated is defined by equation (2.1.3) \( \vec{\mu} = iA \hat{k} \) and it was shown in section 2.2 that the discretized magnetic moment due to a magnetic nuclei is defined by \( \mu_N = g_N \mu_N \hat{I} \). Using these relationships, the effective current that is generated from the magnetic nucleus is,

\[
i = \frac{\mu}{A} = \frac{g_N \mu_N \hat{I}}{\pi r^2} \tag{2.7.3}
\]
The field generated by this current loop at the center can be calculated from the Biot Savart Law, as was done in section 2.6 when dealing with the effects of spin orbit coupling. The differential field is defined by,

\[ dB = \frac{\mu_0 i \, dl \times r}{4\pi r^3} \]  

(2.7.4)

where \( dl \) represents the differential length around the current loop. Integrating over the loop using polar coordinates, plugging in for the current, and evaluating at the center of the loop, yields,

\[ B_{NI} = \int dB = \frac{\mu_0}{4\pi} \frac{2\pi r}{r^2} = \frac{\mu_0 g_N \mu_N I}{4\pi} \frac{2\pi r}{r^2} = \frac{\mu_0}{2\pi r^3} g_N \mu_N I \]  

(2.7.5)

The average field that the electron experiences from the magnetic nucleus can be found by multiplying the nuclear field at the center of the loop \( B_{NI} \) by the average probability that the electron spends at the center of the nucleus. This probability is simply the electron probability density \( |\psi^*(0)\psi(0)| = |\psi^2(0)| \) multiplied by density of occupancy, or simply the volume of the s orbital \( V_s = 4\pi r^3/3 \). Therefore, the average field the electron in an s orbital experiences is,

\[ B_N = B_{NI} \cdot |\psi^2(0)| \cdot V_s \]

\[ = \left( \frac{\mu_0}{2\pi r^3} g_N \mu_N I \right) |\psi^2(0)| \left( \frac{4}{3} \pi r^3 \right) \]

\[ = \frac{2}{3} \mu_0 g_N \mu_N |\psi^2(0)| \hat{I} \]  

(2.7.6)

Plugging this relation into (2.7.1) yields,

\[ E_e = -\mu \cdot B_{\text{eff}} \]

\[ = g_e \mu_b \hat{S} \cdot B + \frac{2}{3} \mu_0 g_e \mu_b g_N \mu_N |\psi^2(0)| \hat{S} \cdot \hat{I} \]

\[ = g_e \mu_b \hat{S}_z B_0 + \frac{2}{3} \mu_0 g_e \mu_b g_N \mu_N |\psi^2(0)| \hat{S}_z \hat{I}_z \]

\[ = m_s g_e \mu_b B_0 + \frac{2}{3} \mu_0 g_e \mu_b g_N \mu_N |\psi^2(0)| m_s m_l \]

\[ = m_s g_e \mu_b B_0 + A_s m_s m_l, \]  

(2.7.7)
where,

\[
A_s = \frac{2}{3} \mu_0 g_e \mu_\beta g_N \mu_N |\psi^2(0)| \tag{2.7.8}
\]

is known as the Fermi contact constant. Note that no matter the orientation of the sample with respect to the externally applied magnetic field, the resultant interaction will always be the same. This is not the case for p orbitals however. Consider again figure 2.16. In this case, one must evaluate the dipole field averaged over the space in which the electron resides, ie: the p orbital wavefunction.

\[
E_e = \mu \cdot B_{\text{eff}}
= g_e \mu_\beta \vec{S} \cdot \vec{B} + g_e \mu_\beta \vec{S} \cdot \vec{B}_N \tag{2.7.9}
\]

Using equation (2.7.2),

\[
E_e = g_e \mu_\beta \vec{S} \cdot \vec{B} + \frac{\mu_0 \mu_e}{4\pi} \frac{r^2 - 3 \mu_e \cdot r (\mu_N \cdot r)}{r^5} - 3 g_e \mu_\beta \vec{S} \cdot (g_N \mu_N \vec{l} \cdot \vec{r})
= g_e \mu_\beta \vec{S}_z B_0 + \frac{\mu_0 \mu_e}{4\pi} \frac{r^2 - 3 g_e \mu_\beta g_N \mu_N \vec{S}_z \vec{l}_z \cos(\theta) r^2}{r^5}
= m_s g_e \mu_\beta B_0 + \frac{\mu_0}{4\pi} g_e \mu_\beta g_N \mu_N m_s m_l \frac{1 - 3 \cos(\theta)}{r^3}
= m_s g_e \mu_\beta B_0 + A_p m_s m_l \tag{2.7.10}
\]

Where

\[
A_p = \frac{\mu_0}{4\pi} g_e \mu_\beta g_N \mu_N \left(1 - 3 \cos(\theta)\right) \tag{2.7.11}
\]

and the notation \(\langle \cdot \rangle\) represents the average nuclear field the electron experiences over the space in which it resides. In both s and p orbital cases, the energy associated with the magnetic moments are due to the electron spin/field interaction \(\mathcal{H}_{SB}\) and the electron spin/nuclear spin interaction \(\mathcal{H}_{SI}\). Applying the Hamiltonian of the form,
Figure 2.16: Dipole field generated from a magnetic nucleus with an unpaired electron in an s and p (two orientations) orbitals.

\[ H = H_{SB} + H_{SI} = g_e \mu_B \mathbf{S} \cdot \mathbf{B} + A \mathbf{S} \cdot \mathbf{l} \quad (2.7.12) \]

to the electron spin wavefunction,

\[ \hat{H} |s, m_s\rangle = E |s, m_s\rangle \quad (2.7.13) \]

results in energies of the general form,

\[ E = g_e \mu_B B_0 m_s + A m_s m_l = m_s g_e \mu_B (B_0 + A m_l) \quad (2.7.14) \]

where the local field of the unpaired electron is defined by

\[ B_{local} = (B_0 + A m_l) \quad (2.7.15) \]

where \( A \) is called the hyperfine coupling constant. (Note that the full Hamiltonian also takes into account the nuclear spin / field interaction. It is omitted here because it is generally small) These first order calculations hold for \( s = 1/2 \) and \( l = 1/2 \). For spin systems greater than this, second order corrections are needed due the quadrupole coupling involved with atoms of spins \( I > 1/2 \) [52]. Many times, the hyperfine coupling constant is an orientation dependent parameter which results from wave functions.
composed of $s$ AND $p$ like character. Therefore, as was done in section 2.2 with the orientation $g$ value, the hyperfine coupling constant can also be represented as a second rank tensor,

$$
A = \begin{bmatrix}
A_{xx} & A_{xy} & A_{xz} \\
A_{yx} & A_{yy} & A_{yz} \\
A_{zx} & A_{zy} & A_{zz}
\end{bmatrix}
$$

(2.7.16)

Where the Hamiltonian becomes,

$$
\mathcal{H} = \mu_\beta \vec{B} \cdot \vec{g} \cdot \vec{S} + \vec{S}^T \cdot A \cdot \hat{I}
$$

(2.7.17)

Because there could be more than one magnetic nuclei nearby the unpaired electron, one must consider the addition of $i$ nearest neighbor magnetic nuclei,

$$
\mathcal{H} = \mu_\beta \vec{B} \cdot \vec{g} \cdot \vec{S} + \sum_i \vec{S}^T \cdot A_i \cdot \hat{I}_i
$$

(2.7.18)

where the index $i$ refers to the hyperfine interaction of the $i$th nucleus with the electronic moment, $\hat{I}_i$ represents the $i$th nuclear spin operator for the different nuclei, and $A_i$ are the hyperfine parameters for those nuclei. Note that the matrix $A$ contains information about the isotropic and anisotropic nature of the electron nuclear hyperfine interactions.

In order to understand how magnetic nuclei affect the resonance line shape detected, consider the imaginary defect system illustrated in figure 2.17 as a simple example. In the figure, the imaginary material system is composed of 8 sites. Each of these sites is composed of a cluster of green atoms (100% abundant, spin $s = 0$ nucleus) with a single purple atom (25% abundant, spin $s = 1$ nuclei) in the vicinity of each of a cluster. Of the 8 sites of clusters, 4 of them have defects indicated by the red electron in a dangling bond. As a result, only sites (2), (4), (5), and (7) are EPR active. Site (5) will have a resonance that is unaffected by the nearby nucleus because it is not magnetic in this case (the purple atom’s nucleus will not be magnetic 25% of the time). Site (2), (4), and (7) will have altered resonance conditions because of the nearby magnetic nuclei. Because the purple atom has a spin $l = 1$ nucleus, there will be $2l + 1 = 3$ configurations of angular momentum denoted by $m_l$, where $m_l \in \{-1,0,1\}$. Site (2) experiences a local field from nearby purple atom in the $m_l = -1$ state. As a result, the unpaired electron will have a higher resonant field because the field generated by the $m_l = -1$ nucleus is directed against the externally applied magnetic field. The resonance field will be higher
because the electron needs to see *additional external field* to compensate for the component that was canceled by the magnetic nuclei. Defect site (7) is similar to that of site (2), except that it is in the presence of a magnetic nucleus in the $m_I = +1$ state. For this case, the electron will have a lower resonant field because it takes *less external field* to satisfy its resonance condition. And finally, at defect site (4), the resonant condition is the same as it would be if there were no magnetic nuclei nearby. The reason for this is because the unpaired electron is near a magnetic nucleus in the $m_I = 0$ state. As a result, the magnetic nucleus projects zero angular momentum along the axis of quantization and therefore, does not couple to the electron spin. Therefore, defect site (4) will have the same resonance condition as defect site (5). Note that all other sites are EPR inactive because there are no unpaired electrons. As illustrated in figure 2.18, the EPR spectrum will consist of 4 total lines (one from each paramagnetic site in this example). Two of the lines will have the same resonance so their contribution will add resulting in a 3 line spectrum.

---

**Figure 2.17:** Imaginary defect system to illustrate hyperfine interactions.
Figure 2.18 illustrates the energy level diagram for the imaginary system illustrated in figure 2.17 and the spectrum observed.

The center line would represent of 50% of the total spectrum and each of the outer lines represent of 25% of the total spectrum. Each of the peaks are separated by the hyperfine coupling constant $A$. This parameter defines the magnetic field at which the resonance condition of the unpaired electron is satisfied. Therefore, the resonances lines appear at,$$
site (7) \rightarrow B_0 + Am_l |m_l = -1 = B_0 - A\n$$
site (4) $\rightarrow B_0 + Am_l |m_l = 0 = B_0$ (same as site (5))
site (2) $\rightarrow B_0 + Am_l |m_l = 1 = B_0 + A$

In summary, the electron spin interacts with its local surroundings in multiple ways. The complete Hamiltonian is composed of the electron spin/field interaction $\mathcal{H}_{SB}$, the electron spin / orbital angular momentum interaction $\mathcal{H}_{SL}$, and the hyperfine interactions $\mathcal{H}_{SI}$,

$$\mathcal{H}_{total} = \mathcal{H}_{SB} + \mathcal{H}_{SL} + \mathcal{H}_{SI} = \mu_\beta B^T \cdot g \cdot S + \lambda \mathbf{L} \cdot \mathbf{S} + \sum_i \mathbf{S}^T \cdot \mathbf{A}_i \cdot \mathbf{I}_i \quad (2.7.19)$$

The ideas expressed in this chapter will be needed to understand EPR’s electrically detected equivalent, EDMR.
CHAPTER 3 – SPIN DEPENDENT TRANSPORT & ITS ELECTRICAL DETECTION

Electronic transport involves various spin dependent phenomena within micro- and nano-electronics. Unlike in EPR where a magnetization is detected from a sum of electron magnetic moments arising from defects with unpaired electrons in a material, spin dependent transport mechanisms are detected by measuring a spin dependent current, voltage, or resistance due to defects within fully processed devices. Some of the electronic devices in which these spin dependent phenomena can be detected are MOSFETs, BJTs, diodes, solar cells, photodiodes, capacitors, quantum dot structures, and single electron transistors. The more commonly known spin dependent transport mechanisms are recombination, tunneling, scattering, and hopping. This chapter is intended to further analyze some of these mechanisms.

3.1 – Spin Dependent Transport

All electrical processes within a micro- or nano-electronic device are spin dependent because they all involve electrons which are magnetic. Some of the more common spin dependent transport mechanism include spin dependent recombination (SDR), spin dependent tunneling (SDT), spin dependent hopping, and spin depending scattering. In this study, we will be concerned with the former two, recombination and tunneling.

Recombination occurs within a semiconductor when a conduction electron pairs with a valence hole. In this process, the electron and hole charge carriers annihilate each other thereby reducing the total current that passes through the device. This reduction in current is sometimes undesirable because it requires the device to operate at a higher power than would be required if recombination wasn’t present. Three of the more common forms of recombination include band-to-band (Auger) [53], radiative [53], and trap assisted (Shockley-Read-Hall (SRH)) [54] [55]. Auger recombination is a band-to-band process where the excess energy in the recombination process is given off to a third charge carrier. In the case of figure 3.1, the Auger process of recombination gives off energy to an electron which eventually relaxes back to the conduction band. This type of recombination limits the lifetime of charge carriers and effectively, the efficiency in solar cells [53]. Radiative recombination is a similar form of band-to-band recombination except that the energy given off in the recombination process involves the emission of a photon or heat with energy equal to that of the bandgap of the material. This process is typically encountered in direct bandgap semiconductor devices such as LEDs but can also be present in an indirect semiconductor device with the emission of a phonon for momentum conservation [53]. The third recombination process involves trap assistance via deep level defects and is present in both direct and indirect bandgap semiconductors. (Recombination can also be aided via dopants that reside in a small
bandgap material.) This form of recombination typically is referred to by the names of the people who first explained the process: Shockley, Read, and Hall [55] [54]. In this process, a conduction electron (or valence hole) is first captured by a deep level defect before it recombines with a valence hole (or conduction electron) as illustrated in figure 3.1. Note that this process must also conserve total momentum. This was a key idea in the Kaplan, Solomon, and Mott (KSM) model for SDR discussed in the upcoming section. However, the momentum that the KSM model took into account was not the conventional linear momentum of the charge carriers, but rather the angular momentum arising from the quantum mechanical spin of the charge carriers (electrons and holes) and defect (which contain unpaired electrons). Because the total momentum must be conserved, angular momentum due to spin must always be taken into account when analyzing any trap assisted process. Therefore, this form of recombination is aptly named spin dependent recombination. This idea of conservation of angular momentum also applied to other trap assisted transport processes such as spin dependent tunneling (SDT). From the background provided so far, it should be evident that magnetic resonance based experiments could be used to study the underlying defects (if spin dependent) responsible for limitations in device performance. The name given to such experiments that involving magnetic resonance to detect spin dependent transport phenomenon is electrically detected magnetic resonance (EDMR).

![Figure 3.1: Illustration of the three main types of recombination in a semiconductor.](image)

The first detection of EDMR via SDR and the first proposed model of SDR was provided by Lepine in 1972 [12]. His model suggested that the observed SDR effect scaled with applied magnetic field and indicated that the size of the effect could be as large as one part in one million change in current at X-band frequencies. This model fit his data nicely however; it failed to describe the significantly larger effect observed in other measurements [13] [56] [57]. A new model was proposed in 1978 by Kaplan, Solomon, and Mott (KSM) [13]. This model predicted a very large SDR effect and was based on the idea
of precursor electron pair formation prior to a recombination event. (The KSM model was proposed in order to describe resonant SDR; however, similar ideas can be directly applied resonant SDT.)

In the KSM model, recombination is assumed to follow the trap-assisted process illustrated in figure 3.1. In this two-step process, a conduction electron (or valence hole) could recombine with a valence hole (or conduction electron) by first getting trapped by a deep level defect which is paramagnetic. This deep level defect could involve a single or multiple dangling bonds in which at least one unpaired electron in resides. The key idea in the KSM model is that when the conduction electron (or valence hole) gets trapped by such a defect, its spin couples with the spin of the unpaired electron \( s \) of the defect, thereby forming an intermediate pair prior to recombination. The total “spin state” of the pair will determine whether or not recombination is possible. Because recombination involves the annihilation of an electron and hole (a process that involves zero change in angular momentum because the spin angular momentum of the electron and hole cancel), only pairs states that are characterized as having zero angular momentum will be involved in the recombination process. Therefore, spin angular momentum will be conserved. If the \( total \) angular momentum of the pair is not conserved, then the pair is able to dissociate; the electron releases back to the conduction band (or hole back to the valence band). The spins states that are formed include the triplet (total spin angular momentum \( s = 1 \)) and singlet (total spin angular momentum \( s = 0 \)) states. (These states will be discussed in more detail in the upcoming sections.) This process is illustrated in figure 3.2. With the application of electromagnetic radiation (that satisfies the magnetic resonance condition \( h\nu = g\mu_B B \)), either of the electrons in the spin pair are able to “flip”. This effectively allows the transitioning of triplet states into singlet states. As a result, the ratio of singlet to triplet states is modified, thereby altering the recombination current.

![Diagram](image)

**Figure 3.2:** Illustration of precursor spin pair formation prior to recombination. In this model for SDR, only singlet pairs that form will responsible for recombination.
This can be observed while monitoring the output current of a device when sweeping through resonance. Because the size of the SDR effect that we observe in our data is consistent with that predicted by the KSM model, a good deal of time will be spent in the next few chapters describing this model in detail. Note that in the figure 3.2, the conduction electron is depicted as getting captured in the vicinity of the defect’s unpaired electron, not in the same orbital in which the unpaired electron resides. If this were the case, only conduction electrons encountering the trap with different spin quantum numbers would get captured because of the Pauli Exclusion Principle (PEP). This is definitely not always the case. It is only required that the conduction electron gets captured in the vicinity of the unpaired electron such that it can couple with the defect electron. For example, the conduction electron may get trapped in an orbital from a neighboring atom. This not only allows for singlet pair formation, but also allows triplet pairs to form. The original KSM model paper begins with manipulating the quantum mechanical projection operators of the singlet and triplet state pairings. Before this model is reviewed any further, one must first understand the origin of these operators and the details of singlet and triplet states. Therefore, the quantum mechanical treatment of a pair of electron spins will first be reviewed.

(A note on notation before proceeding: In the upcoming sections, all bold face characters represent column vectors, \( \mathbf{i}, \mathbf{j}, \mathbf{k} \) are used to denote the x, y, z axes respectfully, and all other characters with a carrot top \( \hat{S} \) represent operators.)

### 3.2 –Pair of Electrons: Singlet and Triplets

The pairing of two electrons is the most basic application of group theory in quantum mechanics [42]. These groupings have been previously investigated by analyzing the two electrons that reside within a He atom [40]. Others have investigated the general addition of angular momentum of two arbitrary particles that have spin [42] [43] [58] [59] [60] [61]. Here, the analysis is slightly modified for the case of a paired conduction electron and defect electron spin as illustrated in figure 3.2. The treatment below is derived considering only pure basis states. A more precise treatment containing mixed states is derived in appendix B, which involves the use of the Pauli matrices.

Consider the case of an electron that gets trapped in the vicinity of a defect with an unpaired electron as illustrated in figure 3.2. If the electrons are in close enough proximity, they may couple (or become entangled) to form a pair state such that their angular momentum may add. Their angular momentum will add because spin is additive. Pairs of electrons whose spin adds to \( s = 1 \) will be designated a triplet state while pairs of electrons whose spin adds to \( s = 0 \) will be designated a singlet state. Let the spin of two electrons \( A \) and \( B \), be represented by, \( s_A \) and \( s_B \). The states of these individual electrons can be represented by,
Because the two electrons couple, their total state can be represented by a spin pair. Using the notation,

\[
|\psi\rangle = |\psi_A\psi_B\rangle = |s_A, m_{sA}; s_B, m_{sB}\rangle
\]

the pair states that can be formed might naively be written as,

\[
|\psi_1\rangle = |\frac{1}{2}, +\frac{1}{2}; +\frac{1}{2}, +\frac{1}{2}\rangle = |\uparrow\uparrow\rangle
\]

\[
|\psi_2\rangle = |\frac{1}{2}, -\frac{1}{2}; -\frac{1}{2}, -\frac{1}{2}\rangle = |\downarrow\downarrow\rangle
\]

\[
|\psi_3\rangle = |\frac{1}{2}, -\frac{1}{2}; +\frac{1}{2}, +\frac{1}{2}\rangle = |\downarrow\uparrow\rangle
\]

\[
|\psi_4\rangle = |\frac{1}{2}, +\frac{1}{2}; -\frac{1}{2}, -\frac{1}{2}\rangle = |\uparrow\downarrow\rangle
\]

Note that the first two pair states are acceptable because they are eigenstates of \(\hat{S}_z\) and \(\hat{S}_z\) and the electrons involved are indistinguishable, i.e.: they have a symmetrical pairing. This is a requirement for identical particles that are coupled in quantum mechanical treatment [40] [60]. For example, if two electrons are swapped in a given state, the resultant state should be identical to the original. This can only happen if the electrons are indistinguishable. Symmetrical pairings occur in the first two states, but not in the latter two. Because spin angular momentum is additive,

\[
s = s_A + s_B \tag{3.2.9}
\]

\[
m_s = m_{sA} + m_{sB} \tag{3.2.10}
\]

using the notation,

\[
total\ state = |total\ spin, total\ spin\ projection\rangle = |s, m_s\rangle
\]

The total state first two pair states can be represented by,
total state 1 = |1, +1⟩ = |↑↑⟩ \quad (3.2.11)
total state 2 = |1, −1⟩ = |↓↓⟩ \quad (3.2.12)

These are not the only valid states that an electron pair can possess however. It was already mentioned that they simply cannot be paired oppositely because the electrons would be distinguishable which is in violation of the accepted formulation of quantum mechanics. However, a pair of electrons can actually form states that are mixtures of oppositely paired spin states. Considering the rules of quantized angular momentum discussed so far, a fair guess for the total spin of the mixed state would be,

total state 3 = |1, 0⟩ \quad (3.2.13)

This is a valid guess because for any particle in general, if \( s = 1 \), then the values of \( m_s \) are in the range of \(-s \leq m_s \leq s\), where \( m_s \in \{-1,0,1\}\). In order to find the pair state mixture representation associated with this total state, the raising and lowering operators can be used. These operators essentially raise or lower the spin angular momentum to the next discretized quantum state. For this reason, these operators are also called ladder operators. The raising and lowering operators are defined respectively as [40] [44],

\[
\hat{S}_+ |s,m_s⟩ = \sqrt{(s(s + 1) − m_s(m_s + 1))} |s,m_s + 1⟩
\]

\[
\hat{S}_− |s,m_s⟩ = \sqrt{(s(s + 1) − m_s(m_s − 1))} |s,m_s − 1⟩
\]

where when applied to an individual spin,

\[
\hat{S}_+ |↑⟩ = 0 \quad (3.2.16)
\]

\[
\hat{S}_− |↓⟩ = 0 \quad (3.2.17)
\]

\[
\hat{S}_+ |↓⟩ = |↑⟩ \quad (3.2.18)
\]

\[
\hat{S}_− |↑⟩ = |↓⟩ \quad (3.2.19)
\]

Note that \( \hat{S}_+ |↑⟩ = 0 \) and \( \hat{S}_− |↓⟩ = 0 \) because the \( |↑⟩ \) state cannot be raised to a spin greater than \( \frac{1}{2} \) and \( |↓⟩ \) cannot be lowered to a spin less than \( −\frac{1}{2} \). Figure 3.3 illustrates the state transition diagram for these operators for a single electron.
One could also apply these operators to the pair of electron spins to obtain the next higher or lower quantum level of the spin pair. Applying the lowering operator to the \( S^- \) will yield the \( S^- |\uparrow\downarrow| = \frac{\hbar}{2} (|\uparrow\downarrow| + |\downarrow\uparrow|) \).

In order to find the pair state associated with this total state, one needs to apply the raising or lowering operators to either of the valid electron pair states using,

\[
\hat{S}^- = \hat{S}_{A^-} + \hat{S}_{B^-} \\
\hat{S}^+ = \hat{S}_{A^+} + \hat{S}_{B^+}
\]

Here, \( \hat{S}_{A^-} \) and \( \hat{S}_{A^+} \) only operate on electron A while \( \hat{S}_{B^-} \) and \( \hat{S}_{B^+} \) only operate on electron B. Applying (3.2.22) to pair state 1 results in pair state 3,

\[
\hat{S}^- |\uparrow\uparrow| = (\hat{S}_{A^-} + \hat{S}_{B^-}) |\uparrow\uparrow| = |\uparrow\uparrow| + |\downarrow\downarrow|
\]
Because the eigenvalues of the total states should equal the eigenvalues of the pair states, the expression obtained from lowering total state 1 can be set equal to expression obtained from lowering pair state 1,

\[ \hat{S}_- |1, +1\rangle = \hat{S}_- |\uparrow\uparrow\rangle \rightarrow \sqrt{2} |1, 0\rangle = (|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle) \] (3.2.24)

Rearranging yields the third valid basis state,

\[ |1, 0\rangle = \frac{1}{\sqrt{2}} (|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle) \] (3.2.25)

Note that the electrons are indistinguishable in this case and therefore represent the third symmetric state with \( s = 1 \). A fair guess for a fourth state would be when \( s = 0 \) This is another good guess considering that a pair of electrons within the same orbital have \( s = 0 \). In this case, the spin of the electrons always have different quantum numbers and therefore, their angular momentum will cancel which renders the pair without spin. If \( s = 0 \), then \( m_s = 0 \) according to \( -s \leq m_s \leq s \). Therefore, the fourth valid state can be represented by,

\[ total\ state\ 4 = |0, 0\rangle \] (3.2.26)

The pair state associate with this total state can be found by realizing that it must also be a mixture of oppositely oriented spins because \( m_s = 0 \). Therefore,

\[ pair\ state\ 4 = c_1 |\downarrow\uparrow\rangle + c_2 |\uparrow\downarrow\rangle \]

In order to find the coefficients \( c_1 \) and \( c_2 \), one must apply the constraints that are associated with orthonormal basis states; orthogonality and normality. It is evident that this state is orthogonal to \( |1, +1\rangle \) and \( |1, -1\rangle \); however, the orthogonality constrain needs to be applied to the mixed state of \( |1, 0\rangle \). Therefore, the orthogonality condition yields,

\[ \langle 1, 0 | 0, 0 \rangle = \frac{1}{\sqrt{2}} (\langle \downarrow\uparrow | + \langle \uparrow\downarrow |) \cdot (c_1 |\downarrow\uparrow\rangle + c_2 |\uparrow\downarrow\rangle) = \frac{c_1}{\sqrt{2}} \langle \downarrow\uparrow | \downarrow\uparrow \rangle + \frac{c_2}{\sqrt{2}} \langle \uparrow\downarrow | \uparrow\downarrow \rangle = 0 \Rightarrow c_1 = -c_2 \]

and the normality condition yields,
\( \langle 0,0,0,0 \rangle = (c_1 \langle \downarrow \uparrow | + c_2 \langle \uparrow \downarrow |) \cdot (c_1 \langle \downarrow \uparrow | + c_2 \langle \uparrow \downarrow |)) = c_1^2 \langle \downarrow \uparrow \downarrow \uparrow | + c_2^2 \langle \uparrow \downarrow \uparrow \downarrow | + 2c_1c_2 \langle \downarrow \uparrow \uparrow \downarrow | = 1 \Rightarrow c_1^2 + c_2^2 = 1 \)

Solving these two equations yields the valid pair state coefficients, \( c_1 = \pm 1/\sqrt{2} \) and \( c_2 = \mp 1/\sqrt{2} \). Therefore, the state associated with \( s = 0 \) can be represented as,

\[
\text{pair state } 4 = \frac{1}{\sqrt{2}} (|\uparrow\uparrow| - |\uparrow\downarrow|)
\]  

(3.2.27)

Note that when the electrons are swapped, the resultant state is the negative of the original state which is considered to be anti-symmetric. As noted from chapter 2, the total wave function of the electron must be anti-symmetric so that the Pauli exclusion principle is not violated. Restated here for convenience, the total wave function is the product of a spatial wave function and a spin wave function,

\[
|\psi_e\rangle = |\psi(x, y, z, t)\rangle \cdot |s, m_s\rangle
\]  

(3.2.28)

Therefore, if the spin wave function is symmetric, then the spatial wave function must be anti-symmetric and vice versa. As a result, the \( s = 0 \) state is the only state that has a spatially symmetric wave function.

In summary, the pairing of two electrons results in three symmetrical states with total spin \( s = 1 \) and one anti-symmetrical state with total spin \( s = 0 \). These states are commonly referred to as the triplet and singlet states, respectively. They all are eigenstates of \( \hat{S}^2 \) and \( \hat{S}_z \). Table 3.1 summarizes the properties of the four states just described and figure 3.4 illustrates the state transition diagram for the ladder operators applied to these states.

<table>
<thead>
<tr>
<th>Label</th>
<th>Total State</th>
<th>Pair State</th>
<th>Symmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_1 )</td>
<td>(</td>
<td>1, +1\rangle )</td>
<td>(</td>
</tr>
<tr>
<td>( T_{-1} )</td>
<td>(</td>
<td>1, -1\rangle )</td>
<td>(</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>(</td>
<td>0, 0\rangle )</td>
<td>( (</td>
</tr>
<tr>
<td>( S_0 )</td>
<td>(</td>
<td>0, 0\rangle )</td>
<td>( (</td>
</tr>
</tbody>
</table>

Table 3.1: Summary of total and pair states for two electrons.
Figure 3.4: Illustrating the state transition diagram for effect of raising and lowering operators on triplet and singlet states. The up arrows represent the paths of the raising operator and the down arrows represent the paths of the lowering operator.

In many cases, it is convenient to represent the total state in matrix form as the product of a coefficient matrix multiplied by the original set of basis states,

\[
\begin{pmatrix}
|1,1\rangle \\
|1,0\rangle \\
|0,0\rangle \\
|1,-1\rangle \\
\end{pmatrix} =
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1/\sqrt{2} & 0 & 0 \\
0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}
\begin{pmatrix}
|\uparrow\uparrow\rangle \\
|\uparrow\downarrow\rangle \\
|\downarrow\uparrow\rangle \\
|\downarrow\downarrow\rangle \\
\end{pmatrix}
\]

(3.2.29)

The coefficient matrix can easily be shown to be unitary. This is important in quantum mechanics because basis transformations such as this do not alter the length of the original basis vectors.

The goal now is construct the projection operators. An understanding of these operators is needed to understand the KSM model for SDR. Consider first the total spin angular momentum magnitude operator,

\[
\hat{S}^2 = (\hat{S}_A + \hat{S}_B)^2 = \hat{S}_A^2 + \hat{S}_B^2 + 2\hat{S}_A \cdot \hat{S}_B
\]

(3.2.30)

where \(\hat{S}_A\) only operates on electron A and \(\hat{S}_B\) only operates on electron B. Application of this operator on the pair of electron spins yields,
Therefore, the total spin angular momentum magnitude operator can be redefined as,

\[ S^2 = \frac{3}{2} + 2 \mathbf{S}_A \cdot \mathbf{S}_B \]  \tag{3.2.32}  

Because the eigenvalues of the total spin magnitude operator \( S^2 \) obey the same rules of discretized angular momentum as for an individual electron, then

\[ S^2 |s, m_s\rangle = \left( \frac{3}{2} + 2 \mathbf{S}_A \cdot \mathbf{S}_B \right) |s, m_s\rangle = s(s + 1) |s, m_s\rangle \]  \tag{3.2.33}  

Therefore,

\[ \mathbf{S}_A \cdot \mathbf{S}_B |s, m_s\rangle = \frac{s(s + 1) - 3/2}{2} |s, m_s\rangle \]  \tag{3.2.34}  

One can use this result to find the projection operators of the singlet and triplet states. The projection operators essentially provide a means to find the probability of either state using the individual spin vectors \( \mathbf{S}_A \) and \( \mathbf{S}_B \). They are called projection operators because they both project the angular momentum vectors \( \mathbf{S}_A \) and \( \mathbf{S}_B \) onto a probability space \( \in \{0,1\} \). From equation (3.2.34),

\[ \mathbf{S}_A \cdot \mathbf{S}_B |s, m_s\rangle |_{s=0} = -\frac{3}{4} \quad \text{(singlet)} \]  \tag{3.2.35}  

\[ \mathbf{S}_A \cdot \mathbf{S}_B |s, m_s\rangle |_{s=1} = +\frac{1}{4} \quad \text{(triplet)} \]  \tag{3.2.36}  

It is desired that if \( \mathbf{S}_A \cdot \mathbf{S}_B = -3/4 \), the singlet projection operator \( \hat{P}_s |s, m_s\rangle \) map that value to 1 indicating the presence of a pure singlet state. It is also desired that if \( \mathbf{S}_A \cdot \mathbf{S}_B = 1/4 \), the singlet projection operator \( \hat{P}_s |s, m_s\rangle \) map that value to 0 indicating that a pure triplet is present. This operator is easily constructed by,
\[
\hat{P}_s = \frac{1}{4} - \mathbf{S}_A \cdot \mathbf{S}_B
\]  (3.2.37)

Similarly, it is desired that if \( \mathbf{S}_A \cdot \mathbf{S}_B = 1/4 \), the triplet projection operator \( \hat{P}_t |s, m_s\rangle \) map that value to 1 indicating the presence of a pure triplet state. It is also desired that if \( \mathbf{S}_A \cdot \mathbf{S}_B = -3/4 \), the triplet projection operator \( \hat{P}_t |s, m_s\rangle \) map that value to 0 indicating the presence of a pure singlet state. This operator is also easily constructed,

\[
\hat{P}_t = \frac{3}{4} + \mathbf{S}_A \cdot \mathbf{S}_B
\]  (3.2.38)

The projection operators provide a means to evaluate the probability of a pure singlet and pure triplet state given the individual spin vectors of \( \mathbf{S}_A \) and \( \mathbf{S}_B \). Rewriting the original basis states in terms of the singlet triplet basis allows one to better understand the probabilities produced by the projection operators.

\[
|\uparrow\uparrow\rangle = |1, +1\rangle
\]  (3.2.39)
\[
|\downarrow\downarrow\rangle = |1, -1\rangle
\]  (3.2.40)
\[
|\uparrow\downarrow\rangle = (|1,0\rangle + |0,0\rangle)/\sqrt{2}
\]  (3.2.41)
\[
|\downarrow\uparrow\rangle = (|1,0\rangle - |0,0\rangle)/\sqrt{2}
\]  (3.2.42)

For example, if two electrons have the same spin state, \( |\uparrow\uparrow\rangle \) or \( |\downarrow\downarrow\rangle \), then 100% of the time, a pure triplet will be formed, i.e. \( T_{+1} \) or \( T_{-1} \). If the spins are oppositely paired, \( |\uparrow\downarrow\rangle \) or \( |\downarrow\uparrow\rangle \), then there will be a 50% chance that the pair is either a pure triplet or a pure singlet, i.e. \( T_0 \) or \( S_0 \). These probabilities are illustrated in table 3.2.

<table>
<thead>
<tr>
<th>electron A</th>
<th>electron B</th>
<th>( \hat{P}_t )</th>
<th>( \hat{P}_s )</th>
<th>Singlet/Triplet State</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑</td>
<td>↑</td>
<td>1</td>
<td>0</td>
<td>( T_{+1} )</td>
</tr>
<tr>
<td>↓</td>
<td>↓</td>
<td>1</td>
<td>0</td>
<td>( T_{-1} )</td>
</tr>
<tr>
<td>↓</td>
<td>↑</td>
<td>1/2</td>
<td>1/2</td>
<td>( T_0 ) or ( S_0 )</td>
</tr>
<tr>
<td>↑</td>
<td>↓</td>
<td>1/2</td>
<td>1/2</td>
<td>( T_0 ) or ( S_0 )</td>
</tr>
</tbody>
</table>

Table 3.2: Probability of pure singlet and pure triplet states given the individual electron spin states.

In reality, the electrons are more accurately represented as a linear combination of up and down states. As a result, all orientations of the spin vectors \( \mathbf{S}_A \) and \( \mathbf{S}_B \) will be possible. This indicates that not only pure singlet and pure triplets will be formed, but rather a mixture of them. This is one of the assumptions utilized in the KSM model for SDR.
### 3.3 – Kaplan, Solomon, and Mott Model

The Kaplan, Solomon, and Mott (KSM) model for spin dependent recombination (SDR) was developed in order to help explain the large change in recombination observed in a device when exposed to a resonant field which could not be explained from earlier polarization models [12]. As mentioned previously, the innovative concept in the KSM model is the formation of spin pair prior to a trap assisted recombination event [13] [14] [15]. As illustrated in figure 3.2, the spin pair consists of either a conduction electron or valence hole in the presence of an unpaired electron(s) within a deep level defect. Because the process of recombination involves zero spin (cancellation of angular momentum of electron and hole), only spin pairs that form singlet states with \( s = 0 \) are able to recombine. Triplet pairs are able to dissociate which allows the electron to return back to the conduction band (or the holes back to the valence band) leaving the defect as it was prior to the pairing so that the process can be repeated. This model neglects the momentum involved with the electron’s spin-orbit interaction discussed in chapter 2. If large atoms are involved, conservation of total momentum \( J = S + L \) will depend on the intrinsic angular momentum due to spin \( S \) as well as the momentum arising from the spin-orbit interaction \( L \). Therefore, the possibility that triplet pairs are involved in recombination may be possible [14]. This is not considered in this study because we utilize devices made of materials whose atoms are too small for spin-orbit coupling to play a role in recombination.

In this analysis, recombination will first be analyzed in the absence of a resonant field, ie: the resonance condition of the electron’s is not satisfied. As was derived in the previous section, the probability of a singlet pair formation may be explained from the singlet projection operator,

\[
P_s = \frac{1}{4} - \vec{S}_A \cdot \vec{S}_B
\]  

(3.3.1)

If recombination is only allowed from singlet pair formation, then the singlet projection operator can be used to define the probability of a recombination event. However, this operator describes the probability of a pure singlet pairing. In reality, the electrons are more accurately represented as a mixture of up and down states. As a result, not only pure singlet and pure triplet states will be formed, but rather a mixture of them. Therefore, one must consider all orientations of the spin vectors \( S_A \) and \( S_B \). To simplify the analysis, the quantum average is utilized and is found by,

\[
\langle P_s \rangle = \frac{1}{4} - \langle \vec{S}_A \cdot \vec{S}_B \rangle
\]  

(3.3.2)
Consider the quantum averages of the electron spin vectors illustrated in figure 3.5. Note that the quantum averages of the $\hat{S}_x$ and $\hat{S}_y$ components for both the up and down spin states are zero and the quantum average of the $\hat{S}_z$ component is simply $m_s$. Because of this, the average magnitude of the spin angular momentum is equal to $|m_s|$. For example,

$$\langle \hat{S}_x \rangle = \langle \hat{S}_y \rangle = 0$$  \hspace{1cm} (3.3.3)  
$$\langle \hat{S}_z \rangle = m_s$$  \hspace{1cm} (3.3.4)  
$$\langle |\hat{S}| \rangle = |m_s|$$  \hspace{1cm} (3.3.5)

The angle $\theta$ between the two spin vectors can be defined by,

$$\hat{S}_A \cdot \hat{S}_B = |S_A||S_B| \cos(\theta) = \frac{3}{4} \cos(\theta)$$  \hspace{1cm} (3.3.6)

and the angle between the quantum average of the spin vectors can be defined by,

$$\langle \hat{S}_A \cdot \hat{S}_B \rangle = \langle |S_A||S_B| \rangle \cos(\theta) = \frac{1}{4} \cos(\theta)$$  \hspace{1cm} (3.3.7)

Figure 3.5: Comparison of electron up and down spin vectors and their vector quantum averages. So, equation (3.3.2) can be rewritten as,
This relation provides a means to find the probability of a singlet pairing for a continuous range of spin orientations for both electrons. Note that the highest probability (50%) of a singlet pairing is when the $\theta = 180^\circ$ (opposite direction). This agrees with the conventional singlet projection operator that provided the probability for the pure singlet state. As the spins become more aligned (i.e. $\theta = 0^\circ$), the probability of a singlet pairing reduces to 0% and the probability of a triplet pairing increases to 100%. This also agrees with the conventional projection operators for pure singlet and triplet states. The distribution for these states as a function of the angle $\theta$ is illustrated in figure 3.6. Using the probability for a singlet pairing, the recombination rate can easily be calculated by,

$$W_R(\theta) = W_S\langle \hat{p}_S \rangle = W_S \left( \frac{1 - \cos(\theta)}{4} \right)$$

(3.3.9)

where $W_S$ represents the recombination rate for a pair in a singlet configuration. As a result, this expression represents the recombination probability per unit time as a function of angle between the electrons spins in the pair. In steady state, the pair creation rate $c(\theta)$ is equal to the pair annihilation rate. In order for a pair to be annihilated, it may either dissociate or recombine. Therefore,

Figure 3.6: Probability of singlet and triplet pairings as a function of the angle between the spin vectors. Note that pure singlets are formed 50% of the time when $\theta = 180^\circ$ and pure triplets are formed 100% of the time when $\theta = 0^\circ$. 
\[ c(\theta) = N(\theta)[W_D + W_R(\theta)] \]  

(3.3.10)

where \( W_D \) is the disassociation probability of the pair (pairs that don’t recombine), \( W_R(\theta) \) is the recombination probability of the pair, and \( N(\theta) \) is the steady state density of existing pairs. \( W_D \) and \( W_R(\theta) \) can be therefore thought of as weighting functions. Their product with \( N(\theta) \) describes the rate at which these respective processes occur.

The next step is to find the pair creation rate \( c(\theta) \). First, remember that it was originally assumed that the electron spin vectors in the pair could be orientated in any direction. In other words, the spins can be thought of as being isotropic. Therefore, one must consider all pairings in a 3-dimensional space. However, equation (3.3.8) implies that all that is needed to determine the probability of the pair state is the relative angle between the spin vectors, not the absolute polar angles \( \theta, \phi \) between the spins. However, one must first find the creation rate as a function of 3-dimensional space using polar coordinates \( \theta, \phi \). Then, this function can be mapped to a 2-dimensional plane so that \( \theta \) can be used as the sole parameter that defines the probability of the pair state. Because it is only the relative angle between the two spin vectors that matters, the coordinate axis can be fixed to one of the electron spin vectors. As illustrated in figure 3.7, fixing the electron spin \( \mathbf{S}_A \) to the coordinate axis allows one to evaluate the orientation of a single spin vector \( \mathbf{S}_B \). This spin vector is able to point in any direction radially originating from the center of the sphere. As a result, the created electron spin pairs will have an isotropic distribution and will be constant as a function of the solid angle \( \Omega \) in units of steradians. That is,

\[ c(\Omega) = A \]  

(3.3.11)

(Steradian is the unit of the solid angle which describes a 2-dimensional angular span in a 3-dimensional space similar to the way a radian describes a 1-dimensional angle in a 2-dimensional plane. Because the surface area of a sphere is \( 4\pi r^2 \), the definition implies that a sphere measures \( 4\pi \) steradians. Therefore, the maximum solid angle that can be subtended at any point is \( 4\pi \) steradians.) The constant \( A \) can be solved for by normalization. Therefore, integrating over the solid angle in the range 0 to \( 4\pi \),

\[ \int_0^{4\pi} A d\Omega = 1 \rightarrow A = \frac{1}{4\pi} = c(\Omega) \]  

(3.3.12)

Using the steradian to polar coordinate transformations,

\[ d\Omega \rightarrow \sin(\theta) d\theta d\phi \]  

(3.3.13)
\[ c(\Omega) \rightarrow c(\theta, \phi) \quad (3.3.14) \]

(3.3.11) can be rewritten as,

\[ c(\Omega) d\Omega = \frac{1}{4\pi} \sin(\theta) d\theta d\phi = c(\theta, \phi) d\theta d\phi \quad (3.3.15) \]

where,

\[ c(\theta, \phi) = \frac{1}{4} \sin(\theta) \quad (3.3.16) \]

Notice that the pair creation rate in 3-dimensional space is independent of the polar angle \( \phi \). However, one must account for all possible angles of \( \phi \) by integrating over this variable on the range of \( 0 \leq \phi \leq 2\pi \). This will effectively map the 3-dimensional space onto a 2-dimensional space as illustrated in figure 3.7. By doing this, \( c \) becomes solely dependent on \( \theta \),

\[ c(\theta) = \int_{0}^{2\pi} c(\theta, \phi) d\phi = \frac{1}{4\pi} \sin(\theta) \int_{0}^{2\pi} d\phi = \frac{1}{2} \sin(\theta) \quad (3.3.17) \]

Figure 3.7: (left) By fixing the electron spin \( S_A \), one can envision the other electron spin \( S_B \) pointing in an isotropic distribution of angles with respect to that fixed axis. Note that of all the orientations that \( S_B \) can be possess, it will have the least probability of aligning exactly with \( S_A \) but the highest probability of aligning perpendicular to it along the disk illustrated around the sphere. Because it is only the relative angle \( \theta \) between the pairs that matters, and not the absolute angles \( \theta, \phi \), one can map the (middle) 3-dimentional coordinate system to a (right) 2-dimensional coordinate system which accounts for all orientations of pairings and only requires the use of the angle \( \theta \).
Including the total creation rate constant $C$, equation (3.3.17) can be rewritten as,

$$c(\theta) = \frac{C}{2} \sin(\theta)$$  

(3.3.18)

This function describes the creation rate of pairs as a function of angle between the pairs. The $\sin(\theta)$ dependence simply implies that pairs will be more likely to be created in orthogonal pairs than parallel pairs. (This logic is very similar to that used to derive the EPR powder pattern spectrum of a defect with axial symmetry.) Plugging in equations (3.3.9) and (3.3.18) back into equation (3.3.10), yields the steady state density of pairs as a function of angle $\theta$,

$$N(\theta) = \frac{C}{2} \sin(\theta) \left[ W_D + W_S \frac{1 - \cos(\theta)}{4} \right]$$  

(3.3.19)

The recombination rate can now be calculated. It is simply the product of the steady state density of pairs present at a given time $N(\theta)$, multiplied by the recombination probability rate per unit time $W_R(\theta)$, integrated over the range of possible angles between the spins, $0 \leq \theta \leq \pi$,

$$R = \int N(\theta) W_R(\theta) = \int_{0}^{\pi} \frac{2C \sin(\theta) W_S (1 - \cos(\theta))}{4W_D + W_S (1 - \cos(\theta))} \, d\theta$$  

(3.3.20)

This integral can be solved by performing two substitution of variables. This results in

$$R = C \left\{ 1 - 2\lambda \ln \left( \frac{2\lambda + 1}{2\lambda} \right) \right\}$$  

(3.3.21)

where $\lambda = W_D / W_S$. This relation describes the recombination rate in the absence of a resonant field. The spin pairs follow a correlated distribution that does not change as a function of applied static magnetic field. In other words, the charge carriers will maintain their relative spin orientation even after being captured. However, when the electrons are in resonance, the charge carriers will not maintain their relative spin orientation because they are able to flip and change state. This will modify the recombination rate denoted by equation (3.3.21).

Now, consider the case in which either of the electrons in the spin pair experiences the magnetic resonant condition. As explained earlier, the electrons are able to flip spin states when the resonance
condition is satisfied. As a result, triplet pairs are able to transition into singlet pairs and singlet pairs into triplet pairs. Because the lifetime of a singlet pairing is short, recombination current in a device will increase. When the microwave power is increased such that it saturates the spin system (which destroys the magnetization, a process undesired in conventional EPR), the spin pair orientation becomes completely randomized because of the continuous flipping of electron spins. (Therefore, it is desired to have the largest possible radiation power applied to a device during an EDMR experiment.) As a result, the recombination probability density function becomes the average of equation (3.3.9),

\[ W_R(\theta) = \langle W_R(\theta) \rangle = W_S \frac{1 - \langle \cos(\theta) \rangle}{4} = \frac{W_S}{4} \quad (3.3.22) \]

Notice the recombination probability per unit time in the resonant case is uniform and independent of the angle between the two spins. As a result, the spin pair will not maintain its relative distribution so the correlation is destroyed. It can still be assumed that the created pairs are formed from an isotropic distribution of electron spins. Using equation (3.3.22), the steady state density of pairs becomes,

\[ N(\theta) = \frac{c(\theta)}{[W_D + W_R(\theta)]} = \frac{C}{2} \frac{\sin(\theta) d\theta}{W_D + \frac{W_S}{4}} = \frac{2C \sin(\theta) d\theta}{4W_D + W_S} \quad (3.3.23) \]

The recombination rate when the electron spins are fully saturated from a resonant field becomes,

\[ R_{SAT} = \int_0^\pi N(\theta) W_R(\theta) d\theta = \int_0^\pi \frac{2C \sin(\theta) W_S}{4W_D + W_S} d\theta = \frac{C}{4\lambda + 1} \quad (3.3.24) \]

The change in recombination in the resonant and non-resonant cases is simply,

\[ \frac{\Delta R}{R} = \frac{R_{SAT} - R}{R} = \frac{C}{4\lambda + 1} - \frac{C \left\{ 1 - 2\lambda \ln \left( \frac{2\lambda + 1}{2\lambda} \right) \right\}}{C \left\{ 1 - 2\lambda \ln \left( \frac{2\lambda + 1}{2\lambda} \right) \right\}}, \quad (3.3.25) \]

where \( \lambda = W_D / W_S \). The \( \Delta R/R \) is maximized when \( \lambda = 0.3 \). Even though this quasi-classical model was proposed in 1978, it provides, to this day, the most widely accepted understanding of spin dependent recombination that is capable of predicting the large effect observed in many SDR measurements. This model is completely independent of applied magnetic field and only predicts a change in recombination.
with the application of EM radiation. However, more recent studies that utilize a purely quantum mechanical treatment of a pair of spins have shown that this is not the case. These studies have treated this issue by studying the dynamics of the spin pair density matrix using the Liouville equation [14] [16] [17]. This method was first utilized by Haberkorn and Dietz in 1980 to investigate SDR [16]. Following studies were performed by Barabanov and Tretiak in 1996 [17] and then in 2003 by Boehme and Lips [14]. The details of these studies will not be discussed here; however, a brief introduction to the density matrix and the Liouville equation is provided in appendix D. However, it should be noted that in the initial study of Haberkorn and Dietz, it was shown that the mere presence of an external magnetic field influences the singlet recombination probability $W_R$ of the spin pair. More specifically, they were able to show that,

$$W_R = \frac{W_S}{4(W_S + W_D)} \quad \text{for} \quad B_0 = \infty$$

$$W_R = \frac{W_S}{2(W_S + 2W_D)} \quad \text{for} \quad B_0 = 0$$

where, once again, $W_S$ is the recombination rate of a pair in a pure singlet state and $W_D$ is dissociation rate of the spin pair. These equations suggest that the SDR phenomenon in not completely field independent. The equations also imply that SDR phenomenon can be studied with the sole application of an external magnetic field. Though these relations may provide some insight about the change in current within a device as a function of magnetic field (for example magnetoresistance), it should still be unclear as to the underlying physical mechanisms responsible for recombination. This requires further investigation of spin dependent transport at extremely low- and zero- magnetic fields.

### 3.4 – Zero-Field Spin Dependent Transport

It should be evident at this point that even though conventional EPR and EDMR both require the ability to flip electron spins, the detection, sensitivity, and analysis of each technique are completely different. What should not be so obvious at this point is why spin dependent transport can be detected at zero magnetic field and in the absence of electromagnetic radiation in fully processed devices. This section describes this phenomenon.

Low magnetic field electronic transport has been a hot topic in recent years. Many have reported on a low-field magnetoresistance phenomenon, especially in organic semiconductors [26] [27] [28] [29] [30]. This phenomenon has typically been associated with the changing of singlet and triplet ratios as a function of magnetic field. One particular study suggested a model to describe this phenomenon which is very similar to the KSM model described in the previous section [28]. This model and other proposed
mechanisms to describe the zero-field phenomenon are plausible but only speculative. Also, zero- and low-field spin dependent transport has been observed in double quantum dots structures for quantum computing applications [20] [19] [22]. In these studies, this low magnetic field phenomenon is attributed to singlet/triplet mixing and nuclear spin dynamics [24] [25].

We have also observed magnetic field induced current change phenomenon in our low field measurements in the absence of electromagnetic radiation. Until our recent study, it was unclear which spin dependent mechanism was responsible for this phenomenon. In our latest work, we were able to directly demonstrate that this low field effect is associated with spin dependent recombination (SDR) in 4H SiC BJTs [18]. Our more recent work demonstrates that zero field detection of SDR can be performed in MOSFETs and zero field detection of spin dependent tunneling (SDT) can be performed in capacitor structures. Also, unlike any of the previous studies at low field, we can to clearly resolve the electron-nuclear hyperfine interactions using magnetic field modulation because of the crystalline devices that we use. To the best of our knowledge, no one has previously measured electron-nuclear hyperfine interactions through this zero-/low-field process. The basic underlying mechanism of the zero-field phenomenon will now be briefly discussed.

From chapter 2, it was determined that for conventional EPR, the unpaired electrons involved experience interactions from the externally applied magnetic field, the electron’s interaction with its motion around the nucleus, and the electron’s interaction with nearby magnetic nuclei. As noted before, the Hamiltonian can be described as,

\[ \mathcal{H}_{total} = \mathcal{H}_{SB} + \mathcal{H}_{SL} + \mathcal{H}_{SL} = \mu_B \mathbf{B} \cdot \mathbf{g} \cdot \mathbf{S} + \lambda \hat{L} \cdot \mathbf{S} + \sum_i \mathbf{S}_i^T \cdot \mathbf{A}_i \cdot \hat{I}_i \] (3.4.1)

Because the material systems (lattice and dopant atoms) utilized in this study involve small atoms (Si, C, N, Al, B, P, O), the electron spin orbit interaction can be neglected. Also, because two (at least) electron spins are involved in most spin dependent transport processes (such as the electron spin pairing mechanism proposed in the KSM model), one must account for each of their interactions with the field and nearby nuclei. Therefore, the total Hamiltonian for a two electron system becomes,

\[ \mathcal{H}_{total} = \mu_B \mathbf{B} \cdot \mathbf{g} \cdot \mathbf{S}_A + \mu_B \mathbf{B} \cdot \mathbf{g} \cdot \mathbf{S}_B + \sum_i \mathbf{S}_A^T \cdot \mathbf{A}_i \cdot \hat{I}_i + \sum_i \mathbf{S}_B^T \cdot \mathbf{A}_i \cdot \hat{I}_i \] (3.4.2)

where \( \mathbf{S}_A \) is the spin vector of electron A and \( \mathbf{S}_B \) is the spin vector of electron B. For ease of analysis, assume for the moment that there is no anisotropy in \( \mathbf{g} \), there are no nearby magnetic nuclei, and that the
electrons \( A \) and \( B \) have the same resonance condition (ie: same \( g \) value). When placed in a magnetic field \( \mathbf{B} = B_z \hat{\mathbf{k}} = B_0 \), the Hamiltonian can be redefined by,

\[
\mathcal{H} = g \mu_\beta (\mathbf{S}_A + \mathbf{S}_B) \cdot \mathbf{B} = g \mu_\beta \vec{S} \cdot \mathbf{B}
\]  

(3.4.3)

where \( \vec{S} \) is the total spin operator. Application of this Hamiltonian on the spin wave function \( |s, m_s\rangle \) yields,

\[
\mathcal{H}|s, m_s\rangle = g \mu_\beta (\hat{\mathbf{S}}_x + \hat{\mathbf{S}}_y + \hat{\mathbf{S}}_z) \cdot (\hat{B}_x + \hat{B}_y + \hat{B}_z)|s, m_s\rangle
\]

\[
= g \mu_\beta \hat{S}_z \hat{B}_z |s, m_s\rangle
\]

\[
= m_s g \mu_\beta B_0 |s, m_s\rangle,
\]  

(3.4.4)

where \( m_s = m_{s_A} + m_{s_B} \) and the triplet and singlet energy states are,

\[
\mathcal{H}|0,0\rangle = 0
\]
\[
\mathcal{H}|1,+1\rangle = g \mu_\beta B_0 |1,+1\rangle
\]
\[
\mathcal{H}|1,0\rangle = 0
\]
\[
\mathcal{H}|1,-1\rangle = -g \mu_\beta B_0 |1,-1\rangle
\]

Note that at zero magnetic field, the singlet and triplet states are all degenerate with zero energy. As a result, the triplet states are able to “mix” with the singlet states. Singlet triplet mixing allows the triplet states to transition into singlet states without application of EM energy because at zero magnetic field, they are energetically accessible. (Dynamic nuclear polarization may play a role in this mixing of the singlet and triplet states as suggested by other studies involving double quantum dots [24] [25].) Because the lifetime of a singlet pairing is much shorter than that of the triplet pairing, more triplet to singlet transitions will be possible and lead to an increase the recombination [13]. The process of triplet states non-radiatively transitioning into a singlet states (or vice versa), is commonly referred to as intersystem crossing (ISC). This is illustrated in figure 3.8.
Figure 3.8: Singlet triplet mixing allows the triplet states to transition into singlet states because all of the basis states are degenerate at zero magnetic field. As a result, recombination will be enhanced at zero-field where the mixing occurs.

Therefore, if the only spin interaction is due to the external magnetic field, then an enhancement of recombination would be expected at zero Gauss. We consistently observe this when performing low-field measurements in our lab. However, sometimes we observe a split peak centered precisely about zero magnetic field. This suggests the involvement of electron spin-spin interactions.

As was discussed in the previous section, the most accepted model of SDR is based upon an intermediate electron pair formation (e.g. conduction electron, or valance hole, getting captured by a deep level defect with unpaired electron(s)) prior to a recombination event. Because two spins are involved in this capture process, the other two mechanisms that need to be taken into account when considering a general model for spin dependent transport are the dipolar and exchange interactions between two electrons in the spin pair [43] [52] [46] [47]. The most general form of the Hamiltonian for the exchange interaction (also sometimes referred to the orbital overlap interaction) is,

$$\mathcal{H}_j = - \sum_{ij} J_{ij} \tilde{S}_A \cdot \tilde{S}_B = - \frac{1}{2} \left[ \tilde{S}_A^T \cdot J \cdot \tilde{S}_B + \tilde{S}_B^T \cdot J \cdot \tilde{S}_A \right]$$  \hspace{1cm} \text{(3.4.5)}$$

where \( J \) describes the exchange interaction between the two electron spins \( \tilde{S}_A \) and \( \tilde{S}_B \) and has the form,
\[ J = \begin{bmatrix} J_{xx} & J_{yx} & J_{zx} \\ J_{xy} & J_{yy} & J_{zy} \\ J_{xz} & J_{yz} & J_{zz} \end{bmatrix} \]  

(3.4.6)

For simplicity, we will assume an isotropic interaction which results in the Hamiltonian to take on the form,

\[ \mathcal{H}_J = -J_o \mathbf{\hat{S}}_A^T \cdot \mathbf{\hat{S}}_B \]  

(3.4.7)

which is sometimes referred to as the Heisenberg-Dirac exchange Hamiltonian. In general, the magnitude of \( J_o \) is dependent upon the Coulombic interactions between two electrons, electron spacing, and wave function overlap. As illustrated by equation (3.4.7), the exchange interaction is dependent on the spin angular momentum of each electron. Therefore, the exchange interaction is a purely quantum mechanical phenomenon that arises because two electrons within the vicinity of each other both have intrinsic angular momentum due to spin. Even though it takes into account the Coulombic interaction between the two electron spins, it does however influence the magnetic interaction of the two electrons in the spin pair. For example, it causes an energy splitting between the triplet and singlet states. It was shown earlier in section 3.2 that,

\[ \mathbf{\hat{S}}_A \cdot \mathbf{\hat{S}}_B = -\frac{3}{4} \quad \text{for a singlet, } s = 0 \]

\[ \mathbf{\hat{S}}_A \cdot \mathbf{\hat{S}}_B = +\frac{1}{4} \quad \text{for a triplet, } s = 1 \]

Using the notation for a pair of electrons,

\[ |\text{total spin, total spin projection}\rangle = |s, m_s\rangle \]

the eigenvalues (or energies) for the exchange Hamiltonian operating on the spin pair eigenstates, \( |s, m_s\rangle \) yield,

\[ \mathcal{H}_J |0,0\rangle = J_o \frac{3}{4} |0,0\rangle \quad \text{(singlet state)} \]  

(3.4.8)

\[ \mathcal{H}_J |1, m_s\rangle = -J_o \frac{1}{4} |1, m_s\rangle \quad \text{(triplet states)} \]  

(3.4.9)
Note that in the absence of a magnetic field, the triplet states are degenerate and are separated in energy from the singlet state by a factor of $J_0$. When placed in a magnetic field, the Hamiltonian of the singlet and triplet pairs are described by,

$$\mathcal{H} = g\mu_\beta \vec{S} \cdot \vec{B} - J_0 \vec{S}_A \cdot \vec{S}_B$$  \hspace{1cm} (3.4.10)

Application of this Hamiltonian on the spin wave function illustrates that the triplet and singlet energy states are separated by the exchange parameter $J_0$, 

$$\mathcal{H}|s,m_s\rangle = \left( m_s g\mu_\beta B_0 - J_0 \left( s - \frac{3}{4} \right) \right) |s,m_s\rangle$$ \hspace{1cm} (3.4.11)

$$\mathcal{H}|0,0\rangle = J_0 \frac{3}{4} |0,0\rangle$$

$$\mathcal{H}|1,+1\rangle = \left( g\mu_\beta B_0 - J_0 \frac{1}{4} \right) |1,+1\rangle$$

$$\mathcal{H}|1,0\rangle = -J_0 \frac{1}{4} |1,0\rangle$$

$$\mathcal{H}|1,-1\rangle = \left( -g\mu_\beta B_0 - J_0 \frac{1}{4} \right) |1,-1\rangle$$

Note that $J_0$ can be positive or negative which indicates that the singlet can lie above or below the triplet states in energy. $J_0 > 0$ typically refers to ferromagnetic coupling and $J_0 < 0$ refers to antiferromagnetic coupling. Figure 3.8 illustrates the energies of the singlet and triplet states as a function of magnetic field for $J_0 < 0$. In this case, the singlet triplet mixing will not occur at precisely zero magnetic field, but rather at a very low positive field and a very low negative field. As a result, a split peak will be observed about zero magnetic field. Note that as the exchange interaction constant $J_0$ is increased, a larger field is required for the mixing of triplet and singlet states. As a result, if $J_0$ is small, a single recombination peak will be observed at zero magnetic field and if $J_0$ is large, a double peak will be observed about zero magnetic field. This is illustrated in figure 3.9.

The other important interaction that needs to be considered is the magnetic interaction between the two electrons. This interaction is very similar to the electron nuclear coupling analysis, with the exception that in this case, the electron magnetically couples with its paired electron. This interaction is commonly referred to as the spin-spin or dipolar interaction. Its Hamiltonian is represented by,
Singlet triplet mixing allows the triplet states to transition into singlet states because at certain magnetic fields, they are energetically accessible. As a result, recombination will be enhanced at these tiny fields (very close to 0G) at which the mixing occurs. Note that if the exchange interaction is present in a measurement, one should expect to observe a split peak about zero magnetic field.

\[
\mathcal{H}_D = \frac{\mu_0}{4\pi} g^2 \mu_B^2 \frac{\hat{S}_A \cdot \hat{S}_B}{r^3} - 3 \frac{(\hat{S}_A \cdot \mathbf{r})(\hat{S}_B \cdot \mathbf{r})}{r^5}
\]

\[
= \frac{\mu_0}{4\pi} g^2 \mu_B^2 \left[ \frac{\hat{S}_A^T \cdot \hat{S}_B}{r^3} - 3 \frac{\hat{S}_A^T \cdot (\mathbf{r} \cdot \mathbf{r}^T) \cdot \hat{S}_B}{r^5} \right]
\]

\[
= \hat{S}_A^T \cdot \left( \frac{\mu_0}{4\pi} g^2 \mu_B^2 \left[ \frac{\mathbf{I}}{r^3} - 3 \frac{(\mathbf{r} \cdot \mathbf{r}^T)}{r^5} \right] \right) \cdot \hat{S}_B
\]

\[
= \hat{S}_A^T \cdot \mathbf{D} \cdot \hat{S}_B
\] (3.4.12)

where,

\[
\mathbf{D} = \begin{bmatrix}
D_{xx} & D_{yx} & D_{zx} \\
D_{xy} & D_{yy} & D_{zy} \\
D_{xz} & D_{yz} & D_{zz}
\end{bmatrix}
\] (3.4.13)

\[
D_{ij} = \frac{\mu_0}{4\pi} g^2 \mu_B^2 \left[ \delta_{ij} \frac{\mathbf{I}}{r^3} - 3 \frac{ij}{r^5} \right]
\] (3.4.14)
Here, $i,j = x,y,z$, the brackets indicates the average position of the electron over the spatial coordinates
of the wave function, and $\delta_{ij}$ is the delta function which is defined by

$$
\delta_{ij} = \begin{cases} 
1, & \text{for } i = j \\
0, & \text{for } i \neq j 
\end{cases}
$$

The matrix $D$ is composed of the dipolar coupling parameters (also referred to as the zero-field splitting
parameters or fine structure tensor). If the principle axes are known, then the dipolar coupling matrix can
be written as a diagonal matrix,

$$
D = \begin{bmatrix}
D_{xx} & 0 & 0 \\
0 & D_{yy} & 0 \\
0 & 0 & D_{zz}
\end{bmatrix}
$$

(3.4.15)

(Once again, the principal axes are the ones that lie along the axes of molecular symmetry within the
material system under investigation.) It should be pointed out that the components $D_{ii}$ are dependent on
each other such that $\text{tr}(D) = D_{xx} + D_{yy} + D_{zz} = 0$. It most cases, the Hamiltonian is written in terms of
the total spin operator so that it can be applied to the singlet and triplet states,

$$
\mathcal{H}_D = \mathbf{S}_a^T \cdot D \cdot \mathbf{S}_b \\
= \begin{bmatrix} \mathbf{S}_{ax} & \mathbf{S}_{ay} & \mathbf{S}_{az} \end{bmatrix} \cdot \begin{bmatrix} D_{xx} & 0 & 0 \\
0 & D_{yy} & 0 \\
0 & 0 & D_{zz} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{S}_{bx} \\
\mathbf{S}_{by} \\
\mathbf{S}_{bz} \end{bmatrix} \\
= S_{ax}D_{xx}S_{bx} + S_{ay}D_{yy}S_{by} + S_{az}D_{zz}S_{bz} \\
= D_{xx}\hat{s}_x^2 + D_{yy}\hat{s}_y^2 + D_{zz}\hat{s}_z^2
$$

(3.4.16)

It is more common to write the diagonal components in terms of the two independent parameters,

$$
D_1 = \frac{3}{2}D_{zz}
$$

(3.4.17)

$$
D_2 = \frac{1}{2}(D_{xx} - D_{yy})
$$

(3.4.18)

(Note that these parameters are typically represented as $D$ and $E$. The notation is replaced here by $D_1$ and
$D_2$ respectively so that the energy $E$ not be confused with the latter zero field splitting parameter $E$.)

72
Using $D_{xx} + D_{yy} + D_{zz} = 0$ and equations (3.4.17) and (3.4.18), the diagonal comments can be rewritten as,

$$D_{zz} = \frac{2}{3}D_1$$  \hspace{1cm} (3.4.19)

$$D_{xx} = -\frac{1}{3}D_1 + D_2$$  \hspace{1cm} (3.4.20)

$$D_{yy} = -\frac{1}{3}D_1 - D_2$$  \hspace{1cm} (3.4.21)

Plugging these values back into (3.4.16) yields,

$$\mathcal{H}_D = D_{xx}\hat{S}_x^2 + D_{yy}\hat{S}_y^2 + D_{zz}\hat{S}_z^2$$

$$= D_1 \left[ \hat{S}_z^2 - \frac{1}{3} (\hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2) \right] + D_2 \left[ \hat{S}_x^2 - \hat{S}_y^2 \right]$$  \hspace{1cm} (3.4.22)

Using the following quantum mechanical operators,

$$\hat{S}_x = \frac{i}{2} (\hat{S}_+ + \hat{S}_-)$$

$$\hat{S}_y = \frac{-i}{2} (\hat{S}_+ - \hat{S}_-)$$

$$\hat{S}_+ = \hat{S}_x + i\hat{S}_y$$

$$\hat{S}_- = \hat{S}_x - i\hat{S}_y$$

it can be shown that,

$$\hat{S}_x^2 - \hat{S}_y^2 = \hat{S}_x^2 + \hat{S}_y^2 = \frac{1}{2} (\hat{S}_x^2 + \hat{S}_y^2)$$  \hspace{1cm} (3.4.23)

And, realizing that $\hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = \hat{S}_z^2$ has eigenvalues of $s(s + 1)$, the Hamiltonian can rewritten as,

$$\mathcal{H}_D = D_1 \left[ \frac{\hat{S}_z^2 - \frac{s(s + 1)}{3}}{3} \right] + \frac{1}{2} D_2 \left[ \hat{S}_+^2 + \hat{S}_-^2 \right]$$  \hspace{1cm} (3.4.24)

And when placed in a magnetic field, the Hamiltonian is described by,
\[ \mathcal{H} = g \mu_B \vec{S} \cdot \vec{B} + D_1 \left[ \hat{S}_z^2 - \frac{s(s+1)}{3} \right] + \frac{1}{2} D_2 [\hat{S}_x^2 + \hat{S}_y^2] \]  \hspace{1cm} (3.4.25)

Using the raising and lowering operators introduced in the section 3.2,

\[ \hat{S}_+ |s, m_s\rangle = \sqrt{(s(s+1) - m_s(m_s + 1))} |s, m_s + 1\rangle \]
\[ \hat{S}_- |s, m_s\rangle = \sqrt{(s(s+1) - m_s(m_s - 1))} |s, m_s - 1\rangle \]

the following operators can easily be constructed,

\[ \hat{S}_+^2 |s, m_s\rangle = (s(s+1) - m_s(m_s + 1)) |s, m_s + 2\rangle \]  \hspace{1cm} (3.4.26)
\[ \hat{S}_-^2 |s, m_s\rangle = (s(s+1) - m_s(m_s - 1)) |s, m_s - 2\rangle \]  \hspace{1cm} (3.4.27)

Utilizing these operators in the Hamiltonian described by equation (3.4.25) results in,

\[ \mathcal{H} |s, m_s\rangle = \left[ m_s g \mu_B B_0 + D_1 \left( m_s^2 - \frac{s(s+1)}{3} \right) + \frac{1}{2} D_2 \hat{S}_z^2 + \frac{1}{2} D_2 \hat{S}_x^2 \right] |s, m_s\rangle \]  \hspace{1cm} (3.4.28)

Operating on the triplet singlet basis functions yields,

\[ \mathcal{H} |1, +1\rangle = \left( g \mu_B B_0 + \frac{D_1}{3} \right) |1, +1\rangle + D_2 |1, -1\rangle \]
\[ \mathcal{H} |1, 0\rangle = -D_1 \frac{2}{3} |1, 0\rangle \]
\[ \mathcal{H} |1, -1\rangle = \left( -g \mu_B B_0 + \frac{D_1}{3} \right) |1, -1\rangle + D_2 |1, 1\rangle \]
\[ \mathcal{H} |1, -1\rangle = 0 \]

Note that the singlet state is characterized as having zero energy while the triplet states are no longer degenerate at zero magnetic field. The splitting of the triplet states are defined by the zero-field parameters \( D_1 \) and \( D_2 \). The Hamiltonian for the triplet states can be written in the triplet basis as,
Here, the matrix is no longer diagonal which means the eigenvalues simply cannot be extracted from the diagonal components. Instead, the eigenvalues (allowed energies $E$) can be found by utilizing the conventional method to extract eigenvalues from a matrix, that is, using the characteristic equation. This can be found by evaluating, $\det(\mathcal{H} - EI) = 0$. This will result in a polynomial with the roots equal to $E$. These will be the allowed energies of the system and are given by [46],

\[ E_{r_0} = -\frac{2}{3}D_1 \]  
\[ E_{r_{+1}} = \frac{D_1}{3} + \sqrt{g^2\mu_B^2B^2 + D_2^2} \]  
\[ E_{r_{-1}} = \frac{D_1}{3} - \sqrt{g^2\mu_B^2B^2 + D_2^2} \]  
\[ E_{s_0} = 0 \]

and allowed energies for EPR are given by [46].

\[ \Delta E = E_{r_{+1}} - E_{r_0} = E_{r_0} - E_{r_{-1}} = D_1 \pm \sqrt{g^2\mu_B^2B^2 + D_2^2} \]

These energies are illustrated in figure 3.10. It should now be clear why many spin dependent transport phenomenon can be observed at or very near zero-field. The simple addition of the exchange and dipolar interactions allows one to understand the splitting that occurs at zero Gauss. Solving for the energies becomes more difficult when one takes into account all of the interactions that the electron spin pair experience. Here, they will interact with the externally applied field, the neighboring magnetic nuclei, and with each other. Grouping all of the interactions into a single Hamiltonian yields,

\[ \mathcal{H} = g_e\mu_B (\vec{S}_A + \vec{S}_B) \cdot \vec{B} + \sum_i A_i (\vec{S}_A + \vec{S}_B) \cdot \hat{I}_i + J_0 \vec{S}_A^T \cdot \vec{S}_B + \vec{S}_A^T \cdot \vec{D} \cdot \vec{S}_B \]  

(3.4.35)
Figure 3.10: Singlet triplet mixing due to dipolar coupling between two electrons. As a result, recombination will be enhanced at these tiny fields (very close to 0G) at which the mixing occurs.

Neglecting the latter two terms in equation 3.4.35, one is able to better understand why hyperfine interactions can be detected using the zero-field phenomenon. It should be clear at this point that when the local spin pair experiences zero magnetic field, a recombination response can be detected due to the mixing of singlet and triplet energy levels. However, the defect site will not contribute to recombination when the externally applied magnetic field is zero and in the vicinity of a magnetic nucleus. The defect will only contribute to recombination when the externally applied magnetic field cancels the hyperfine field of the nucleus. As a result, a recombination response can be detected at a field that corresponds to the hyperfine field. A simple example is illustrated in figure 3.11.

It should be pointed out that the zero-field hyperfine side peaks will not have the same splitting as they would in the high field case. An exact correspondence should not be expected for several reasons. Among them, at extremely low external fields, the nuclear magnetic moment quantization axis is likely different than what it would be at high fields. At exactly zero-field, the axis of quantization is defined by the field the electron experiences from its local magnetic environment while at relatively high magnetic fields (a 3 kG field is often enough for this to be so), it is defined primarily by the externally applied field. At low fields (< 100 G) however, the field the electron experiences is the vector sum of the electron’s local magnetic surroundings and externally applied magnetic field. We speculate that at these very low fields, the electrons and nuclei will tend reorient their axes of quantization so they remain parallel with each other because of the fields they exert on each other. Therefore, the local magnetic field that the electron experiences from the neighboring nucleus will be different when the externally applied magnetic
field is low. As a result, the electron nuclear hyperfine peaks will appear to have shifted from the high field case. This concept is illustrated figure 3.12.

\[ B_{\text{local}} = B_0 + B_{N1} = 0 \quad \text{Figure 3.11: Example illustrating a } 66.67\% \text{ abundant spin } I = \frac{1}{2} \text{ magnetic nucleus that is in the vicinity of the electron spin pair. At zero applied magnetic field, a recombination response will be observed due to singlet triplet mixing. However, the magnetic nucleus will exert a magnetic field on the spin pair. Therefore, additional recombination responses can be detected when the externally applied magnetic field cancels the hyperfine field.} \]

\[ B_{\text{local}} = B_0 = 0 \quad \text{recombination} \quad B_{\text{local}} = B_0 + B_{N2} = 0 \]

\[ B_0 = -B_{N1} \quad B_0 = 0 \quad B_0 = -B_{N2} \]

Figure 3.12: Reorientation of the electron and nuclear quantization axes at high, moderate, and low magnetic fields.

\[ B_e \ll B_{\text{ext}} \quad B_N \ll B_{\text{ext}} \ll B_e \quad B_{\text{ext}} \approx 0 \]

\[ \mu_N \]

\[ m_I = -1/2 \]

\[ \text{quantization axis} \]

\[ B_{\text{ext}} \]

\[ B_{\text{local}} \]
It was just shown that the electron exchange interaction (Coulombic) was responsible for splitting the singlet and triplet energy levels, where the triplets are degenerate at zero magnetic field. The dipolar interaction (magnetic) is a similar interaction in that $D$ in this case takes into account the dipole field that each electron exerts on one another. This interaction is responsible for ridding the triplet states of their degeneracy at zero magnetic field. In either case, mixing of triplet and singlet states leads one to “induce” recombination without the use of electromagnetic radiation to flip the spins of electrons. Rather, mixing the singlet and triplet states will enhance the recombination current in a fully processed device. This suggests the use of this phenomenon in many areas of research.

The first and most obvious application of the zero-field phenomenon is to use it as a spectroscopy tool to study SDR or SDT in fully processed devices. In just about all devices that we have tested, we have about always observed a zero-field signal that we have tentatively attributed to SDR or SDT. However, in our recent study which utilized a BJT with a well characterized pn junction, we were able to clearly demonstrate that the zero field response is in fact due to SDR (details discussed in chapter 7). Also, because we observe a split peak at zero field (indicating the presence of spin-spin dipolar or exchange interactions), we feel that this is further evidence of precursor pair formation prior to recombination. We also notice that the amount of inflection is highly dependent upon applied electric field bias, a result indicating that the applied bias is affecting the exchange and or dipolar interactions of the spin pair. The splitting at zero field into two or three levels provides information about which type of interaction is occurring between the electrons involved in recombination. In addition, the technique is well suited for extracting electron-nuclear hyperfine interactions as previously discussed. The hyperfine interactions are arguably the most useful parameter that can be obtained from a conventional EPR and EDMR measurement. Our technique demonstrates that all of these interactions can be obtained very easily and inexpensively.

The zero field phenomenon may also possibly be used for the electrical detection of zero-field electron nuclear double resonance (ENDOR). Conventional ENDOR entails the use of continuous wave electromagnetic radiation and a constant magnetic field to fully saturate the electron spin system. Then, the frequency (RF) of an additional magnetic field is swept and used to flip the spins of the magnetic nuclei which allows one to extract detailed information about the atomic structure of the defect site. ENDOR is essentially a way to detect NMR via EPR. With zero-field SDR, the electron spin system is effectively saturated without the use of electromagnetic radiation to flip electron spins. This suggests the possibility of electrically detecting the NMR transitions in a device by sweeping the frequency of a single EM radiation source at extremely low magnetic fields.

This new physics can also be applied to device reliability. One particular example is to study spin dependent leakage currents through a dielectric via spin dependent hopping or tunneling. Because trap-to-
trap tunneling conserves angular momentum, spin dependent tunneling is enabled when unpaired electrons in neighboring defects form singlet pairs. If this is the case, then coupled electron spins residing in neighboring traps would exhibit a singlet triplet mixing phenomenon at zero magnetic field. As the pair states mix, the electrons would be capable of tunneling from one defect to another, thereby creating a random percolation path through the dielectric. However, this mechanism only works when the unpaired electrons residing in neighboring traps are able to couple. The response at zero magnetic field (one or multiple peaks and distance between the peaks) could provide much insight into the number and separation of traps along the percolation path within the dielectric. This makes the technique very attractive for those interested in studying hopping and tunneling mechanisms in thin film dielectrics.

Also, we were able to show that the low-field resonant and zero-field responses behave similarly as a function of bias and time. This demonstrates that both responses arise from the same underlying mechanism, that being defects that allow for tunneling. Thus, the data implies that the zero-field response could be utilized to study time dependent dielectric breakdown (TDDB) and mapping energy levels of the traps involved in the hopping process. The other device reliability problem in which this technique can be useful is the study of the bias temperature instability (BTI) in MOSFETs. We demonstrate the power of this technique with application to Si MOSFETs in chapter 7.

The zero-field phenomenon could also be used to study spin dependent transport in devices that involve material systems with defects that are difficult to saturate (i.e. difficult to increase the spin temperature) with EM radiation. As mentioned earlier in this section, it is important to saturate the spin system to observe the maximum spin dependent response. In certain material systems such as III-Vs where very broad magnetic resonance lines are inevitable, very large oscillating $B_1$ magnetic fields are needed to saturate the spin system. Many times, a fully saturated spin system cannot be achieved with reasonable EM radiation power levels, thereby making the electrical detection of magnetic resonance almost impossible. Detecting spin dependent transport at zero-field allows one to detect an essentially fully saturated response because the randomization resulting from a high spin temperature at resonance is effectively achieved by the mixing of triplet and singlet states at zero magnetic field.

Another application of this new physics is the development of very inexpensive and portable spectrometers. Because large electromagnets, high power/frequency electromagnetic radiation sources, and cooling systems are not required, the footprint of the low- / zero-field spectrometer can be made extremely small with a price tag that is about two orders of magnitude cheaper than that of a commercial X-band spectrometer. Also, because the size of many spin dependent phenomena are independent of the applied magnetic field, there is not much to sacrifice when making measurements at low-field. These spectrometers would allow for simultaneous detection of resonant and zero-field spin dependent transport.
And finally, there are many potential applications, outside of the area of research, that this new physics can be utilized. For example, it has the potential to be incorporated in the design of new magnetic field sensors and magneto-isolation circuits. These applications are further discussed in chapter 7.

We think that the zero-field phenomenon hasn’t been exploited as much as it should because no one has realized its true potential. The true potential is only uncovered when one utilizes magnetic field modulation. The use of this signal processing technique allows spin dependent transport mechanisms to pass at a certain frequency, phase, and direction. Because it this technique in particular that allows us to observe the hyperfine interactions at very low fields, it, along with other signal processing techniques, will be discussed in the upcoming chapter.
CHAPTER 4 – SIGNAL PROCESSING OF SPIN DEPENDENT SIGNALS

As advancements are continually made in semiconductor technology, device performance significantly improves due to the fewer defects present within the devices. As a result, the size of the magnetic resonance signals observed is reduced. Therefore, spectrometer artifacts and device noise have more of an impact when attempting to achieve high SNR magnetic resonance spectra. Therefore, many signal processing techniques and algorithms have been utilized in order to increase SNR in these measurements [62] [63] [64] [65] [66] [67] [68] [69]. Some of the noise sources that are associated with our EDMR measurements include the ambient noise from the surrounding hardware and most importantly, the internal shot, thermal, and flicker noise arising from within the device under observation. There are many things that can be done to improve the SNR in terms of hardware such as proper grounding, shortening of cables to reduce stray capacitance, and utilizing low noise preamplifiers. Many of these things have been done in preliminary work for this thesis project. The focus of this section will be on advancements made in software.

4.1 – Noise Observed in EDMR

There are many sources of noise in EDMR as illustrated in figure 4.1. The sources of noise arise from within the device under test (DUT) itself as well as from the spectrometer. Some of these noise sources include electromagnetic interference (EMI), thermal, and quantization noise that arise from the hardware of the spectrometer. However, it is the shot, flicker, and thermal noise within the device that ultimately limits the detection of low number of defects.

![Diagram of sources in EDMR](image)

Figure 4.1: Sources of noise in an EDMR spectrometer. Shot, flicker, and thermal noise arise from the DUT, electromagnetic interference (EMI) arises from spectrometer hardware, and additional thermal noise arises from the preamp, and the sampler or digitizer is responsible for quantization noise.

There have been many attempts to model the noise within a device, all of which contain three different sources of noise. The three primary sources of current noise are flicker (1/f noise), thermal (Johnson) noise, and shot noise which all are usually present in semiconductor devices [70] [71]. Thermal noise is a white process meaning that its spectral density remains constant for all frequencies. This type of noise arises from the random motion of molecules at temperatures greater than 0 K. The thermal current $I_{thermal}$ can be described by...
\[ I_{\text{thermal}} = \sqrt{\frac{4kT\Delta f}{R}} \] (4.1.1)

where \( k \) is Boltzmann’s constant, \( T \) is the temperature in Kelvins, \( \Delta f \) is the bandwidth in Hertz, and \( R \) is the internal resistance of the device. Shot noise is also a white noise process that arises due to the fact that current is not actually continuous in nature but rather a summation of individual charge carriers. Shot noise arising from current can be described by

\[ I_s = \sqrt{2qI\Delta f} \] (4.1.2)

where \( q \) is the electric charge and \( I \) is the current through the device. The final source of noise observed in semiconductor devices is flicker noise (1/f noise) which arises from recombination noise or other semiconductor impurities and is usually associated with direct current [71]. It is not a white noise process, but rather frequency dependent. Flicker noise current due to currents in a semiconductor device may be described as

\[ I_f(f) = \sqrt{KF \cdot I^{AF} \Delta f} \] (4.1.3)

where \( KF \) is the flicker noise coefficient, \( AF \) is the flicker noise exponent, and \( f \) is the frequency. Note that of the three noise processes just described, this is the only one that is a function of frequency. It is important to note that \( KF \) varies from device to device and \( AF \) is typically 1. Typical values for silicon devices are \( KF = 10^{-16} \) and \( AF = 1 \) [71]. All of these noise sources arise in almost all semiconductor devices, though some dominate others in particular devices.

The left image in figure 4.2 illustrates the frequency spectrum for a MOSFET biased under three different conditions. The top figure represents 0.0 volts applied to the diode (0 \( \mu A \)), the middle figure represents 1.0 volt applied to the diode (0.002 \( \mu A \)), and the bottom figure represents 2.0 volts applied to the diode (5 \( \mu A \)). The top figure basically represents the noise generated by the preamp because the diode was essentially off. This value is consistent with the current preamplifier’s specifications as listed in the manual [72]. The middle figure indicates that not a significant amount of current was generated by the device to contribute to the noise spectrum with a 1.0 volt forward bias. Though, when 2.0 volts was applied to the diode, flicker noise (which is usually present with direct current) becomes apparent in the bottom figure. In order to model the noise at this applied voltage, we simulated the flicker, shot, thermal,
and preamp noise as illustrated in the right plot if figure 4.1. The top plot illustrates the simulations for each of these noise components. The bottom plot illustrates the summation of these components plotted against the noise spectrum of the device. It is obvious that there is a strong quantitative correlation between simulation and experiment which indicates that all sources of noise are most likely contributing to the total noise.

The data collected in the previous figure was gathered with the SR570 preamp with a 200kHz bandwidth setting. One way to reduce the noise is to reduce the bandwidth in the measurement. All amplifiers are limited in gain for a particular bandwidth. As the amplifier gain increases, the bandwidth decreases and as the gain decreases, the allowable bandwidth increases. This tradeoff is typically referred to as the gain bandwidth product and is usually specified with a gain of one. Therefore, if an amplifier is speced at a 500MHz, this simply means that the amplifier will have unity gain at 500MHz. The reason for this is because there are voltage swing limitations within the amplifier. Understanding this tradeoff is very important when attempting to maximize SNR in conventional EDMR measurements. Recently, there has been a very promising modification in the detection scheme that goes away from this conventional detection in the audio range but rather in the RF region. The one example is that of EDMR using RF reflectometry [73].

Typical EDMR measurements entail detecting spin dependent currents in the fA to nA range. This requires preamp gains in the range of $10^{15}$ to $10^7$ Volt/Amp to observe a voltage response that can be sampled by an analog-to-digital converter (ADC).

![Image](image_url)

Figure 4.2: (left) Frequency spectrum of a MOSFET under 3 different biasing conditions. The numbers in red indicate the amount of current flowing through the device. (right) simulated flicker, thermal, shot, and preamp noise for a MOSFET that has 5 uA of current passing through it. The bottom plot compares the summation of the simulated noise sources to the frequency spectrum of a MOSFET that has 5 uA of current passing through it. Note that the amount of flicker noise that is present is dependent on the amount of DC current flowing through the device.
At such high gains, the preamplifier bandwidth is limited to a few kHz all the way down to a few Hz. One may then attempt to amplify with multiple stages to increase the bandwidth. The first amplifier stage is the most sensitive because it adds its own thermal noise to the system. In a really sensitive measurement like EDMR where very small currents need to be amplified, it is very important to give as much amplification to the signal at the analog front end so that it can be lifted above the front end amplifier noise. If it is not, then any additional gain stages will amplify the preamp noise as well as the signal by the same amount. As a result, the signal will never be pulled out of the noise floor. One may ask why not just perform the measurement at DC. The problem with this is that flicker noise arises from about just all instruments and devices. This is specifically why magnetic field modulation is utilized. It essentially gives the resonance condition a frequency dependence. If a preamplifier is to be designed for EDMR, a 4-stage amplifier should be utilized as illustrated in figure 4.3.

Figure 4.3: Illustration the basic components of a preamplifier that should be used with an EDMR spectrometer.

A few rules of thumb should be followed when performing an EDMR measurement with this form of preamplifier (e.g. SR570).

(1) The first stage gain setting should be set as high as possible such that it doesn’t overload the preamp. If a particular gain setting is desired which overloads the preamp, use the input offset to cancel this current. (Note that adding current to cancel the DC input current will NOT completely remove the flicker noise associate the large input currents. The technique simply allows for the utilization of larger gain settings without overloading the preamplifier front end circuitry.)

(2) Use a second stage high pass filter to block the DC component (and much of the flicker noise) that gets passed from the first stage. This will allow for the modulated signal to be further amplified without overloading the following gain stage. Also, this removes the DC offset and allows for the best dynamic range setting of the ADC. The cutoff frequency of the filter should be set to $\approx$10x lower than the modulation frequency to prevent attenuation to the signal.

(3) Chose a modulation frequency such that it is greater than the flicker noise but lower than bandwidth of the preamplifier. (200-1.8kHz modulation frequency works well with the SR570)
(4) Use low pass (antialiasing) filters to prevent aliasing of the ADC. The cutoff frequency should be set higher than the modulation frequency but lower than the Nyquist rate (Sampling Rate in Hz / 2). If really high gains on the preamplifier are used, a low pass filter may not be necessary if sampling fast enough because the bandwidth limitation of the preamp will provide an 8dB/octave attenuation rolloff.

(5) Ensure that the sample rate on the ADC is more than 10x that of the modulation frequency. (This provides 10 samples of data per modulation period, a value 5 times more than obtained when using the Nyquist rate.)

(6) Ensure that the signal entering the ADC does not exceed the quantization limits for the ADC. This can be determined by the red indicator lights on the Stanford Instruments lock-in amplifiers or visually on the Labview GUI when utilizing the virtual lock-in amplifier. This will cause distortion in the EPR spectrum.

An important point to make is that the noise generated by the preamp dominates the noise in EDMR for small device currents, while flicker noise dominates at larger device currents. This suggests that the SNR of an EDMR measurement is independent of device size down to a particular size of the device because flicker and shot noise sources are directly proportional to the amount of current flowing through it. At some point, the preamp noise will dominate the noise spectrum. This is precisely what we have observed in our earlier measurements. Figure 4.4 illustrates (left) an individual EDMR spectra recorded for three differently sized devices and (right) a plot of the SNR versus device size.

![Figure 4.4: Individual EDMR scans of identical devices with gate areas of 500 μm², 40,000 μm², 160,000 μm² and SNR of EDMR spectra for each measurement.](image-url)
It is obvious from the figure that each EDMR trace has roughly equivalent SNR even though the device size varies by orders of magnitude. These figures illustrate the possibility of observing defects in transistors of very small size. A few years back, the ultimate goal was to observe a single spin. This has been accomplished by the observation of tunneling electrons in double quantum dots used for quantum computing applications [20] [21]. It is worth noting that for these measurements, current flowed only when the electrons in the double quantum dot were in resonance. As a result, there was no background noise (due to large DC currents) to deal with which is why the detection of this single spin was possible. In the case of defect detection in microelectronics, single defect detection is much more difficult (and may be unrealizable for some devices) because of the background currents that are involved. Semiconductor manufacturing companies have various processing methods which attempt to reduce the number of defects, and hence recombination. The electrons that were previously captured are now getting through without recombining with holes. This increases the DC throughput current in the device and hence, increase the noise as well. This makes detecting spin dependent currents in the fA range extremely difficult to detect in modern microelectronics designed for optimizing DC current efficiency.

4.2 – Lock-in Amplification

Lock-in amplification was invented in 1954 by Robert H. Dicke and is one of the most reliable techniques used to remove noise in sensitive signal measurements. Continuous wave EPR and EDMR typically utilize a sinusoidal modulation of the applied magnetic field so to encode the MR signal in a narrow frequency band. The amplitude of the detected modulated signal is a measure of the MR signal. A lock-in amplifier (LIA) is then used to demodulate the amplitude modulated signal to DC, thus exploiting the sensitivity enhancement available from the phase and frequency detection. This widely used method effectively attenuates much of the noise associated with the 1/f noise typically observed in these experiments.

Today, digital lock-in amplifiers (LIAs) are preferred over analog designs for many reasons [74]. The primary reason digital LIAs are more widely utilized is because the output is less prone to drift because there are no analog components with performance which may vary with time and temperature. Also, digital LIAs have more stable and precise oscillators. They are capable of achieving perfect demodulation of in-phase and quadrature components, and they can be purchased at relatively lower prices than that of an analog design. A few years back, the one situation in which an analog LIA would be preferred over a digital one would be when high frequencies are utilized (ie: > 1MHz). This is because, a few years back, processor speeds in digital design lacked the capability of fast computation rates. However, because of advancements made in hardware and software, Zurich Instruments recently designed a 600 MHz digital LIA with 1.8GS/s which is now on the market.
The main components of a LIA consist of an optional bandpass filter, a mixer or multiplier, a phase lock loop (PLL), a phase shifter, and a low pass filter [74] [75]. A PLL, composed of a voltage (analog) controlled oscillator (VCO) or numerically (digital) controlled oscillator (NCO) with feedback loop filter, is one of the most important components of a LIA and is used to track the phase and frequency of an external reference signal so that successful synchronous demodulation of amplitude modulated signals can be achieved. However, this component is not required if the reference and modulation signals are generated by the same source. This is precisely the LIA configuration upon which our Labview virtual lock-in amplifier (VLIA) is based. It doesn’t utilize a phase lock loop (PLL) because the modulation signal and the reference signal are generated by Labview software (i.e. the same internal clock). Therefore, these signals have exactly the same frequency and reference phase every time they are generated.

The EDMR signal that is received at the output of the VLIA is actually the derivative of the spin dependent current. This is the case because the amplitude modulated EDMR signal flips sign after the DC field is swept past the center field of the resonant response. As a result, the amplitude modulated sinusoid will exhibit a 180° phase shift which results in a zero crossing at the resonant field when demodulated down to DC. Therefore, the observed EDMR spectrum resembles the derivative of the actual EDMR absorption response. It is actually more desirable that the derivative is obtained because information can be extracted more easily from the spectra. EDMR spectra may contain very detailed structure such as those resulting from hyperfine interactions; these interactions may be very difficult to identify in a simple EDMR absorption curve. The derivative reveals modest changes in slope which allows for easier identification of somewhat subtle features.

The basic operation of a LIA is an easy concept to understand and can be explained mathematically in the time domain or in the frequency domain. The latter will be described here. The amplitude modulated discrete-time signal of interest \( x(n) \) can be generally defined as,

\[
x(n) = d(n) \cos(\omega_d n + \theta_d) + \sum_{\omega_u=0}^{\infty} \sum_{\theta_u=0}^{2\pi} u_{\omega_u,\theta_u}(n) \cos(\omega_u n + \theta_u) \quad (4.2.1)
\]

where \( n \) is the discrete-time index, \( d(n) \) is the deterministic signal under observation modulated with normalized digital carrier frequency \( \omega_d \) and phase \( \theta_d \), and \( u_{\omega_u,\theta_u}(n) \) is the random noise component with normalized digital frequencies \( \omega_u \) and phases \( \theta_u \). In order to analyze the LIA process in the frequency domain, the signal \( x(n) \) must first be transformed from the time domain to the frequency domain \( X(\omega) \) via the discrete-time Fourier transform (DTFT) described by [76]
\[ X(\omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} \quad (4.2.2) \]

The DTFT differs from the continuous-time Fourier transform (CTFT) in the sense that, as its name implies, it computes the Fourier transform of a sequence of discrete samples. Note that the DTFT is a continuous function of the frequency variable \( \omega \). Therefore, it is common to use discrete Fourier transform (DFT) when digital computation needs to be performed in the frequency domain. For the case of this derivation, the DTFT will be sufficient to analyze the VLIA. In order to find the DTFT of the input signal \( x(n) \), a spectral shifting property can be utilized (which is straightforward to derive) instead of computing it directly. An arbitrary discrete time sequence \( g(n) \) multiplied by a sinusoid of frequency \( \omega_d \) and phase \( \theta_d \) is transformed into the frequency domain using the following relationship,

\[ g(n) \cos(\omega_o n + \theta_o) \leftrightarrow \frac{1}{2} \left[ G(\omega - \omega_o) e^{j\theta_o} + G(\omega + \omega_o) e^{-j\theta_o} \right] \quad (4.2.3) \]

where \( G(\omega) \) is the DTFT of the discrete time sequence \( g(n) \). This relation states that \( \frac{1}{2} \) of \( G(\omega) \) is shifted in frequency \( -\omega_o \) and \( \frac{1}{2} \) of \( G(\omega) \) is shifted in frequency \( +\omega_o \). The phase \( \theta_o \) is usually taken to be 0 so the exponential phase constants vanish in (4.2.3). Using this relation and the fact that the DTFT is a linear operation, the DTFT of the input signal \( x(n) \) can be computed as follows

\[ X(\omega) = \frac{1}{2} \left[ D(\omega - \omega_d) e^{j\theta_d} + D(\omega + \omega_d) e^{-j\theta_d} \right] \]

\[ + \frac{1}{2} \sum_{\omega_u=0}^{\infty} \sum_{\theta_u=0}^{2\pi} \left[ U_{\omega_u,\theta_u}(\omega - \omega_u) e^{j\theta_u} + U_{\omega_u,\theta_u}(\omega + \omega_u) e^{-j\theta_u} \right] \quad (4.2.4) \]

where \( D(\omega) \) is the DTFT of \( d(n) \) and \( U_{\omega_u,\theta_u}(\omega) \) is the DTFT of \( u_{\omega_u,\theta_u}(n) \). This signal is first fed into a bandpass filter. For ease of analysis, it will be assumed that an ideal bandpass filter is used and is defined by

\[ |H_{BP}(\omega)| = \text{rect} \left( \frac{\omega - \omega_o}{2B} \right) \quad (4.2.5) \]
where $B$ is the bandwidth of the filter and is assumed to be narrow. The filtering operation is computed in the time domain by convolving the input $x(n)$ with the filter $h_{BP}(n)$ and is simply computed in the frequency domain by multiplication of their Fourier Transforms.

$$x(n) * h(n) \leftrightarrow X(\omega) \cdot H(\omega) \quad (4.2.6)$$

Using this property, the DTFT $X_{BP}(\omega)$ of the bandpassed signal $h_{BP}(n)$ becomes,

$$X_{BP}(\omega) = \frac{1}{2} \left[ D(\omega - \omega_d)e^{j\theta_d} + D(\omega + \omega_d)e^{-j\theta_d} \right]$$

$$+ \frac{1}{2} \sum_{\theta_u=0}^{2\pi} \left[ U_{\omega_u,\theta_u} (\omega - \omega_d) e^{j\theta_u} + U_{\omega_u,\theta_u} (\omega + \omega_d) e^{-j\theta_u} \right] \quad (4.2.7)$$

This result assumes that the ideal bandpass filter completely removes the noise components at all frequencies except $\omega_d$. In practice, an ideal filter is unrealizable so there will always be noise present over the transition bands of the filter that pass without attenuation. As a result, the noise that now resides on $x(n)$ is the noise that is present over the bandwidth of the filter at random phases. The bandpassed signal $x_{BP}(n)$ is then fed through a mixer or multiplier. This operation multiplies the reference signal $r(n)$, defined by

$$r(n) = \cos(\omega_d n + \theta_d) \quad (4.2.8)$$

with the bandpassed signal $x_{BP}(n)$. The phase shifter is used to adjust the phase of the reference signal so that it is equal to the phase of the incoming signal $\theta_d$. Using the property given in equation (4.2.3) and grouping like terms, the DTFT of the mixed signal $x_M(n)$ becomes

$$X_M(\omega) = \frac{1}{2} \left[ D(\omega) + U_{\omega_u,\theta_u} (\omega) \right]$$

$$+ \frac{1}{4} \left[ D(\omega - 2\omega_d)e^{j2\theta_d} + D(\omega + 2\omega_d)e^{-j2\theta_d} \right]$$

$$+ \frac{1}{4} \left[ U_{\omega_u,\theta_u} (\omega - 2\omega_d)e^{j2\theta_u} + U_{\omega_u,\theta_u} (\omega + 2\omega_d)e^{-j2\theta_u} \right] \quad (4.2.9)$$

This result assumes that the mixing operation completely removes all noise components with dissimilar phase and frequency (ie: all $\omega_u \neq \omega_d$ and all $\theta_u \neq \theta_d$) as that of the reference signal. In reality, these
noise components are not completely removed but merely attenuated so they will partially pass through the mixing operation. The mixed signal \( x_M(n) \) is then fed through a low pass filter \( x_{LP}(n) \) which completely removes the higher order harmonics produced by the mixing operation and attenuates the remaining noise riding on the signal. The DTFT \( X_{LP}(\omega) \) of this low pass filtered signal \( x_{LP}(n) \) becomes,

\[
X_{LP}(\omega) = \frac{1}{2} [D(\omega) + U_{\omega_d,\theta_d}(\omega)]
\]  

(4.2.10)

The inverse discrete-time Fourier transform (IDTFT) is described by

\[
x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\omega n} d\omega
\]

(4.2.11)

and is used to transform the frequency domain signal \( X_{LP}(\omega) \) to its time domain representation \( x_{LP}(n) \)

\[
x_{LP}(n) = \frac{1}{2} [d(n) + u_{\omega_d,\theta_d}(n)]
\]

(4.2.12)

Therefore, the dominant source of noise observed at the output of the LIA is the noise present over the bandwidth of the modulated signal with similar phase. The output low pass filter is unable remove the noise that resides over the bandwidth of the signal but can further attenuate the leakage noise that is present due to the non-ideal performance of the filter and mixing operations. The amount of noise that is removed is controlled via the time constant parameter. The time constant is inversely proportional to the cutoff frequency of the filter so less noise will pass for higher time constants. The cutoff frequency \( f_c \) is defined by the frequency at which the power of the input signal is attenuated to \( \frac{1}{2} \) of the total power. The time constant \( \tau \) for a first order filter is usually defined by

\[
\tau = \frac{1}{2\pi f_c}
\]

(4.2.13)

Caution must be taken when choosing the order of the output low pass filter because higher order digital filters usually result in an undesirable phenomenon known as ringing. As the order of a low pass filter is increased, the transition frequency band of the filter narrows. As a result, the shape of the filter approaches an ideal filter which is defined by the \( rect(\omega) \) function in the frequency domain. The IDTFT of a \( rect(\omega) \) function is the \( sinc(n) \) function, defined by \( \sin(n)/n \), which has side lobes whose
amplitude increases with filter order. The amplitude of these side lobes determines the amount ringing that occurs in the filtering operation. More ringing occurs for filters of higher order. Butterworth filters typically experience ringing when the order is greater than 2. Bessel filters on the other hand are more preferable because ringing is not so much of a problem for higher order filters.

LIAs are extremely valuable to magnetic resonance because they significantly enhance the capability of detecting the spin dependent recombination current in modern devices. However, despite the significant improvement that lock-in detection can provide, the amount of noise that remains in the spectra is significant enough that, in almost all cases, other forms of signal processing are still needed to interpret the spectra. The most common technique used on top of lock-in detection is signal averaging.

4.3 – Signal Averaging

Signal averaging is another widely used form of noise reduction in sensitive signal measurements and is therefore commonly utilized in continuous wave ESR and EDMR experiments. The technique has been around for some time [77] and is used to improve the SNR of repetitive signals obscured by noise. Today, signal averaging is commonly used in many sensitive experimental methods in which signals are repetitively acquired in the presence of noise. Some of these experimental methods include medical imaging, electrocardiography, or electroencephalography [78] [79] [80]. The basic idea of signal averaging is to successively add a series of scans so that the random nature of the noise cancels itself out over time. A derivation of the amount of noise removed by signal averaging is given below which will aid in the understanding of the enhancement achieved from adaptive signal averaging discussed in section 4.4. Similar derivations may be found elsewhere [81].

If an individual scan sequence $x(n)$ is composed of a deterministic desired component $d(n)$ and a random noise component $u(n)$, then the average of scans $x_N(n)$ is simply the sum of all the scan sequences divided by the number of scans $N$. Dropping the sample index $n$,

$$
\bar{x}_N = \frac{1}{N} \sum_{i=1}^{N} x_i = \frac{1}{N} \sum_{i=1}^{N} (d_i + u_i) = d + \frac{1}{N} \sum_{i=1}^{N} u_i = d + \bar{u}_N
$$

(4.3.1)

where $\bar{u}_N$ is the average of the noise. In order to determine the enhancement in SNR for the averaging technique, the variance of $\bar{x}_N$ needs to be calculated. Because the desired scan $d$ is deterministic and not random, the variance of $\bar{x}_N$ is equal to the variance of $\bar{u}_N$. This variance $\sigma^2_{uN}$ of the averaged noise reduces to,

$$
\sigma^2_{uN} = E[(\bar{u}_N - \mu)^2] = E[\bar{u}_N^2 - 2\bar{u}_N\mu + \mu^2] = E[\bar{u}_N^2] - 2\mu E[\bar{u}_N] + \mu^2 = E[\bar{u}_N^2] - \mu^2
$$

(4.3.2)
where $E[\cdot]$ is the expectation operator and $\mu$ is the theoretical mean of the noise sequence $u$. $E[\bar{u}_N^2]$ is defined as the second moment $\bar{u}_N$ of can be found by

$$E[\bar{u}_N^2] = E \left[ \left( \frac{1}{N} \sum_{i=1}^{N} u_i \right)^2 \right] = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} E[u_i u_j] = \frac{1}{N^2} \left[ \sum_{i=1}^{N} E[u_i^2] + \sum_{i=1}^{N} \sum_{j=1}^{N} E[u_i u_j] \right] = \frac{1}{N^2} \left[ \sum_{i=1}^{N} \sigma_u^2 + \sum_{i=1}^{N} \sum_{j=1}^{N} \mu^2 \right]$$

(4.3.3)

Using the relation $E[\bar{u}_N^2] = \sigma_{uN}^2 + \mu^2$ and assuming the noise is statistically independent and identically distributed (iid), the second moment of $\bar{u}_N$ becomes

$$E[\bar{u}_N^2] = \frac{1}{N^2} \left[ \sum_{i=1}^{N} \sigma_u^2 + \sum_{i=1}^{N} \sum_{j=1}^{N} \mu^2 \right] = \frac{1}{N^2} \left[ \sum_{i=1}^{N} \sigma_u^2 + \sum_{i=1}^{N} \sum_{j=1}^{N} \mu^2 \right] = \frac{\sigma_u^2 + \mu^2}{N^2} + \mu^2 \frac{N(N-1)}{N^2} = \frac{\sigma_u^2}{N} + \mu^2$$

(4.3.4)

Plugging this result back into equation (4.16) yields,

$$\sigma_{uN}^2 = E[\bar{u}_N^2] - \mu^2 = \frac{\sigma_u^2}{N}$$

(4.3.5)

The standard deviation of the averaged noise $\sigma_{uN}$ is simply equal to the square root of the variance

$$\sigma_{uN} = \frac{\sigma_u}{\sqrt{N}}$$

(4.3.6)

This result shows that the technique of signal averaging reduces the standard deviation of the noise by a factor of $\sqrt{N}$ where $N$ is the number of scans in the average. When $N$ is very large, the average converges to the desired signal because the variance of the noise is almost reduced to nothing. Later on, this derivation will be compared to that of an expedited averaging process that we created for our spectrometers. Even with lock-in amplification and signal averaging, the EDMR technique can still be quite a time consuming task to acquire high SNR in the scans. It is therefore desirable to somehow improve acquisition time using more advanced techniques.
4.4 – Adaptive Signal Averaging

Lock-in detection is not always sufficient alone to achieve a reasonable SNR, so signal averaging is almost always utilized. In many cases, for example, in the area of electrically detected MR in solid state electronics, extensive signal averaging may be required to achieve a reasonable SNR which can be very time consuming. A variety of filters have been used in addition to lock-in amplification in attempt to attenuate the noise on the individual scans prior to being averaged in hopes to expedite the noise reduction process. For example, conventional low pass smoothing and polynomial fitting filters are commonly used for this purpose in MR [65]. However, these filters encounter problems when the noise and signal have similar frequency characteristics. When this is the case, signal attenuation occurs so filters of this nature are often undesirable. (This can simply be ameliorated by sweeping the magnetic field a much slower rate; however, this adds a significant of time to the experiment.) An alternative option would be to use an adaptive filter which utilizes a reference signal to filter with. But how should one utilize an adaptive filter when one searches for signals in MR?

We have developed a way to utilize an adaptive filter in the averaging process without any prior knowledge of the signal under observation. The adaptive filter works by using the real-time conventional scan average as the desired response in an adaptive linear prediction (ALP) configuration. We term this technique adaptive signal averaging (ASA) [62] [63] [64]. Not only can this technique be applied to MR experiments, but can also be utilized in any experiment where signal averaging is utilized. By utilizing an adaptive filter, noise can be removed from a signal while retaining the quantitative and qualitative features in the desired signal, even when the noise and the signal have similar frequency characteristics. This paper describes the advancements made to the technique and provides a statistical derivation of its performance enhancement over conventional signal averaging. It also demonstrates that a noisy desired signal can in fact be effective in adaptive filtering applications.

The technique of ASA starts as any data acquisition system does which requires signal averaging. As low SNR scans are acquired, an average of these scans is continually updated. Once a weak signal becomes apparent in the average, the adaptive filter is switched on using the conventional average as the desired signal in the ALP. (Note that even though the conventional average at this point is quite far from what is desired, it still contains useful information about the signal and nature of the noise of the scans.) At this time, the previous scans can all be post processed by the algorithm if the spectrometer software allows for it. From then on, the subsequent scans are adaptively filtered and separately averaged. At the same time, the conventional average is continually updated with the unfiltered scans and is used as the desired signal in the ALP. Thus, the quality of the desired signal is continually improving and the effectiveness of the ALP is consequently also improving. (Note that because the noise of the filtered scans is reduced, the noise in the filtered average will be reduced more quickly than that of the noise in the
conventional average. Hence, the averaging process is essentially expedited via ASA. In order to achieve an additional two fold reduction in noise variance for an individual scan, the backward prediction of the desired signal is performed on every scan and then averaged with the forward prediction. (The forward prediction essentially refers to the filtering process just described (i.e. filtering the data from left to right) and the backward prediction simply refers to filtering the data from right to left.) As a result, both of these filtering processes are performed after the acquisition of every scan. The forward and backward predictions effectively produce two unique versions of the estimated spectrum because each estimated sample is calculated based on the ALP predicting in two different directions on the same data array. In order to preserve the edges of the data array, the weight vector and the autocovariance matrix are initialized with the weights and inverse autocovariance matrix from the scan before. Only when the filter is first turned are the weights initialize with zeros and the autocovariance matrix with the product of the identity matrix and the regularization parameter. A simplified block diagram of ASA is illustrated in figure 4.5.

ASA excels specifically when filtering very narrow signals: an area where conventional low pass or smoothing filters fail to perform. Conventional moving averaging filters lose their effectiveness when compensating their length in order to preserve narrower line shapes. Also, it should be noted that the ASA reduces to a conventional adaptive filter when a desired signal is known prior to experiment. If it isn’t known and the SNR for an individual scan is reasonably high, the ASA can be turned on after the first scan is acquired.

Figure 4.5: (left) Simplified block diagram of adaptive signal averaging and (right) adaptive linear predictor used in the adaptive signal averaging technique. Note that the desired signal is replaced with the unfiltered, conventional average.
However, the ASA is most useful for signals which require many scans to average out the noise, e.g. > 50. The reason for this is because the adaptive filter is typically turned on after about 15 to 25 scans have been averaged. (The reason for this turn-on time is because the SNR improves as the square root of the number of scans, most of the noise attenuated in signal averaging occurs within this time frame.) As a result, the effectiveness of ASA will reveal itself after an additional 15 to 25 scans have been adaptively filtered and then averaged. (Note however that the effectiveness of the ASA can actually be observed without the additional 15 to 25 scans if the spectrometer software is flexible enough to allow the user to post process the first 15 to 25 at that time.)

4.4.1 – ASA Algorithm

Consider the noisy signal \( x(n) \) that is to be repetitively acquired and averaged. It is composed of a noise free, deterministic desired signal \( d(n) \) and a random noise component \( u(n) \).

\[
x(n) = d(n) + u(n)
\] (4.4.1)

where \( n \) is the sample index of the scan. The average of these scans is represented by,

\[
\bar{x}_N(n) = \frac{1}{N}\sum_{k=1}^{N} x_k(n)
\] (4.4.2)

where \( k \) represents the scan index and \( N \) represents the total number of scans acquired.

The ALP process utilizes a tapped delay line signal \( x(n) \), of length \( p \), that is fed into a finite impulse response (FIR) filter \( w_n \), also of length \( p \)

\[
x(n) = [x(n - 1), x(n - 2), ..., x(n - p)]^T
\] (4.4.3)

\[
w_n = [w_1, w_2, ..., w_p]^T
\] (4.4.4)

The concept of a \textit{tapped delay line} simply refers to a shift register (with \( p \) memory locations) in which past data samples are shifted out of the register as new samples are shifted in. It essentially “windows” the \( p \) most recent samples acquired in the scan. The prediction or estimate \( d_{est}(n) \) of the desired signal is simply computed by the inner product of these two vectors.

\[
d_{est}(n) = x^T(n) \cdot w_n
\] (4.4.5)
In MR and many other fields, phase is an extremely important figure of merit. Therefore, it is desired that any filter that is used should have a zero phase characteristic. This can be accommodated by implementing a zero phase FIR adaptive filter. This is done by constraining the filter weights to be symmetric with respect to the center tap (i.e., the center weight of the FIR filter). Alternatively, one can alter the tap delay line so that samples symmetric with respect to the center tap are added. This is advantageous because it reduces the number of multiplications computed every iteration. Another feature that we incorporated into the filter is an adaptive bias because it has shown to improve the performance of the ASA technique. It can easily be implemented by inserting a 1 at the end of the zero phase tapped delay line. The adaptive bias enables the algorithm to optimize a single filter weight that is able to adjust the estimate to a more accurate value than that of the conventional predictor. With these two constraints, Eqs. (4.4.21), (4.4.22), and (4.4.23) can be rewritten as the following, respectively:

\[ \chi(n) = \left[ x(n), (x(n + 1) + x(n - 1)), \ldots, (x(n + p) + x(n - q)) , 1 \right]^T \]  \hspace{1cm} (4.4.6)

\[ \omega_n = [\omega_1, \omega_2, \ldots, \omega_q]^T \]  \hspace{1cm} (4.4.7)

\[ d_{est}(n) = \chi^T(n) \cdot \omega_n \]  \hspace{1cm} (4.4.8)

where \( q = (p - 1)/2 \) and \( p \) is assumed to be odd in this case. (The change in notation of the weight vector from \( w_n \) to \( \omega_n \) was performed in order to emphasis that these are vectors of two different lengths.)

In order to use an adaptive filter in an averaging scheme, one must replace the desired response \( d(n) \) with the conventional average of \( N \) scans, \( \bar{x}_N(n) \) given by Eq. (4.4.22). Therefore, the instantaneous error \( e(n) \) which is used to update the weights of the adaptive algorithm is formed by,

\[ e(n) = \bar{x}_N(n) - d_{est}(n) \]  \hspace{1cm} (4.4.9)

The block diagram of the modified ALP is illustrated in figure 4.5.

There are many forms of adaptive filters but the two most widely used and efficient are the least mean squares (LMS) and recursive least squares (RLS) adaptive filters. These filters are advantageous in MR because neither algorithm requires an estimate of the noise or signal statistics. This is desired for MR experiments because these statistics are typically unknown, significantly vary as a function of time, and also differ from experiment to experiment. The main advantage of the RLS algorithm over the LMS algorithm is that it has about an order of magnitude faster convergence time, though, in most cases, the LMS algorithm is known to have better tracking performance [82]. To increase the tracking performance of the RLS algorithm, we utilized the exponentially weighted RLS (EWRLS) algorithm by incorporating an exponentially weighing factor \( \lambda \) into the system. The exponential weighting factor \( \lambda \) effectively
controls the memory of the system and is chosen to be in the range $0 < \lambda < 1$. The EWRLS algorithm attempts to minimize the exponentially weighted sum of squared errors cost function which is given by,

$$\xi(n) = \sum_{i=0}^{n} \lambda^{n-i}|e(i)|^2$$  \hspace{1cm} (4.4.10)

In order to minimize this cost function, the gradient is taken with respect to the weights of the FIR filter and set equal to zero.

$$\nabla \xi(n) = \frac{d\xi(n)}{dw_n} = 2 \sum_{i=0}^{n} \lambda^{n-i}e(i) \frac{de(i)}{dw_n} = -2 \sum_{i=0}^{n} \lambda^{n-i}e(i)x(i) = 0$$  \hspace{1cm} (4.4.11)

Plugging in for the error and using the associative property of vectors yields the following set of linear equations after rearranging,

$$\left[ \sum_{i=0}^{n} \lambda^{n-i}x(i)x^T(i) \right] w_n = \sum_{i=0}^{n} \lambda^{n-i}d(i)x(i)$$  \hspace{1cm} (4.4.12)

This result can be simplified by realizing that the terms in the brackets on the left of equation (4.4.30) is an exponentially weighted deterministic autocorrelation matrix $R_x(n)$ of the input signal and the right hand side of equation (4.4.30) is an exponentially weighted deterministic cross correlation vector $r_{dx}(n)$ of the desired signal and the input signal. By this realization, equation (4.4.30) can be written in matrix form.

$$R_x(n) \cdot w_n = r_{dx}(n) \rightarrow w_n = r_{dx}(n) \cdot R_x^{-1}(n)$$  \hspace{1cm} (4.4.13)

Therefore, the weight vector is found by multiplying the cross correlation vector $r_{dx}(n)$ with the inverse correlation matrix $R_x^{-1}(n)$ Calculation of this inverse is computationally intense so it is not desirable to calculate it every time a new sample is presented to the system. Therefore, one way to reduce the computational time is to realize that $R_x(n)$ and $R_x^{-1}(n)$ can be solved recursively. It can be easily shown that,

$$R_x(n) = \lambda R_x(n-1) + x(n)x^T(n)$$  \hspace{1cm} (4.4.14)
Now that $R_x(n)$ can be solved for in terms of $R_x(n-1)$, there needs to be a way to compute the inverse of this matrix. This is called the matrix inversion lemma. The inverse of the exponentially weighted autocorrelation matrix in equation (4.4.31) can be solved using Woodbury’s identity [82]. Woodbury’s identity states that matrix $A$ of equation (4.4.33) can be inverted with the relation shown in equation (4.4.34). This identity only holds if $A$ and $B$ are positive-definite $p$-by-$p$ matrices, $D$ is a positive-definite $n$-by-$p$ matrix, and $C$ is an $p$-by-$n$ matrix. The relation is easily shown by computing $AA^{-1} = I$, where $I$ is the identity matrix.

\[
A = B^{-1} + CD^{-1}C^T
\]

(4.4.15)\[
A^{-1} = B - BC(D + C^TBC)^{-1}C^T B
\]

(4.4.16)

Note that the following derivation is for real valued data. The transpose operations would be replaced with the hermitian operator for imaginary valued data. Comparing equations (4.4.32) and (4.4.33), it can be realized that,

\[
A = R_x(n)
\]

(4.4.17)\[
B^{-1} = \lambda R_x(n - 1)
\]

(4.4.18)\[
C = x(n)
\]

(4.4.19)\[
D = 1
\]

(4.4.20)

Then, plugging equations (4.4.35) - (4.4.38) into (4.4.32), the exponentially weighted inverse autocorrelation matrix can be computed recursively as follows.

\[
R_x^{-1}(n) = \lambda^{-1} R_x^{-1}(n - 1) + \frac{\lambda^{-2} R_x^{-1}(n - 1)x(n)x^T(n)R_x^{-1}(n - 1)}{1 + \lambda^{-1}x^T(n)R_x^{-1}(n - 1)x(n)}
\]

(4.4.21)

This equation is usually reduced into simpler form,

\[
R_x^{-1}(n) = \lambda^{-1}[R_x^{-1}(n - 1) - g(n)z^T(n)]
\]

(4.4.22)

where,

\[
z(n) = R_x^{-1}(n - 1)x(n)
\]

(4.4.23)
\[ g(n) = \frac{z(n)}{\lambda + x^T(n)z(n)} = R^{-1}_x(n)x(n) \]  

(4.4.24)

The next step is to solve for the weight update. As stated earlier, the weight vector is found by multiplying the cross correlation vector \( r_{dx}(n) \) with the inverse correlation matrix \( R^{-1}_x(n) \). To reduce computation, \( r_{dx}(n) \) is solved recursively in a similar fashion to that of \( R_x(n) \) and is shown below.

\[ r_{dx}(n) = \lambda r_{dx}(n-1) + d(n)x(n) \]  

(4.4.25)

The weight vector is found by computing the product of the autocorrelation matrix \( R^{-1}_x(n) \) obtained in equation (4.4.39) and the recursive cross correlation vector \( r_{dx}(n) \) formed by equation (4.4.43) realizing that \( R^{-1}_x(n-1) \cdot w_{n-1} = r_{dx}(n-1) \).

\[ w_n = R^{-1}_x(n)r_{dx}(n) \]
\[ = w_{n-1} + R^{-1}_x(n)x(n)[d(n) - x^T(n)w_{n-1}] \]
\[ = w_{n-1} + g(n)a(n) \]  

(4.4.26)

where \( g(n) \) was defined previously and \( a(n) \) is the a priori error. The priori error is the error that occurs when using the previous set of filter coefficients \( w_{n-1} \) and is shown below,

\[ a(n) = d(n) - x^T(n)w_{n-1} \]  

(4.4.27)

The EWRLS algorithm just derived is summarized in Appendix C in a procedural manner. It is easy to see that the computation has been reduced significantly from the conventional LS algorithm because of the recursive nature of the autocorrelation and cross correlation functions. Actually, the inverse autocorrelation function only needs to be computed once for initialization. In fact, \( R^{-1}_x(n) \) can be initialized by

\[ R^{-1}_x(n = 0) = \delta I \]  

(4.4.28)

where \( \delta \) is a constant and \( I \) is the identity matrix. The initialization of \( \delta \) depends on the SNR of the signal under observation which is described elsewhere [82] [83]. For our case, a \( \delta = 1 \) was sufficient for successful operation of the algorithm. It is easy to see that the EWRLS algorithm is more computationally
complex than that of the LMS algorithm but, as a result, has better performance. The ASA has the ability to utilize either algorithm but this discussion of results will only involve the RLS algorithm because of its significant performance over the LMS algorithm.

4.4.2 – ASA Performance Enhancement

To evaluate the performance of ASA, one must compare the effects of averaging the estimated signal \( d_{\text{est}}(n) \) to that of the averaging the original unfiltered signal \( x(n) \). Similar to the representation of the unfiltered signal given by Eq. (4.4.19), the estimated signal \( d_{\text{est}}(n) \) can be represented by a desired signal component \( d(n) \) and an error term \( v(n) \),

\[
d_{\text{est}}(n) = d(n) + v(n) \tag{4.4.29}
\]

once again, where \( n \) is the sample index of the scan. Note that the error \( v(n) \) arises from two separate sources: the estimation error of the filter and the error in the conventional average at the time of estimation. Because \( d(n) \) is deterministic for both \( x(n) \) and \( d_{\text{est}}(n) \), it can be dropped from this analysis. Therefore, the increase in performance of ASA can be derived by simply by analyzing the statistical behavior of the averaged random variable components \( u(n) \) from Eq. (4.4.19) and \( v(n) \), from Eq. (11), which have variances of \( \sigma_u^2 \) and \( \sigma_v^2 \) respectively. With this notation used, the noise sequence on the unfiltered average will be denoted by \( u_N(n) \) and the noise sequence on the filtered average will be denoted \( v_M(n) \). Here, \( N \) represents the number of unfiltered scans that are averaged and \( M \) represents the number of filtered scans that are averaged, where \( N > M \). The reason \( M \) scans are averaged and not \( N \) is because, as mentioned previously, the adaptive filter is turned on after the conventional average builds up a reasonable desired response so that a better estimation can be achieved. Therefore, \( L = N - M \) will denote the number of scans that are averaged before the filter is turned on. The noise variance of the unfiltered average and filtered average will be denoted by \( \sigma_{uN}^2 \) and \( \sigma_{vM}^2 \), respectively. By comparing these variance measures, one is able to quantify the performance enhancement that is actually achieved by ASA.

In conventional signal averaging, the noise variance of the average \( \sigma_{uN}^2 \) is determined by the noise variance of the individual scans \( \sigma_u^2 \) divided by the number of scans in the average \( N \).

\[
\sigma_{uN}^2 = \frac{\sigma_u^2}{N} \tag{4.4.30}
\]
This is a commonly known relationship which can easily be derived and is therefore not shown here. In MR, it is fair to assume that the noise observed at the output of the lock-in amplifier to be Gaussian in nature. Even for experiments that don’t exhibit this noise characteristic, the sum of any iid random variables will tend to a Gaussian distribution according to the central-limit theorem [84]. Therefore, the probability distribution of the noise riding on the conventional average can be described by the Gaussian distribution given below,

\[
f(u_N) = G(\mu, \sigma_{uN}^2) = \frac{1}{\sqrt{2\pi\sigma_{uN}^2}} \exp\left\{-\frac{(u_N - \mu)^2}{2\sigma_{uN}^2}\right\}
\]

(4.4.31)

where \(\mu\) is the mean of the random variable and the sample index \(n\) has been dropped.

The derivation for the random variable \(\sigma_v^2\) is not so easy however because the estimation error sequence \(v(n)\) for each filtered scan is unknown and scan dependent. The first step in deriving \(\sigma_v^2\) is to find an expression for the probability distribution of the estimation error \(f(v)\). Once this distribution is known, a noise variance measure can be extracted for each filtered scan. Then, the performance enhancement of ASA can be realized by deriving the variance measure \(\sigma_{vM}^2\) that results from averaging \(M\) scans, each with scan dependent noise variances equal to \(\sigma_v^2\).

The key idea in the derivation of the probability distribution \(f(v)\) is to realize that the error term \(v(n)\) has a random mean that is precisely \(u_N(n)\) with variance about that mean of \(\sigma_z^2\) due to the estimation error of the adaptive filter. This is the case because the filter is estimating the signal based on the noisy conventional average which has a noise sequence equal to \(u_N(n)\). Therefore, if we assume Gaussian statistics (a valid assumption considering the results of the simulation illustrated in the next section), the distribution of the estimation error variable can be written in conditional form as,

\[
f(v|u_N) = G(\mu_N, \sigma_z^2) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left\{-\frac{(v - \mu_N)^2}{2\sigma_z^2}\right\}
\]

(4.4.32)

where once again, \(u_N\) is the random mean and \(\sigma_z^2\) is the estimation error variance due to the adaptive filter. Because the statistics of \(u_N\) are known, one can find the marginal distribution of \(f(v)\) by integrating over the product of the two distributions with respect to the random variable \(u_N\),

\[
f(v) = \int_{-\infty}^{\infty} f(v|u_N)f(u_N)du_N
\]

(4.4.33)
Evaluation of this integral is tedious and will be derived here. In order to solve for this distribution, one must evaluate,

\[ f(v) = \frac{1}{2\pi \sigma_z \sigma_{uN}} \int_{-\infty}^{\infty} \exp \left\{ -\frac{(v - u_N)^2}{2\sigma_z^2} - \frac{(u_N - \mu)^2}{2\sigma_{uN}^2} \right\} du_N \]

The first step in this derivation is to rearrange the exponent so that only the \( u_N \) terms remain in the integral. By doing this, one can integrate over the entire distribution which equates to 1 and allows one to solve for a closed form solution for this probability density function. In order to do this, first factor out the exponent of the exponential \( E \) and group the like \( u_N \) terms to get,

\[ E = -\left\{ u_N^2 \left( \frac{\sigma_z^2 + \sigma_{uN}^2}{2\sigma_z^2 \sigma_{uN}^2} \right) - u_N \left( \frac{\mu \sigma_z^2 + \nu \sigma_{uN}^2}{\sigma_z^2 \sigma_{uN}^2} \right) + \left( \frac{\nu^2}{2\sigma_z^2} + \frac{\mu^2}{2\sigma_{uN}^2} \right) \right\} \]

Now, factor the \( u_N \) terms by completing its square. For example, if the above exponent is represented by \( A^2 - 2AB + C \), we wish to find \( B^2 \) such that \( A^2 - 2AB + C = (A - B)^2 + C - B^2 \). Now equate,

\[ A^2 = u_N^2 \left( \frac{\sigma_z^2 + \sigma_{uN}^2}{2\sigma_z^2 \sigma_{uN}^2} \right), \quad 2AB = u_N \left( \frac{\mu \sigma_z^2 + \nu \sigma_{uN}^2}{\sigma_z^2 \sigma_{uN}^2} \right), \quad C = \left( \frac{\nu^2}{2\sigma_z^2} + \frac{\mu^2}{2\sigma_{uN}^2} \right) \]

Squaring the \( 2AB \) term and dividing by \( A^2 \) results with \( 4B^2 \). Therefore,

\[ B^2 = \frac{(2AB)^2}{4A^2} = \frac{(\sqrt{2}\mu \sigma_z^2 + \sqrt{2}\nu \sigma_{uN}^2)^2}{4\sigma_z^2 \sigma_{uN}^2 (\sigma_z^2 + \sigma_{uN}^2)} \]

Now, the exponent \( E \) can be written as \( (A - B)^2 + C - B^2 \)

\[ E = -\left\{ u_N \sqrt{\frac{\sigma_z^2 + \sigma_{uN}^2}{\sqrt{2}\sigma_z \sigma_{uN}}} - \left( \frac{\sqrt{2}\mu \sigma_z^2 + \sqrt{2}\nu \sigma_{uN}^2}{2\sigma_z \sigma_{uN} \sqrt{2} \sigma_z^2 + \sigma_{uN}^2} \right) \right\}^2 + \left( \frac{\nu^2}{2\sigma_z^2} + \frac{\mu^2}{2\sigma_{uN}^2} \right) - \left( \frac{\sqrt{2}\mu \sigma_z^2 + \sqrt{2}\nu \sigma_{uN}^2}{4\sigma_z^2 \sigma_{uN} (\sigma_z^2 + \sigma_{uN}^2)} \right)^2 \]

With the exponent manipulated in this fashion, the probability distribution of the estimation error can be rewritten as
\[ f(v) = \frac{1}{2\pi \sigma_z \sigma_{uN}} \cdot \exp \left\{ -\left( \frac{v^2}{2\sigma_z^2} + \frac{\mu^2}{2\sigma_{uN}^2} \right) + \left( \frac{\sqrt{2} \mu \sigma_z^2 + \sqrt{2} v \sigma_{uN}^2}{4 \sigma_z^2 \sigma_{uN}^2 (\sigma_z^2 + \sigma_{uN}^2)} \right)^2 \right\} \cdot \int_{-\infty}^{\infty} \exp \left\{ -\left( \frac{\sigma_z^2 + \sigma_{uN}^2}{2\sigma_z^2 \sigma_{uN}^2} \cdot \left( u_N - \frac{\mu \sigma_z^2 + v \sigma_{uN}^2}{\sigma_z^2 + \sigma_{uN}^2} \right) \right)^2 \right\} du_N \]

Rewriting the integral \( I \) to isolate the random variable \( u_N \) yields,

\[ I = \int_{-\infty}^{\infty} \exp \left\{ -\frac{\sigma_z^2 + \sigma_{uN}^2}{2\sigma_z^2 \sigma_{uN}^2} \cdot \left( u_N - \frac{\mu \sigma_z^2 + v \sigma_{uN}^2}{\sigma_z^2 + \sigma_{uN}^2} \right) \right)^2 \right\} du_N \]

Note that this exponent has the form of a Gaussian distribution with the exception of the proper multiplicative constant. It can be rewritten in Gaussian format by

\[ I = \frac{1}{\sigma_z^2 + \sigma_{uN}^2} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi (\sigma_z^2 + \sigma_{uN}^2)}} \cdot \exp \left\{ -\frac{\left( u_N - \frac{\mu \sigma_z^2 + v \sigma_{uN}^2}{\sigma_z^2 + \sigma_{uN}^2} \right)^2}{2 \cdot \left( \frac{\sigma_z^2 + \sigma_{uN}^2}{\sigma_z^2 + \sigma_{uN}^2} \right)^2} \right\} du_N \]

By writing the integral in this form, the expression reduces to a constant because the integration takes place over the entire density function which equates to 1. Therefore, \( f(v) \) can be written as,

\[ f(v) = \frac{1}{2\pi \sigma_z \sigma_{uN}} \cdot \exp \left\{ -\left( \frac{v^2}{2\sigma_z^2} + \frac{\mu^2}{2\sigma_{uN}^2} \right) + \left( \frac{\sqrt{2} \mu \sigma_z^2 + \sqrt{2} v \sigma_{uN}^2}{4 \sigma_z^2 \sigma_{uN}^2 (\sigma_z^2 + \sigma_{uN}^2)} \right)^2 \right\} \cdot I \]

\[
= \frac{1}{\sqrt{2\pi (\sigma_z^2 + \sigma_{uN}^2)}} \cdot \exp \left\{ -\left( \frac{v^2}{2\sigma_z^2} + \frac{\mu^2}{2\sigma_{uN}^2} \right) + \left( \frac{\sqrt{2} \mu \sigma_z^2 + \sqrt{2} v \sigma_{uN}^2}{4 \sigma_z^2 \sigma_{uN}^2 (\sigma_z^2 + \sigma_{uN}^2)} \right)^2 \right\}
\]

With quite a bit of algebraic manipulation, the exponent can be rearranged to yield a closed form Gaussian probability density function for \( f(v) \) represented by

\[ f(v) = \frac{1}{\sqrt{2\pi (\sigma_z^2 + \sigma_{uN}^2)}} \exp \left\{ -\frac{(v - \mu)^2}{2(\sigma_z^2 + \sigma_{uN}^2)} \right\} = G(\mu, \sigma_z^2 + \sigma_{uN}^2) = G(\mu, \sigma_v^2) \tag{4.4.34} \]

103
Here, the distribution has mean equal to $\mu$ and scan dependent variance equal to $\sigma_z^2 + \sigma_{uN}^2$ where $N$ indicates the number of scans. As one might initially guess, the error variable $\nu(n)$ has a mean equal to $\mu$ (the mean of $u_N(n)$) and a variance equal to the sum of the adaptive filter error variance $\sigma_z^2$ and the noise variance on the conventional average $\sigma_{uN}^2$. From these statistics, it is clear that this random variable is dependent upon the scan number. The mean of this random variable is constant; however, the variance depends on the filtered scan number $k$ and thus, can be written as,

$$\sigma_v^2 = \sigma_z^2 + \sigma_{uN}^2 = \sigma_z^2 + \frac{\sigma_u^2}{N} = \sigma_z^2 + \frac{\sigma_u^2}{k + L - 1} \tag{4.4.35}$$

The second moment (which will be used later) can be written as,

$$E[\nu^2] = \sigma_v^2 - \mu^2 = \sigma_z^2 + \frac{\sigma_u^2}{k + L - 1} - \mu^2 \tag{4.4.36}$$

Because lock-in amplification is utilized, the mean is equal to zero and therefore, $E[\nu^2] = \sigma_v^2$. Note that the variance (and second moment) are dependent upon the noise variance of an individual scan $\sigma_u^2$ divided filtered scan number $k$ and offset by $L - 1$. The reason for the offset is to account for the fact that the filter is turned on after a few scans have already been averaged. For example, assume that the filter is turned on after 15 scans have been averaged (e.g.: $L = 15$). On the first filtered pass when $k = 1$, the error variance will be $\sigma_z^2 + \sigma_u^2/15$ and on the eighty fifth filtered pass, the error variance will be $\sigma_z^2 + \sigma_u^2/85$. It should now be intuitive why the filter is turned on after 15 to 25 scans into the averaging process. The second term in Eq. (4.4.53) should be reduced as much as possible because it is precisely this term that slightly biases the filtered average. The one parameter under our control to reduce this term is the turn-on time $L$. If we have chosen $L$ such that $\sigma_u^2/(k + L - 1) < \sigma_z^2$, then our filtered estimate would be unbiased for all future scans. In most of the signals we observed in MR, a value of $15 < L < 25$ typically reduced the noise enough to avoid this bias. Even if the filter is applied before 15 scans are acquired, the bias that is introduced will die away as more scans are averaged. As the number of scans increases, $\sigma_u^2/(k + L - 1) \ll \sigma_z^2$ which implies that the estimation error will be solely due to $\sigma_z^2$. Because one can tune the adaptive filter so that $\sigma_z^2 \ll \sigma_u^2$, averaging a series of scans with variance $\sigma_z^2$ will provide a more dramatic attenuation of noise than averaging a series of scans with variance $\sigma_u^2$. However, this simple idea assumes that the estimation error variance is constant for all scans whereas in reality, it is not. As indicated by Eq. (4.4.53), it is dependent upon scan number. Therefore, one must investigate what happens when this realization is taken into account.
Now that the distribution of $v(n)$ is known, the final step is to investigate what happens when these filtered scans are averaged. First, let $v_M(n)$ represent the noise on the average of the $M$ filtered scans and be defined by,

$$v_M(n) = \frac{1}{M} \sum_{k=1}^{M} v_k(n) \quad (4.4.37)$$

where $k$ represents the filtered scan index and $M$ represents the total number of scans in the filtered average. The variance of this random variable can be derived from the following common variance relation,

$$\sigma^2 = E[(v_M(n) - E[v_M(n)])^2] \quad (4.4.38)$$

where $E[v_M(n)]$ can be easily shown to be equal to $\mu$ which in our case equals zero. Assuming that the noise on each scan is iid, Eq. (4.456) can be expanded with the use of Eqs. (4.4.53) and (4.4.55) to yield,

$$\sigma^2 = \frac{1}{M^2} \sum_{k=1}^{M} \sum_{i=1}^{M} v_k(n)v_l(n) = \frac{1}{M^2} \left( \sum_{k=1}^{M} \sigma^2_z + \sum_{k=1}^{M} \frac{\sigma^2_u}{k + L - 1} \right) \quad (4.4.39)$$

Making the substitution, $c = k + L - 1$ and realizing that $\sigma^2_z$ is approximately constant for all scans because the filter is initialized once, Eq. (4.4.57) can be written as

$$\sigma^2 = \frac{\sigma^2_z}{M} + \frac{\sigma^2_u}{M^2} \sum_{c=L}^{N-1} \frac{1}{c} \quad (4.4.40)$$

Finally, using the harmonic series approximation,

$$\sum_{c=1}^{N} \frac{1}{c} \approx \ln(N) + \gamma \quad (4.4.41)$$

where $\gamma = 0.5772156$ is the Euler-Mascheroni constant, the approximated variance of the filtered average is given by,
\[
\sigma_{vM}^2 \approx \frac{1}{M} \left[ \sigma_z^2 + \frac{\sigma_u^2}{M} \ln \left( \frac{N - 1}{L - 1} \right) \right]
\] (4.4.42)

The performance enhancement achieved by that of ASA is simply given by the ratio of the variance of the conventional average over the variance of the filtered average,

\[
\frac{\sigma_{uN}^2}{\sigma_{vM}^2} \approx \frac{M}{N} \left[ \frac{\sigma_z^2}{\sigma_u^2} + \frac{1}{M} \ln \left( \frac{N - 1}{L - 1} \right) \right]^{-1}
\] (4.4.43)

When many scans are averaged, (ie: when \( N >> 1 \)), then \( M/N \approx 1 \), and the ratio becomes,

\[
\frac{\sigma_{uN}^2}{\sigma_{vM}^2} \approx \frac{\sigma_u^2}{\sigma_z^2}
\] (4.4.44)

This expression illustrates that the improvement gained by ASA is approximately equal to the ratio between the noise variance on an individual trace \( \sigma_u^2 \) and the estimation error variance of the filter \( \sigma_z^2 \). This implies that the gain achieved can be quite significant because the estimation error variance of the adaptive filter \( \sigma_z^2 \) is typically relatively small. Note that this relationship also suggests that the performance enhancement is dependent upon the nature of noise. Typically, the larger \( \sigma_u^2 \) the worse the estimation will be and hence, the larger \( \sigma_z^2 \) will be. The estimation of the desired signal however, will always be better than or equal to that of the noisy input because the filter is optimized to minimize the sum of squared errors. Therefore, the level of the noise in the filtered average will always be lower than that of the original average, despite being averaged with fewer scans.

**4.4.3 – ASA Simulation and Results**

In this study, we applied the ASA on simulated continuous wave magnetic resonance line shapes which have varying line widths and amplitudes. These spectra were simulated by taking the derivatives of Gaussian line shapes in which noise was added. (Derivatives were taken because continuous wave MR spectra are typically reported as the slope of the absorption spectrum when magnetic field modulation is utilized.) The desired spectrum in this case is composed of 6 lines, each with amplitude of either 0.1 or 1 and variances of either 40 samples, 200 samples, or 800 samples. The line widths and amplitudes were varied to test the various combinations of lines typically observed in MR spectra. The noise sequences that were generated were \( 1^{st} \) order autoregressive (AR) processes (constant coefficient equal to 0.25) with white Gaussian noise \( (\sigma_w^2 = 0.1) \) driving the process, i.e. \( u(n) = 0.25u(n-1) + w(n) \). AR noise was
used in order to mimic the low pass filtering stage at the back end of the lock-in amplifier. With 40 data points representing the narrowest peak-to-peak line width, we fixed the filter length to 25 taps. Filters with lengths longer than that of 40 taps would slightly distort this narrow line shape while filters smaller than 10 taps were not as effective in removing the noise. A good rule of thumb would be to keep filter length 1/3 to 2/3 times the minimum feature width expected. Also, in order for the adaptive filter to be effective, the sampling rate of the spectrometer should be set to acquire at least 30 samples per minimum line width expected.

Our first simulations involved demonstrating the effects of varying the turn-on time \( L \) and exponential weighting factor \( \lambda \). Figure 4.6 (a) illustrates the results when the exponential weighting factor \( \lambda \) was varied and the turn-on time \( L \) was held constant at 15 scans. For these simulations, we simply subtracted the desired signal from the unfiltered and filtered averages after each scan and recorded a variance measurement. The blue trace illustrates the reduction in noise obtained from conventional signal averaging and the red traces illustrate the reduction in noise obtained from ASA for different values of \( \lambda \) \((0.55 < \lambda < 1\), at increments of 0.05).

![Figure 4.6](a) This figure illustrates the reduction in noise power for both the unfiltered and filtered averages as a function of scan for different values of the exponential weighting factor. The range of \( \lambda \) is \(0.55 \leq \lambda \leq 0.95\) with 0.05 increments. (b)This figure illustrates the reduction in noise power for both the unfiltered and filtered averages as a function of scan for different values of the turn on time \( L \). The turn on time was varied from \(2 \leq L \leq 40\) at increments of 4 scans.

From the figure, it is clear that as \( \lambda \) approaches 1, the more effective the filter is at attenuating the noise. Note however that when \( \lambda = 1 \), signal attenuation was observed for the two narrower line shapes. As a result, we typically do not recommend setting \( \lambda = 1 \) when narrow line shapes are present. Figure 4.6 (b) illustrates the results when the exponential weighting factor was fixed \( \lambda = 0.9 \) (a conservative value to ensure signal attenuation wouldn’t be encountered for the narrow line shapes) and the turn-on time \( L \) was varied \((2 < L < 40\), at increments of 4 scans). Note that when the filter is turned on earlier than the
recommended minimum time of 15 scans, the small bias discussed earlier is observed when compared to a simulation that followed this recommendation, e.g. \( L = 18, 22, 24, 28, 32, 36, \) and 40. Note however that this bias is small and will die away as more scans are added to the filtered average.

Figure 4.7 (a-d) illustrates representative traces taken from the simulations just described. In these data sets, the turn-on time \( L \) was set to 15 scans, the exponential weighting factor was set to 0.9, and the filter length remained at 25 taps. Figure 4.7 (a) illustrates the desired signal and a representative noise sequence for this experiment. Figure 4.7 (b) illustrates the 100\(^{th}\) individual unfiltered scan compared to its filtered counterpart. The desired signal used when filtering this scan consisted of the average of the first 100 individual unfiltered scans.

![Figure 4.7](image)

Figure 4.7: (a) The top plot illustrates the desired simulated signal and the bottom plot illustrates a representative noise sequence \( u(n) \) that was added to create \( x(n) \). The three smaller signals to the left are enlarged in a few of the plots in the figure to better visualize results of the simulation. (b) The bottom plot illustrates the 100\(^{th}\) individual unfiltered scan and the top figure illustrates its filtered version using the average of 100 unfiltered scans as the desired signal in the ALP. Note that only two of six line shapes are apparent in the individual unfiltered trace while five of the six line shapes are revealed in an individual filtered trace. (c) The top plot illustrates the average of 85 filtered scans and the bottom figure illustrates the average of 100 unfiltered scans. Note that all 6 line shapes are almost completely resolved in the filtered average. (d) The top plot illustrates the average of 985 filtered scans and the bottom figure illustrates the average of 1000 unfiltered scans. Note that even now, the conventional average still does not match the SNR achieved when averaging 85 filtered scans.
Note that the significant reduction in noise allows one to clearly observe five of the six line shapes in an individual trace. It should be pointed out that even though the adaptive filter acts as a low pass filter, the predicted estimate also has relatively high frequency content associated with it. This implies that a smoothing filter (with a narrow window) can be utilized on top of the ASA for additional noise attenuation. (We did not incorporate this feature in this study as it is intended solely to illustrate the performance of the ASA.) Figure 4.7 (c) illustrates the average of 100 unfiltered scans and the average of 85 filtered scans. The reduction in noise power was measured to be 14 times lower (an improvement in SNR of about 3.74) in the filtered average than in the unfiltered average. This large reduction in noise in the filtered average allows for the observation of the small narrow signal on the far left of the spectrum. For this signal to reveal itself in the conventional average, many hundred additional scans were needed. Figure 4.7 (d) illustrates the average of 1000 unfiltered scans and the average of 985 filtered scans. Note that reduction in noise power in the filtered average at after the 1000th scan was about 10 times lower than that of the unfiltered average. Only now does the small narrow signal on the far left of the spectrum make an appearance in the unfiltered average.

Figure 4.8 (a) illustrates the histogram of the 85th sequence of prediction errors $v(n)$ with a superimposed Gaussian fit. This close fit justifies the assumption made earlier in section III about this error term as having a Gaussian distribution. Figure 4.8 (b) illustrates the histogram of the average (985 scans) prediction errors $\langle v \rangle_{985}(n)$ with a superimposed Gaussian fit. This close correspondence is expected as suggested by the central-limit theorem.
We tracked the noise power as a function of scan number for the unfiltered \((10 \cdot \log(\sigma_{uN}^2))\) and filtered averages \((10 \cdot \log(\sigma_{mN}^2))\) as illustrated in the top plot of figure 4.9. The bottom plot of figure 4.9 illustrates the factor at which these two noise attenuation measures differ as a function of scan number. Note that a factor greater than 10 was consistently observed throughout the averaging process. This is consistent with the data that we collected when first performing ASA on continuous wave electrically detected MR reported in 2008 [62].

ASA has been demonstrated to be an extremely useful and efficient tool for the averaging MR scans. It is very effective in its estimation abilities even when using a noisy desired signal. It is capable of reducing the noise variance by a factor of at least 10 in a single trace and the average of the filtered scans was shown to converge by a similar factor, much faster than that of conventional signal averaging. With such great reduction in noise, the ASA greatly reduces the time required for averaging. This tool not only has great potential in MR measurements, but it would also be beneficial in any field where noisy data sets of a repeatable measurement are averaged to acquire high sensitivity measurements with high SNR. Some of these applications include, but not limited to, are medical imaging, electrocardiography, or electroencephalography.

Figure 4.9: The top plot illustrates the reduction in noise power for both unfiltered and filtered averages as a function of scan. The bottom plot illustrates the noise reduction factor between these two variance levels as a function of scan. Note that the filtered average maintained a noise power level that was consistently 10 times lower than that of the unfiltered average. This would correspond to a SNR improvement by a factor of more than 3.
4.5 – Higher Order Harmonic Detection

We have also developed a way to enhance the resolution of the EDMR spectrum by utilizing a conventional EPR phenomenon commonly referred to as passage effects [45] [85] [86] [87]. As mentioned earlier, magnetic field modulation is utilized in both EPR and EDMR to remove much of the noise in the measurement, however, it introduces artifacts that are observed in the received signal. These artifacts are not necessarily harmful to the EDMR measurement, but rather can be useful.

As the quasi static $B_0$ magnetic field is swept through an absorption line shape, the magnetic field modulation (applied in the same direction) brings the system in and out of resonance at a frequency equal to that of the modulation reference signal. An example is simulated in figure 4.10. (In this simulation, relaxation processes were not taken into account.) The left plot illustrates the field passage and the right plot illustrates time passage through resonance. In the left figure, the red line is representative of a Gaussian absorption line shape and the superimposed oscillating blue line is that of the magnetic field. The resonant line shape $x$ is a function of magnetic field $B$ as is defined by a Gaussian,

$$x(B) = C \exp \left\{ -\frac{(B - B_r)^2}{2\sigma^2} \right\}$$  (4.5.1)

where $B_r$ is the resonant field, $C$ is an arbitrary constant (equal to 1 in this case), and $\sigma^2$ is the variance of the line shape (ie: line width of the signal). The magnetic field $B$ of the simulation was defined by,

$$B(t) = \left[ CF - \frac{SW}{2} \right] + \frac{SW}{ST} t + B_m \cos(\omega_m t)$$  (4.5.2)

where $t$ is the time index, $CF$ is the center field, $SW$ is the sweep width, $ST$ is the sweep time, $B_m$ is the modulation amplitude, and $\omega_m$ is the modulation frequency. The time passage signal $x(t)$ can be obtained by sampling the Gaussian spectrum along $B(t)$,

$$x(t) = x(B(t)) = \exp \left\{ -\frac{(B(t) - B_r)^2}{2\sigma^2} \right\}$$  (4.5.3)

The matlab function $x(t) = interp1(B, x(B), B(t))$ was used to extract equation (4.5.3) by evaluating $x(B)$ on a new axis $B(t)$, using linear interpolation. As illustrated in the figure, the two highlighted periods of the modulation waveform illustrate why harmonics are observed in these measurements. The first highlighted section (magenta) crosses the resonant center field (green line) two times over one
modulation cycle whereas the second section (orange) crosses the center resonant field three times over a given modulation cycle. As illustrated in the time passage plot to the right, these “resonant center field crossings” introduce multiple phase reversals over one period of the modulation cycle. These phase reversals introduce new frequencies in the received signal.

The simulation in figure 4.10 utilized a magnetic field modulation of 2 Hz. The purpose of this was solely to visually illustrate the effects that magnetic field modulation has on the passage of a Gaussian line shape. This simulation was run again using a modulation frequency of 1kHz. Figure 4.11 illustrates these results. The top two plots illustrate the FFT of equation (4.5.3) for the (left) in-phase/real component and the (right) quadrature/imaginary components for the entire scan. The bottom figure illustrates the (1st column) time domain signals for the 1st, 2nd, and 3rd harmonics, (2nd column) in-phase demodulated 1st, 2nd, and 3rd harmonics as a function of field, and (3rd column) out-of-phase demodulated 1st, 2nd, and 3rd harmonics as a function of field. (Note that the signals in the first column were acquired by first demodulating that specific harmonic, low pass filtering to remove the higher harmonics, and then modulated back to its respective frequency band. This enables one to view the time domain representation of that signal which is solely due to that particular frequency band.) As illustrated in the figures, the resultant signal will consist of not only the fundamental frequency, but will also contain higher order harmonics at integer values of the fundamental. The total signal can be represented the Fourier series approximation [45] [88],

\[
x(t) = a_0(B) + \sum_{n=1}^{\infty} [a_n(B) \cos(n\omega_n t) + b_n(B) \sin(n\omega_n t)]
\]

\[
= a_0(B) + \sum_{n=1}^{\infty} [a_n(B) \cos(n\omega_m t) + b_n(B)\sin(n\omega_m t)] 
\]

(4.5.4)

where \(\omega_m\) is the modulation frequency, and the Fourier coefficients are defined by

\[
a_0 = 0
\]

(4.5.5)

\[
a_n(B) = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cos(n\omega_m t) dt
\]

(4.5.6)

\[
b_n(B) = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \sin(n\omega_m t) dt
\]

(4.5.7)

where \(a_n(B)\) and \(b_n(B)\) are the Fourier coefficients are a function of magnetic field.
Figure 4.10: The effect modulation has on the resultant signal when passing through resonance. (top) Passage through field and (bottom) passage through time. The frequency simulated is extremely low to illustrate the effects of the modulation over the Gaussian line shape.
Figure 4.11: Simulation of the effects of magnetic field modulation over the Gaussian line shape illustrated in figure 4.10.

These coefficients will depend upon the resonant response of the device and the settings of the spectrometer. Some have attempted to solve these coefficients with simplified assumptions of the line shapes [45]. Note the simulation illustrated in figure 4.11 is just a particular example of the type of line shapes that can be extracted from the higher order harmonics. The harmonics illustrated in the FFT plots in figure 4.11 represent the average frequency response over the entire scan. In passage through time, these harmonics will vary in amplitude as a function of time. Figure 4.12 illustrates the FFT of the in-phase and quadrature components performed over a “chunk” (100k samples) of the data gathered from the modulated signal centered over the absorption signal. As illustrated, there is no in-phase (blue) fundamental component at 1kHz, only the quadrature component. This is additional evidence that there is
a zero crossing precisely at the resonance field when utilizing 1st harmonic magnetic field demodulation. In fact, there are zero crossings for all the odd harmonics at the resonant field as illustrated in both figure 4.11 and 4.12. At the same time, all of the even components are all maximized at the resonant field as illustrated in the two figures. It is worth noting that the amplitude of the harmonics is reduced as the frequency is increased. This simulation was performed without the introduction of noise which makes all the harmonics that are present visible in the simulation. In a real experiment, noise will ultimately determine the highest order harmonic that can be demodulated. In the bipolar junction transistor used in this study, we were actually capable of demodulating up to the 5th harmonic. These higher order harmonics have actually been useful in many areas. Higher order harmonics have been used for reconstruction of the first derivative EPR spectra to improve SNR [89]. Also, higher order harmonic detection can sometimes provide valuable information when certain conditions are met. More specifically, detection of various “passage” conditions are possible. The conditions for these effects are discussed in chapter 6.

Figure 4.12: FFT of the modulated signal over the center field for (blue) in phase and (red) quadrature components. Note that there is very low power signal at the fundamental frequency and a maximum for the second harmonic for the in phase component. This also demonstrates why a zero crossing is observed precisely at resonance when demodulating at the fundamental frequency when perfectly in phase.
CHAPTER 5 - EXPERIMENTAL

5.1 – Devices Used

Our initial study of the zero-field phenomenon utilized devices made of silicon carbide (SiC). The reasoning for this, simply put, is because SiC was the first material system in which we observed this phenomenon. SiC is a relatively new semiconductor with potential uses in high power electronics and operation in high temperature environments. After many years of research, SiC has finally been commercialized into a power MOSFET package which can now be purchased from digikey for $33. The main advantage that SiC has over its conventional silicon (Si) counterpart is that it has a larger band gap energy. The band gap energy of SiC (the 4H polytype is 3.2 eV) is almost three times as large as the bandgap of Si (1.12 eV). Of the three main polytypes of SiC (3C, 4H, and 6H) studied for use as a semiconductor, the most widely used (and the one used in this study) is the 4H SiC polytype. In the 4H polytype, each silicon atom is covalently bonded to four carbon atoms and each carbon atom is covalently bonded to four silicon atoms. Consequently, there are an equal number of carbon and silicon atoms in the 4H SiC polytype unit cell. This polytype is preferred over the 150 other SiC based polytypes for power electronics because of its higher carrier mobility and its lower dopant ionization energy [90].

It is important to note that SiC technology is still in its infancy. As a result, there are still many problems associated with it. For example, hollow-core dislocations in the c-axis direction may be present in some SiC crystals [90]. These defects, known as micropipes, can severely affect the performance of a device. Also, there are many problems associated with the semiconductor/oxide interface region of SiC/SiO₂ based MOSFETs [90]. A poor interface results from the mismatched hexagonal 4H SiC polytype with the SiO₂ oxide. A poor interface results from the presence of a high number defects at the interface which reduces the channel mobility and alters the threshold voltage of a device. Because the SiC system is still plagued with many defects, the EDMR signals we typically observe are large and have high signal to noise ratios (SNR). Therefore, these imperfections in the SiC system make it the perfect candidate for understanding and developing the connection between this newly discovered low field defect identification technique and conventional EDMR.

In our preliminary work [18], low field EDMR is applied to the base/collector diode of a 4H SiC n-p-n bipolar junction transistor (BJT) that has an area of 500 x 500 μm² fabricated by Cree. This device was chosen specifically because it has pn homojunctions whose theoretical qualitative recombination response is well understood as a function of bias voltage. Because we show that recombination is a spin dependent process, this understanding will allow us to help characterize this low field spin dependent phenomenon. The BJT is characterized as having a beta value of about 40. These devices consist of a 40 μm thick n-type collector doped at 1.4x10¹⁵ cm⁻³, a 1 μm thick p-type base doped at 4x10¹⁷ cm⁻³, and a 2
μm thick n-type emitter doped at 3x10^{19} \text{ cm}^{-3} with the base implant to emitter distance of 5 μm. The dopants in this device are aluminum and nitrogen. EDMR results on this device at X-band frequencies have been well characterized and documented by our group [8]. We also investigated the zero- and low-field spin dependent transport of two different metal oxide semiconductor field effect transistors (MOSFETs). The first MOSFET used in this study was an older generation 4H SiC n-channel MOSFETs fabricated by Cree in 2003. This device was characterized as having a mobility of 5.7 cm²/V·s, a thermal ONO anneal, gate area of 200 x 200 μm², and the source and drain regions were doped nitrogen. The second MOSFET used in the study was a more recent device also fabricated by Cree in 2007. This device was a lateral 4H SiC n-channel MOSFETs characterized as having a mobility of 0.5 cm²/V·s, did not have an annealing treatment, had gate areas of 100 x 5 μm², and the source and drain regions were doped with nitrogen. These MOSFETs were utilized in this study because the dominating defect, a silicon vacancy, was well characterized in both of these devices at X-band [3]. The fourth device that was utilized in this study was a 50Å amorphous SiC capacitor that was fabricated by Intel for interlayer dielectrics. This device was used to demonstrate that the zero-field response not only could be used to detect SDR, but also SDT through a dielectric. The final device that was used in this study was a Si power pMOSFET with 48 nm SiO₂ dielectric thickness from Austria Microsystems. These devices had gate areas of 1x10⁶ μm² and were fabricated using a 20 minute wet oxidation at 900°C. The oxidation steps were followed by a N₂ anneal at 925°C. This device was used to show that NBTI mechanisms can be studied with zero-field SDR in silicon devices.

These five devices are illustrated in figure 5.1. Each was mounted on a custom designed PCB which has exposed leads that are wire bonded to the pads of the device. The elongated PCB allows for the device to be biased when placed inside either of the high field or low-field spectrometers while simultaneously monitoring the current.

Figure 5.1: Photos taken of the devices used in this study: 4H SiC BJT, amorphous SiC:H capacitor, older generation 4H SiC MOSFET, newer generation 4H SiC MOSFET, and Si MOSFET.
5.2 – Low-/Zero-Field Spectrometer Development

The measurements made in this study were performed on custom built low-field and high-field EDMR spectrometers. The high field EDMR spectrometer (X-band) is composed of a Varian power supply and a 4 in. diameter high-field electromagnet which generates the sweeping $B_0$ magnetic field. The power supply was controlled by a PI controller integrated in a LakeShore 475 DSP Gaussmeter with connected Hall probe. Because the power supply (built into the magnet) requires a floating analog control input, an optoisolation circuit was built for proper functioning. We utilized the bipolar optoisolation circuit obtained from Analog Devices application note HCNR200 as illustrated in figure 5.2.

![Figure 5.2: Bipolar optoisolation circuit obtained from Analog Devices application note HCNR200 used to isolate grounds of the controller and the Varian magnet.](image)

This circuit was used to isolate grounds of the controller and the magnet. In the figure, VIN represents the input control signal from the Lakeshore 475 DSP Gaussmeter and the VOUT signal is connected to the Varian power supply control input. Note that the ground of the left circuit is connected to Earth ground through the Gaussmeter and the ground of the right circuit is left floating. Labview computer software was designed to trigger the Gaussmeter for repetitive scans so that signal averaging could be utilized. The $B_1$ oscillating field was generated by a Micro-Now X-band microwave source which was connected to a waveguide and TE$_{102}$ microwave cavity. Magnetic field modulation was implemented by sending a digitally generated sinusoidal voltage through a SONY audio amplifier which pumped current through a set of Helmholtz coils mounted on the outer walls of the microwave cavity. The spin dependent currents were first amplified and then conditioned using a Stanford Instruments SR-570 current preamplifier before being quantized by a 16-bit NI PCI-6259 M series DAQ card at 500 kS/s. Once digitized, the modulated current was then demodulated in software via a custom designed, virtual lock-in amplifier. The X-band spectrometer is illustrated in figure 5.3.
The low- / zero- field EDMR measurements were made on a custom built spectrometer which utilizes a Stanford Instruments SG382 RF signal generator with an in-house built tuned resonant circuit in which a coil generates the oscillating $B_1$ magnetic field. The $B_0$ magnetic field was generated through a custom built Helmholtz electromagnet designed with 3D CAD tools and machined using a water jet and 3d printer. A PI controller, implemented in Labview software, controls a programmable Kepco BOP 50-2M power supply which powers the magnet and utilizes a LakeShore 450 Gaussmeter with Hall probe as the feedback sensor. As is the case for the X-band spectrometer, we use a Stanford Research Systems SR570 current preamplifier for amplifying and conditioning the device currents while sampling the analog output with a 16-bit NI PCI-6259 M series DAQ card. We utilize magnetic field modulation (audio frequencies) via Helmholtz coils which modulate the $B_0$ magnetic field. The modulated spin dependent currents are demodulated using a custom designed virtual lock-in amplifier written in Labview. This spectrometer is illustrated in figure 5.4.

The software for each spectrometer was written in Labview 2009. The software controls the repetitive sweeping of the magnet field, signal averaging, lock-in amplification, and any additional signal.
processing that is desired by the user. All measurements made at low and high field were made at room temperature unless indicated otherwise. Much time has been devoted to designing the low-field EDMR spectrometer over the past year. The construction of this new spectrometer has allowed us to perform low-field resonance and zero-field transport experiments. Because the developments of these advancements in spectrometer design were a significant portion of the work performed, a few paragraphs will now be dedicated to this work.

5.2.1 – Magnetic Field Control

In continuous wave EDMR measurements, one of the most valuable pieces of information that can be extracted from a spectrum are the electron nuclear hyperfine interactions. In the systems under study, the hyperfine interactions are generally observed as side peaks around a dominating center line in defects associated with low abundance nuclear moments. If these hyperfine side peaks are small and narrow, they will get averaged out if the magnetic field control is not properly designed. Most spectrometers use a proportional integrator (PI) controller with the use of a Hall probe to sweep the magnetic field. (This controller is essentially a subset of PID controllers where the D is used to indicate derivative.) The Hall probe and Gaussmeter act as the sensor in a closed loop control system where P and I are constants that make the overall transfer function of the system stable. For an analog controller, the electrical components are prone to drift due to changes in temperature which will result in the shifting of the field as a function of scan. As a result, blurring or complete removal of narrow hyperfine side peaks may result when signal averaging is utilized. Although the drift is compensated for with temperature control in the Hall probe, over the period of a few dozen scans, drifts in our spectrometers (using an off the shelf MicroNow 8320A field controller) have been measured up to ±1.5G. Although these drifts do not typically lead to significant errors in average values, the shifts are completely intolerable if hyperfine lines are relatively weak and are much less than 1 Gauss in width. The first strategy to overcome this problem is to switch to a digital controller. Digital controllers are less prone to drift because the field control is performed numerically and not dependent upon temperature dependent electronics. (There will still be a slight error in measurement of the Hall probe due to its analog nature. However, it is usually << 1G for slight changes in temperature. We also ensured that the Gaussmeter purchased also have a numerical temperature compensation feature which would account for this slight error.) We designed a PI controller in Labview that communicates with a BOP 50-2M power supply to power the custom built Helmholtz magnet. PI controller that was implemented for our low field spectrometer is illustrated in figure 5.5.
Figure 5.5: PI controller developed in Labview to control the sweeping $B_0$ magnetic field.

As illustrated in the figure, negative feedback is utilized in PID controllers. Negative feedback essentially provides the controller with a measure of the error in magnetic field as a function of time which quantifies its performance. At the beginning of the experiment, the computer initializes an array of samples that represent a linearly increasing magnetic field as a function of time. This parameters used to create this array are center field (G), sweep width (Gauss), sweep time (sec), and the desired number of samples to represent the final data set. This array is represented by the desired field variable in the figure. When the experiment is started, the computer will retrieve one sample of measured magnetic field every iteration of the software loop. The controller first finds the error in field by taking the difference of the desired field and the measured field. The controller will then compute a numeric voltage by multiplying the error by a constant $P$ and the sum of all errors by constant $I$. The proportional constant $P$ represents the weight for the most present sample and the integration constant $I$ represents the weight for integrated errors. (Note that the integration component of the controller is absolutely necessary because it is precisely this component that compensate for any errors that prevent the current sweep from being perfectly linear. One particular example would be the deviation from Ohm’s Law $V = IR$ when the resistance of a wire changes because of the changes in temperature it experiences from the increases current being sent through it. For this reason alone, a simple ramp generator should never be used to control a magnetic field. If one does, erroneous results will be obtained in hyperfine parameters, relative field measures, and or line shape.) In analog control theory, the output voltage of a controller would be,

$$V(t) = Pe(t) + I \int_{t=0}^{t} e(i) dt$$  \hspace{1cm} (5.4.1)$$

and in the digital domain, the output voltage is replaced by a numerical representation of the voltage needed to compensate for the measured error,
\[ V(n) = Pe(n) + \sum_{i=0}^{n} e(i)T_s \]  

(5.4.2)

where \( n \) represents the digital sample number and \( T_s \) is the sample period. In the digital domain, the smallest amount of time between consecutive samples is the sampling period. Therefore, the differential time can be replaced by the sampling period constant. As a result, the digital version of a PI controller can easily be implemented in a virtual environment. Because the spectrometer (plant) typically has an unknown transfer function, the values of \( P \) and \( I \) cannot be directly calculated for stable operation. However, these constants are usually “tuned” for optimal operation. The final addition to the controller was the limiter block which simply limits the computed output voltage. This is necessary to limit voltage (and current) that is applied to the electromagnet. PID control theory is a topic within itself so no further discussion will be presented here but can be found elsewhere [91]. Utilizing a PI controller to control the field in a continuous wave MR spectrometer allows one to generate a perfectly linear field. However, in many cases, inconsistencies in latency arise from the communication of the computer, Gaussmeter, and programmable power supply due to the varying buffering and queuing processes that occur within the hardware component interfaces. As a result, the magnetic field sweep will begin at slightly different times relative to the instant at which the user (or software) initiated the scanning sequence. This will result in the acquisition of data sets that appear to be shifted in magnetic field which, however, are actually shifted in time. If these signals are averaged, smoothing or removal of hyperfine interactions is inevitable. Even with the new temperature compensated LakeShore 475 DSP Gaussmeters (rated with stability of 0.1 Gauss) functioning with our Labview software, we observed shifts in the spectrum up to 1G due to these timing inconsistencies. Obviously, these errors can be reduced if one realizes that the error is minimized if the magnetic field is swept very slowly. However, slowly sweeping the magnetic field is not desirable because not only do the scans take much longer to acquire, but the acquired date is more prone to DC drift due to drafts, changes in temperature, and voltage fluctuations in the hardware. These DC drifts are more harmful than problems caused by high frequency noise because they are harder to average out. Because the magnetic field samples are measured simultaneously with the measured EMDR samples, this problem can be corrected for via a magnetic field correction algorithm. This correction algorithm is illustrated in figure 5.6. Put very simply, this algorithm first performs a search for the field it expects to see at a given time (using the same desired field array used for the PI controller), and then performs a linear interpolation of the two EMDR data points that were acquired about that magnetic field. It then time aligns the interpolated EMDR sample with the particular field that was searched for. This is performed for every sample in the desired magnetic field array. This technique actually also provides a good means to
decimate the data set in the same process. (This is a very simplified description of the algorithm. It is more complicated because the ADC samples data in blocks. Therefore, the magnetic field samples are linearly interpolated while exponential averaging is performed on the EDMR data blocks to essentially decimate the data set.) As a result of this algorithm, inconsistencies in hardware latencies can be corrected for so that small hyperfine interactions can be better resolved.

For the past 8 years, our group has consistently detected a $g = 2.0030$ signal with appeared to be just one set of broad, and poorly resolved, hyperfine side peaks in almost all of the 4H SiC devices that we have experimented with. Only after this algorithm was implemented were we able to sharpen the hyperfine side peaks which has led to the discovery of a much smaller set of peaks slightly further away from the inner set. This along with the utilization of fast passage EDMR has allowed us to definitely identify a silicon vacancy defect in 4H SiC MOSFETs [3].

![Figure 5.6: Magnetic field correction algorithm. Because the magnetic field samples read in by the computer will not always align as a function of time, a field correction algorithm is needed.](image)
5.2.2 – RF Resonant Circuit

As discussed earlier, the size of the SDR effect via EDMR is not highly dependent upon the externally applied magnetic field. Therefore, much information is retained when performing EDMR at lower magnetic fields. To achieve this resonance, lower frequency electromagnetic radiation can be implemented through a resonant circuit. By passing a sinusoidal current through a surface coil or solenoid at radio frequencies, one is able to generate an oscillating $B_1$ magnetic field which is capable of flipping electron spins at lower magnetic fields. However, circuits working in the 10’s to 100’s of MHz frequencies require precision circuit design. If “long” cables are used, then transmission line analysis must be used when designing such a resonant circuit. But what defines a “long” cable? This requires an understanding of what is referred to as the electrical wavelength of a cable. The wavelength $\lambda$ of a signal propagating down a transmission line is defined by,

$$\lambda = \frac{c}{v}$$

(5.2.3)

where $c$ is the speed of light (2.998x10$^8$ m/s) and $v$ is the frequency of the signal. Transmission line analysis should be used when cables used are considered electrically long. These lines are characterized when the length of the line being used is longer than $\frac{1}{4}$ of the electrical wavelength, multiplied by the velocity factor (VF) of the cable being used. (This can be found on the spec sheet of any cable. The lower the value, the slower the wave will travel along the cable.) If a cable satisfies the following equation, then it is considered electrically long.

$$L > 0.25 \frac{c}{v} VF$$

(5.2.4)

For example, if it is desired to feed a 100MHz sinusoid through a cable that has a velocity factor of 0.5, then the longest cable that can be used without the need for transmission line analysis, is 0.375 m (1.23 ft). A frequency of 100MHz is actually slightly too low because the resonance of this condition will sometimes overlap with zero-field spin dependent transport signal discussed in section 3. Assuming $g = 2$, the resonant field can be approximated by,

$$B(G) = 357.234 \cdot v \ (GHz)$$

(5.2.5)

Therefore, 100MHz corresponds to a resonance at 35.7 G. In some of the data illustrated in the next section, the zero-field transport signal can sometimes be as wide as 100G or larger (centered at 0G).
Doubling the frequency to 200MHz will double the field at which resonance occurs. The 71.4 G center field provides enough separation between the resonant signal and zero-field transport signal. This remaining portion of this section describes the resonant circuit that was designed to meet this specification.

The components used for the RF section of the low-field spectrometer are a Stanford Research Systems SG384 RF Signal Generator (50Ω output impedance), HD Communications Corp. HD29347 Solid State High Power Module RF Amplifier (10 W, 50Ω input/output impedance), and a home built resonant circuit and coil. Much thought needs to be taken into account when designing the RF resonant circuit because it needs to be impedance matched to the RF amplifier at the resonant frequency so that maximum power can be delivered to the coil. Maximum current delivered to the coil will allow a maximum oscillating magnetic field to be generated across the sample. This is necessary when attempting to saturate the spin system for EDMR. A conventional tank circuit (parallel or serial resonant LC circuit) cannot be used because there is no way to match the impedance to 50Ω. For example, consider the simple series RLC circuit (the R is included to model the resistance of the coil) as illustrated in figure 5.6. The total impedance into the circuit can easily be calculated to be,

$$Z_{eq} = R + jX = R + j\left(\omega L - \frac{1}{\omega C}\right)$$

(5.2.6)

where \( R \) is the resistance and \( X \) is the reactance (capacitive and inductive components). Note that the impedance is composed of real (resistive) and imaginary (reactive) components. Similar to how the power is dissipated by a resistive component,

$$P = I^2R$$

(5.2.7)

The power stored in each of the reactive components is defined by,

$$P_c = I^2X_c = I^2\frac{1}{\omega C}$$

$$P_L = I^2X_L = I^2\omega L$$

(5.2.8) (5.2.9)

The circuit will be in resonance when the frequency of the RF source allows the reactive components to store the same power. This happens when,
\[ \frac{1}{\omega C} = \omega L \quad (5.2.10) \]

As a result, according to equation (5.2.6), the impedance will solely be due to the resistive components at the resonant frequency of the circuit. Using the previous equation, the resonant frequency of the circuit can be calculated by,

\[ 2\pi v_0 = \omega_0 = \frac{1}{\sqrt{LC}} \quad (5.2.11) \]

The quality factor of a circuit is defined as

\[ Q = \frac{\text{power stored}}{\text{power dissipated}} = \frac{I^2X}{I^2R} = \frac{X}{R} \quad (5.2.12) \]

The Q of the series resonant circuit can be defined many ways,

\[ Q = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 RC} = \frac{1}{R} \sqrt{\frac{L}{C}} \quad (5.2.13) \]

It is desired to make the Q as high as possible because it is desired to store as much energy within the circuit while minimizing the dissipated energy. Note that the quality factor is also equivalently defined by

\[ Q = \frac{v_0}{\Delta v} \quad (5.2.14) \]

where \( \Delta v \) represents the bandwidth of the circuit. (The bandwidth of the circuit is simply a measure of the frequency range in which the circuit responds with a voltage that is \( V > 0.707 V_{\text{max}} \).) As a result, the quality factor represents the sharpness of the circuit response as a function of frequency.

As mentioned before, one needs to design the circuit such that the real component of the impedance matches that of the output impedance of the RF amplifier (50\( \Omega \)) at resonance. In the case of the series resonant circuit, the real component will solely be due to \( R \) or the resistance of the wire that makes up the inductive coil. This resistance is \( << 1\Omega \) which implies that a separate component needs to be added in order to meet the 50\( \Omega \) matching requirement. It is undesirable to add an additional resistor
because it will dissipate more energy thereby reducing the Q. It is also undesirable to add an inductor because the additional wire increases resistance, thereby also reducing Q. Therefore, the only remaining passive component that can be added to meet this goal is a capacitor. If it is added in series with the series RLC circuit, it will look identical to the original resonant circuit with a modified capacitance. Therefore, a capacitor is typically added in parallel as illustrated in figure 5.7.

![Figure 5.7: Series resonant circuit and series resonant circuit with additional capacitor for impedance matching.](image)

One can show that the input impedance to this circuit is

\[
Z_{eq} = \frac{-R}{\omega C_2} - j \left[ R^2 + \left( \omega L - \frac{1}{\omega C_1} \right) \left( \omega L - \frac{1}{\omega} \left( \frac{C_1 + C_2}{C_1 C_2} \right) \right) \right] \omega C_2 \left[ R^2 + \left( \omega L - \frac{1}{\omega} \left( \frac{C_1 + C_2}{C_1 C_2} \right) \right)^2 \right] \tag{5.2.15}
\]

The impedance is written in this way because a few approximations can be made,

\[
\frac{C_1 + C_2}{C_1 C_2} \approx \frac{1}{C_1} \quad \text{if } C_2 \gg C_1
\]

\[
\omega L - \frac{1}{\omega C_1} = 0 \quad \text{when } \omega = \omega_0
\]

Using these approximations with the constraint \(Re\{Z_{eq}\} = 50\Omega\),

\[
Re\{Z_{eq}\} \approx \frac{1}{R(\omega C_2)^2} = 50 \Omega
\]
where $C_2$ can easily be calculated using the following,

$$C_2 = \frac{1}{\omega} \sqrt{\frac{1}{50R}}$$  \hspace{1cm} (5.2.16)

The imaginary constraint $Im\{Z_{eq}\} = 0$ can be used to find $C_1$. The imaginary component can be simplified to,

$$Im\{Z_{eq}\} \approx \omega^2 L + R \omega - \frac{1}{C_1} = 0$$

Note that there will be two resonant frequencies because of the quadratic nature of the equation. If it is assumed that $\omega L \gg R$ in the previous relationship, the $R \omega$ can be neglected so that $C_1$ can be found by,

$$C_1 = \frac{1}{\omega^2 L}$$  \hspace{1cm} (5.2.17)

Note that $C_1$ determines the resonant frequency of the circuit whereas $C_2$ matches the real impedance to the RF amplifier. Typically, $C_1$ and $C_2$ are made variable. The inductor in our case was a solenoid. The inductance of a solenoid is calculated by,

$$L = \mu_0 \frac{N^2 A}{l}$$  \hspace{1cm} (5.2.18)

where $N$ is the number of turns, $A$ is the area of the coil, and $l$ is the length. The magnetic flux density within the center of the solenoid is defined by,

$$B = \mu_0 \frac{Ni}{l}$$  \hspace{1cm} (5.2.19)

where $i$ is the current in the wire. It is important that when designing the coil to keep the number of turns small because the inductance increases with the number of turns squared. As a result, an increase in inductance will lower the maximum resonant frequency that can be obtained as indicated by equation (5.2.11). Additionally, increasing the number turns in a solenoid will not actually increase the generated magnetic field (when coils are spaced as close as possible) because the field is also inversely proportional
to the length of the solenoid. However, increasing the length (and number of turns) makes the oscillating field more uniform over the sample while simultaneously making the coil more mechanically stable. Using the previous equations, we simulated the response of the resonant circuit with a desired frequency of 209.94 MHz (corresponds to a EPR resonance of 75G) and a 4 turn solenoid with a radius of 5mm. The real and imaginary components of the impedance are illustrated in figure 5.8.

The most difficult aspect of designing any high frequency circuit are the practical implementations involved. Many of the previous circuits that were built in the lab suffered from stray capacitance and inductance which significantly affected the performance of the circuit. For the newly constructed circuit, expensive non-magnetic, low loss, variable capacitors were purchased from Voltronics, SMA connectors/cables were used instead of BNC connectors/cables, and the cables and connections of all the components in the resonant circuit were made as short as possible. The final circuit that was used for this study is illustrated in figure 5.8. The generated magnetic field of this solenoid was calculated by measuring the induced voltage on a sniffer coil. The sniffer coil was constructed simply by wrapping a few turns of wire around the existing solenoid used to generate the oscillating magnetic field. Its sole purpose was to sense the field that the solenoid generated. The induced voltage of a coil can be investigated with the use of Faraday’s Law of Induction,

![Figure 5.8: Simulated response of the resonant circuit with a resonant frequency of 210 MHz. The top plot illustrates the real component and the bottom plot illustrates the imaginary component. Note that at resonance, the real (resistance) component is almost matched to 50Ω and the imaginary component (reactance) almost goes to zero.](image)
\[ V_{EMF} = N_s \frac{d\Phi(t)}{dt} \]  

(5.2.20)

where \( N_s \) represents the number of turns on the sniffer coil. Expanding this relation for use with a solenoid (equation 5.2.19) yields,

\[ V_{EMF} = N_s A \frac{dB(t)}{dt} = N_s (\pi r^2) \left( \mu_0 \frac{N}{l} i \right) \frac{d}{dt} \cos(2\pi v_0 t) \]  

(5.2.21)

Here, \( v_0 \) is the frequency of the current in the solenoid and \( r \) is the radius, \( N_s \) is the number of turns, and \( l \) is the length of the sniffer coil. The magnitude of the current \( |i| \) that is forced through the sniffer coil can therefore be written as,

\[ |i| = \frac{V_{EMF} l}{2\mu_0 NN_s \pi^2 r^2 v_0} \]  

(5.2.22)

Plugging this relation back into equation (5.2.19) yields the field that the sniffer coil senses,

\[ B = \frac{V_{EMF}}{2N_s \pi^2 r^2 v_0} \]  

(5.2.23)

This relates the magnetic field generated at the center of the solenoid as a function of the measured induced voltage in units of Tesla. A \( x10 \) oscilloscope probe is used to measure the induced voltage on the sniffer coil because it reduces the capacitance that is presented to the load. This is important to use when working with high frequency circuits so that the circuit does not overload. Therefore, the voltage measured on the oscilloscope will be attenuated by a factor of 10. Multiplying by 10 to compensate for this loss and multiplying by \( 10^4 \) to convert the units of field to Gauss, the measured magnetic field \( B_1 \) becomes,

\[ B_1 = \left( \frac{V_{pp}}{2} \right) x10^5 \]  

(5.2.24)

Next, the response of the circuit was investigated. Figure 5.9 illustrates the data that was gathered when plotting the induced voltage of the sniffer coil as a function of driving frequency. The data in the figure was linearly interpolated so that a more accurate Q value could be calculated. From the figure, the circuit
responded maximally at 212.35 MHz. The bandwidth of the circuit is defined as the frequency range in which the response of the resonant circuit is \( > \frac{1}{2} P_{\text{max}} \). Because the voltage is proportional to the square root of the power, the bandwidth can also be defined as the frequency range in which the response of the resonant circuit is \( > \frac{1}{\sqrt{2}} V_{\text{max}} = 0.707 \cdot V_{\text{max}} \). From the data illustrated in the figure, the bandwidth of the circuit was measured to be 10.5 MHz which corresponds to a loaded Q value of slightly greater than 20. Note that it is very hard to get Q values much higher than this when working with such low frequencies.

![Figure 5.9: (left) Resonant circuit designed for the low-field spectrometer and (right) data gathered to test the response of the circuit. A sniffer coil was used to detect the EMF that was generated by the solenoid.](image)

### 5.2.3 – Lock-In Amplifier

In order to acquire the data more effectively, a virtual (software) lock-in amplifier (VLIA) was developed that allows the user to demodulate up to 4 harmonics at any desired phase simultaneously. Creating this in the virtual environment allows the user to be as flexible as needed. Figure 5.10 illustrates the block diagram of the VLIA and figure 5.11 illustrates the block diagram of each instantiated modulator block. The software allows the user to control multiple parameters including sampling rate, fundamental modulation frequency, and up to 4 demodulation harmonic frequencies. For each demodulator, the user has the ability to independently set the demodulation phase, low pass filter order, and time constant. All digital filters used in the VLIA are based on the Bessel filter topology so that minimum ringing would be present. The lack of steepness attributed to this type of filter was compensated by increasing the filter order. Because lock-in amplification requires the data to be sampled at a high rate, decimation is necessary after the final low-pass filtering stage to reduce the number of samples that are used to represent the final demodulated spectrum. This reduction factor is also a configurable. The VLIA utilizes a 16-bit NI PCI-6259 DAQ card that has a maximum sampling rate of \( F_s \approx 1MS/s \).
Figure 5.10: Block diagram of the VLIA developed in Labview. As indicated by the figure, the user has the ability to demodulate up to any 4 harmonics simultaneously.

Figure 5.11: Block diagram of the demodulator used in the VLIA. As indicated by the figure, the user has the ability to control the fundamental frequency, the demodulated harmonic, demodulator phase, low pass filter order and time constant, and decimation/averaging method.
The architecture of the ADC allows reading and writing data in blocks of length \( N \) for every software loop iteration \( i \). For our application, \( N \) is fixed by software to be 10 times less than that of the sampling rate of the ADC. Therefore, the loop rate (in Hz) of the software is defined by \( F_s / N = 10 \text{ Hz} \). The data is usually sampled at \( F_s > 250 \text{ kHz} \) so \( N \) is typically > 25 kS. This is far more samples than are needed to represent the demodulated spectrum. The VLIA allows the user to select between one of three options to reduce the amount of data from \( N \) samples to \( K \) samples. Here, \( K \ll N \) and \( N \) is an integer multiple of \( K \).

For the first option, the VLIA allows the user to perform a simple decimation where every \( Rth = (N/K)th \) data sample is selected while all other samples are discarded. The second option allows the user to average the entire block of \( N \) samples read in by the ADC. This option is not typically desired because it limits the effective sample rate to the loop rate of the software. (For the same reason, it is not recommended to use the GPIB interface to read data in from a hardware lock-in amplifier.) The third option and the one that should be utilized for all applications, is to perform a windowed average over the contiguous blocks of the data that are read in by the ADC. This process is illustrated in figure 5.12.

![Diagram](image)

Figure: 5.12: Illustration of the average and resampling block. This simplified illustration demonstrates how this block reduces the number of samples from \( N \) to \( K \) using a windowed averaging method. In this particular example, \( N = 10 \) and \( K = 5 \).
In the windowed averaging method, contiguous sub blocks of samples of size $P = 2R$ are averaged to reduce the block size from $N$ samples to $K$ samples. The samples are averaged using a zero-phase implementation so that they are perfectly time aligned to the original sampled data. This method is more advantageous than the conventional decimation in the sense that the averaging allows for additional reduction in noise. For example, consider the case where the loop rate of the software is configured to be 10 Hz. If it is desired by the user to reduce the block size from $N = 25kS$ to $K = 25 S$, then the sample rate is decreased from $N \times 10 = 250kS/s$ to $25 \times 10 = 250 S/s$. In the same process, the averaging of contiguous samples will also reduce the variance of the noise on each sample by a factor of $P = 2N/K = 2000$. Therefore, this method is almost always used because it simultaneously decimates the data while also reduces random noise. The VLIA was shown to successfully operate at the modulation frequencies up to and including 100kHz which is the common choice used for conventional ESR. More information regarding the technical aspects of lock-in amplification can be found in section 4 and the results obtained using this lock-in are illustrated in the following chapter.
CHAPTER 6 – HIGHER ORDER HARMONIC DETECTION OF EDMR

6.1 – Passage Effects

As noted in chapter 4, higher order harmonics can sometimes provide valuable information if certain conditions are met. More specifically, “passage” conditions are observed when certain experimental conditions are met. For devices that have defects that exhibit long spin-relaxation times, one can typically observe a “fast passage” absorption spectrum by simply increasing the modulation frequency. A particular material system that exhibits this long relaxation times is that of SiC [92]. Fast passage is a category of a passage effects which are typically observed when one performs continuous wave magnetic resonance with magnetic field modulation on a material system such that its demodulated components do not behave as expected. The observation of passage effects depends on the spin-lattice relaxation time $T_1$ of the material system under observation, the modulation frequency and amplitude, and microwave field amplitude. Because time $T_1$ is highly dependent upon temperature, decreasing the temperature will increase $T_1$ and therefore increase the chances of observing a passage effect.

In conventional continuous wave EPR, the Bloch equations discussed in section 2 are assumed to be in quasi-steady state and therefore, the differential magnetization components be approximated with zeros. Therefore, a solution to the three coupled differential equations can be solved analytically. This however is not the case for fast passage. Because these equations are not in steady state, one must obtain a solution numerically. It turns out that if numerically calculated (with the addition of magnetic field modulation) for a system with long relaxation time $T_1$, one can demonstrate that particular harmonics (including the possibility of the first) may take on the form of an absorption spectrum. In other words, the magnetization does not instantaneously follow the modulated magnetic field. As a result, the typical derivative is not observed. This is well understood in conventional EPR measurements [85] [86] [87] but has only been reported once in EDMR by us [3]. Here, we will consider the passage cases of adiabatic, rapid, and fast passage [87].

First, take for example the $z$ component of the magnetic field in the rotating reference frame during measurement. It is given by the slowly varying effective component (derived in chapter 2.5 – rotating reference frame) and the modulated component,

$$B_z = \left( B_0 - \frac{\omega}{\gamma} \right) + B_m \cos(\omega_m t) \quad (6.1.1)$$

where $\omega$ is the angular velocity of the applied electromagnetic radiation, $B_m$ is the modulation amplitude, and $\omega_m$ is the modulation angular frequency. The modified effective field is now,
\[ B_{\text{eff}} = B_z \hat{k} + B_x t' = \left[ \left( B_0 - \frac{\omega}{y} \right) + B_m \cos(\omega_m t) \right] \hat{k} + B_1 t' \]  

(6.1.2)

where \( B_1 \) is the amplitude of the oscillating magnetic component of the EM wave and \( t' \) represents the \( x' \) axis in the rotating reference frame as discussed in section 2. The type of passage condition in the experiment is defined by how fast resonance is swept through. In the case of continuous wave magnetic resonance, the passage rate is not determined by the slowly varying magnetic field, but rather the modulation in and out of resonance. Therefore, one must find the angular velocity of the magnetic field \( B_{\text{eff}} \) through the resonant center field. Because the \( z \) component of this field is oscillating, the angular velocity \( \Omega \) of the effective field in and out of resonance will determine the state of passage. Because the rate of change in position versus time is equal to velocity, the rate of change in angular position \( \theta \) versus time is equal to the angular velocity. The angle of the magnetic field in this case is defined by the angle the effective field makes with the rotating \( x \) axis,

\[ \theta = \tan^{-1} \left( \frac{B_z}{B_x} \right) = \tan^{-1} \left( \frac{B_0 - \frac{\omega}{y} + B_m \cos(\omega_m t)}{B_1} \right) \]  

(6.1.3)

In order to find angular velocity \( \Omega \), one must differentiate (4.5.7) with respect to time. Using,

\[ \Omega = \frac{d}{dt} \theta(u) = \frac{d\theta}{du} \frac{du}{dt} = \frac{d}{dt} \tan^{-1}(u) \frac{du}{dt} = \frac{1}{1 + u^2} \frac{du}{dt} \]

where \( u = B_z/B_1 \), and

\[ \frac{du}{dt} = \frac{d}{dt} \left( \frac{B_z}{B_1} \right) = \frac{1}{B_1} \frac{d}{dt} \left[ \left( B_0 - \frac{\omega}{y} \right) + B_m \cos(\omega_m t) \right] = \frac{B_m \omega_m}{B_1} \left( -\sin(\omega_m t) \right) \]

Then,

\[ \Omega = -\frac{B_1 B_m \omega_m \sin(\omega_m t)}{B_1^2 + \left( \left( B_0 - \frac{\omega}{y} \right) + B_m \cos(\omega_m t) \right)^2} \]  

(6.1.4)
This expression can be simplified significantly by realizing that, at resonance, the angle of the effective field with respect to the transverse plane is small. Therefore, the tangential small angle approximation can be used because \( \theta \) is small,

\[
\tan(\theta) \approx \theta \quad \rightarrow \quad \tan^{-1}(\theta) \approx \theta \quad \text{for small } \theta
\]

Therefore, at resonance, (6.1.3) can be approximated by,

\[
\theta \approx \frac{\left( B_0 - \frac{\omega}{\gamma} \right) + B_m \cos(\omega_m t)}{B_1}
\]

(6.1.5)

Differentiating this angle with respect to time yields,

\[
\Omega \approx \frac{d\theta}{dt} = \frac{1}{B_1} \frac{d}{dt} \left[ \left( B_0 - \frac{\omega}{\gamma} \right) + B_m \cos(\omega_m t) \right] = \frac{B_m \omega_m}{B_1} \left( -\sin(\omega_m t) \right) \approx \frac{B_m \omega_m}{B_1}
\]

(6.1.6)

This implies that resonance is swept through at a rate \( B_m \omega_m / B_1 \). This relation indicates that there are 3 ways to increase the rate of passage when utilizing magnetic field modulation. (1) Increase the modulation frequency, (2) increase the modulation amplitude, and or (3) decrease the oscillating \( B_1 \) magnetic field. Note however, that if \( B_m \) is increased too much, then too much time is spend out of resonance allowing time for relaxation of the spins.

The first passage condition is that of adiabatic passage. This occurs when the Larmor precession frequency in the rotating reference frame \( \omega_{eff} \) is much greater than the angular velocity of the effective field \( \Omega \),

\[
\omega_{eff} = \gamma B_{eff} \gg \frac{B_m \omega_m}{B_1}
\]

(6.1.7)

Gradually changing conditions allow the system to adapt to its configuration. This condition is the desired form of resonance, because the magnetization follows the effective field. Therefore, the spin packets have time to recover before consecutive modulation cycles pass through resonance. This is pictorially illustrated in figure 6.1.
At resonance, $B_{\text{eff}} = B_1 i'$, and equation (6.1.7) becomes,

$$\frac{\gamma B_1^2}{\omega_m B_m} \gg 1 \quad (\text{adiabatic passage condition}) \quad (6.1.8)$$

The next passage condition is rapid passage. This occurs when the angular velocity $\Omega$ of the effective field is much greater than the frequency of the spin relaxation process $T_1$.

$$\frac{B_m \omega_m}{B_1} \gg \frac{1}{T_1} \quad (6.1.9)$$

which leads to the more conventional form of the rapid passage case,

$$\frac{B_m \omega_m}{B_1} T_1 \gg 1 \quad (\text{rapid passage condition}) \quad (6.1.10)$$

Essentially, this condition sweeps through resonance in a time faster than it takes for the spins to relax. The third passage condition is that of fast passage. This condition occurs if the modulation period is much less than that of the spin relaxation process $T_1$. For example,
\[
\frac{1}{\omega_m} \ll T_1 \tag{6.1.11}
\]

which leads to,

\[
1 \ll T_1 \omega_m \quad (\text{fast passage condition}) \tag{6.1.12}
\]

In fast passage, an absorption spectrum can be observed after demodulating the first or higher order harmonics at various phases [3] [85] [86] [87]. It should not be so obvious as to why this occurs. In fast passage conditions, the Bloch equations cannot be solved analytically because the system is not in steady state and therefore, the derivatives cannot be approximated with zeros. The equations must be numerically solved for with the addition of magnetic field modulation. Others have demonstrated that this phenomenon can be simulated accurately enough to corroborate their results [87].

As mentioned before, these ideas are well understood and commonly observed in EPR. However, it is not so obvious that this effect would be observed in EDMR because of the extremely different detection mechanism. In EDMR, one detects a spin dependent current and in EPR, one detects a magnetization through coil or microwave cavity. It should be clear that the spin relaxation times of $T_1$ and $T_2$ play a role in the EPR magnetization signal (see section 2), however, it isn’t clear how these relaxation processes affect EDMR signals. It turns out that there is not any literature regarding this issue. We were actually the first to report on the observation of fast passage EDMR in 4H SiC MOSFETs [3]. We think this method works so well on the SiC devices because the individual defects in the SiC lattice have long spin-lattice relaxation times [92]. This time is temperature dependent. It describes the time at which the electron spin magnetization will return to equilibrium after excitation. These long $T_1$ times may be due in part to the small spin-orbit coupling and the stiffness of the SiC lattice which may lead to favorable phonon distribution for long $T_1$. As a result, the polarized spins not only couple to the lattice, but may also be coupled to the other electron in the spin pair described in section 3.

There are many advantages of fast passage EDMR. The first is that it sharpens the EDMR spectrum. The reason for this is because only those defects sensitive to the fast passage condition given by (4.5.16) are demodulated. Therefore, background signals do not get detected, and therefore, partially buried hyperfine peaks can be better resolved. The second advantage of fast passage detection is that the demodulated signal results in an absorption signal rather than a derivative. This is important when hyperfine interactions are present because it allows the experimentalist to more accurately extract the integrated intensities and locations of the hyperfine side peaks. A final advantage of fast passage is that the detection allows one to detect higher order harmonics away from flicker noise, thereby reducing the
noise in the measurement. (However, reduction in amplitude of higher order harmonics and gain bandwidth tradeoffs in the preamplifier are the two disadvantages of higher order harmonic detection.) There are also a few disadvantageous of fast passage detection that should be pointed out. The first disadvantages is that it is very sensitive to modulation phase. If one is slightly out of phase, it is possible to detect a signal which has a fast passage and non-fast passage component. However real, this will result in distorted looking spectra. A rule of thumb is to first perfectly phase adjust the first harmonic, then attempt to move to higher order harmonic detection of the fast passage signal. This is easily done with a digital lock-in amplifier because its perfect synchronization of signal generation and sampling scheme. We have actually developed and utilized a multiband virtual lock-in amplifier (discussed in the previous sections) to demodulate the harmonics of a various 4H SiC devices through resonance.

6.2 – 4H SiC BJT Response: Defect Aggregate

An initial study of a 4H SiC BJT revealed a spectrum with various peaks that are unresolved [8]. In order better resolve these interactions, we proceeded to attempt to resolve them by utilizing the advantages of fast passage effects. We found that by demodulating the 2nd harmonic, quadrature component at a particular frequency, one is able to detect the fast passage condition. Figure 6.2 illustrates these results. Note that the demodulated in-phase and quadrature components change shape as a function modulation frequency. For example, at 6kHz, the in-phase signal flips sign while the quadrature signal changes from a second derivative to an absorption spectrum. This is the transition into fast passage. Figure 6.3 compares the well resolved fast passage spectrum to that of the spectra acquired using the conventional first harmonic detection. The left column of figures illustrate 3D surface plots constructed from linearly interpolated EDMR scans acquired at high field and frequency for different orientations of the device within the applied magnetic field. The top right plot illustrates the conventional 1st harmonic detection of the EDMR signal with extensive signal averaging and bottom right figure illustrates the 2nd harmonic fast passage detection of the EDMR signal.

The fast passage detection allows for better resolution of overlapping hyperfine interactions. As illustrated in the figure, the detected spectrum has a dominating center with an isotropic $g = 2.0030 \pm 0.0002$. It has multiple sets of equally spaced nearly isotropic hyperfine side peaks indicated by B, C, D, E, F, and G in the figure. (These peaks are actually slightly anisotropic; they move very slightly when the sample is rotated in the magnetic field.) The specific defect that generates this spectrum is unknown; however, it likely some sort of defect aggregate. A program was written to perform a least squares fit on the fast passage signal so that a better amplitude measure of the individual hyperfine peaks could be extracted. The amplitudes of thirteen Gaussian functions were fit to the fast passage spectrum. The means
of these functions were chosen by the user with a mouse pointer and the variances were fixed to 1.5G (the modulation amplitude during experiment). Thirteen peaks were chosen for simulation because there consistently appears to be this number in the spectrum. However, other peaks may be buried within the spectrum which makes a detailed identification of the defect very difficult.

The results of the least squares fitting are illustrated in figure 6.4 and table 6.1. As illustrated in the figure, the data was almost perfectly replicated with thirteen Gaussian functions. The relative integrated intensity of each peak is listed in table 6.1. (Note that the intensities of all the peaks add to 100%.) Some of the more interesting ratios include the larger hyperfine peaks with respect to the center line.

Figure 6.2: EDMR results on a 4H SiC BJT. This plot illustrates (top left) the in-phase, 2nd harmonic detected spectrum as a function of modulation frequency, (top right) the quadrature, 2nd harmonic detected spectrum as a function of modulation frequency, and (bottom) peak-to-peak amplitude of the quadrature, 2nd harmonic detected spectrum as a function of modulation frequency. Note that a 6 kHz, there is a change in line shape for both the in-phase and quadrature demodulated spectra. The in-phase signal flips sign while the quadrature signal changes from a second derivative to an absorption spectrum. This is the transition into fast passage.
For example, the integrated intensity of both B peaks to the center A peak is 32.97:67.03 and the ratio of integrated intensity of both E peaks to the center A peak is 10.86:89.14. At first glance of the spectrum, it appears that these sets of peaks may be associated with a spin $I = 1/2$ magnetic nucleus at the defect site. However, the ratios do not match up with the ratio of magnetic to non-magnetic nuclei of silicon $^{29}\text{Si}:^{28}\text{Si}/^{30}\text{Si} = 4.68:95.32$ or carbon $^{13}\text{C}:^{12}\text{C} = 1.07:98.93$. This suggests that the simulated spectrum may actually be composed of additional peaks that lie at the same field or are buried beneath neighboring peaks. Therefore, the defect spectrum is likely a defect aggregate [8]. The large number of multiple peaks observed in the spectrum indicate that the dopant atoms of nitrogen ($I = 1, \approx 100\%$ abundant) and aluminum ($I = 5/2, \approx 100\%$ abundant) may play a role.

Figure 6.3: Data obtained on 4H SiC BJT at X-band frequencies. The BJT signal is composed of a dominant isotropic center line with multiple sets of isotropic (and slightly anisotropic) side peaks. The left column of figures illustrate 3D surface plots constructed from linearly interpolated EDMR scans acquired for different orientations of the device within the applied magnetic field. The top right plot illustrates the conventional 1$^{st}$ harmonic detection of the EDMR signal with extensive signal averaging and bottom right figure illustrates the 2$^{nd}$ harmonic fast passage detection of the EDMR signal. The spectra indicate that a defect aggregate is most likely involved.
Another reason why the spectrum is probably associated with multiple defects is the presence of a large number of hyperfine interactions. As illustrated in figure 6.5, the side peaks associated with these interactions are extremely small (more than 70x smaller than the dominating center line) and highly anisotropic with splitting of many hundred Gauss. Because our low-field spectrometer cannot generate magnetic fields to cancel these large hyperfine interactions, we were not able to study these interactions using the zero-field mixing effect. Even though the multiple defects present in the BJT defect are not completely understood, it is a great device for the investigation of the zero-field SDR phenomenon, because the hyperfine pattern is complex and the anticipated recombination current versus junction voltage is well understood to first order.

Fig. 6.4: Least squares fit of fast passage spectrum illustrated in figure 6.2 using Gaussian functions.

<table>
<thead>
<tr>
<th>Peak</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>55.12</td>
<td>12.96</td>
<td>2.05</td>
<td>1.94</td>
<td>3.8</td>
<td>0.78</td>
<td>0.35</td>
</tr>
<tr>
<td>Integrated Intensity Left / Right (%)</td>
<td>14.15</td>
<td>2.16</td>
<td>2.05</td>
<td>2.92</td>
<td>1.18</td>
<td>0.54</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: Amplitudes of the simulated Gaussian functions illustrated in figure 6.3.
Figure 6.5: 650G wide X-band (left) spectra and (right) its derivative illustrating very small anisotropic hyperfine interactions. These hyperfine interactions are not studied in this work because the low-field spectrometer does not allow for such high fields to be generated to cancel the fields produced by these interactions.

6.3 – 4H SiC MOSFETs: The Silicon Vacancy Defect

For several years, our group has made a tentative argument that the silicon vacancy center is the dominant recombination center detected in 4H SiC MOSFETs based upon EDMR studies [2], [93]. However, these studies could not provide definite identification of the vacancy center because they were unable to resolve any hyperfine interactions. With many spectrometer improvements and utilization of new detection techniques, in particular the utilization of fast passage effects, we have been able to better resolve the $^{13}$C hyperfine side peaks associated with the silicon vacancy center [3].

Last year, we definitively identified a silicon vacancy as being the dominant defect associated with recombination in 4H SiC MOSFETs [3]. We utilized the hyperfine parameters reported for the well characterized conventional EPR spectrum of the silicon vacancy [94] [95] [96] to interpret the EDMR spectrum. With the external magnetic field applied in the direction of the crystalline c axis, the isotropic Si hyperfine coupling constant $A_{iso} = 2.96$ G for the Si vacancy in either the cubic or hexagonal site of the 4H SiC lattice. The hyperfine parameters for carbon are $A_x = 27.35$ G, $A_y = A_z = 10.1$ G for the cubic vacancy site and $A_x = 28.36$ G, $A_y = A_z = 11.1$ G for the hexagonal vacancy site. The hyperfine parameters are much larger for the magnetic carbon atoms simply because these atoms reside closer to the unpaired electron(s) at the defect site. The defects within a device will contain vacancy centers at both the cubic and hexagonal sites. As a result, the splitting observed will be the average of the hexagonal and cubic parameters. Because the carbon and silicon atoms can be magnetic in multiple possible combination
around the defect site, the relative amplitudes can be calculated with the product of two binomial distributions,

\[
P(r^{29}\text{Si AND } s^{13}\text{C}) = \left[ \binom{R}{r} p^{(29\text{Si})}^r \left(1 - p^{(29\text{Si})}\right)^{R-r} \right] \cdot \left[ \binom{K}{k} p^{(13\text{C})}^k \left(1 - p^{(13\text{C})}\right)^{K-k} \right]
\]  

(6.3.1)

where \(R = 12\) equals the number of nearest neighbor \(^{28}\text{Si}\) or \(^{30}\text{Si}\) (non-magnetic) atoms, \(r\) equals the number of nearest neighbor \(^{29}\text{Si}\) atoms (magnetic atoms), \(p^{(29\text{Si})}\) equals the probability of \(^{29}\text{Si} = 4.7\%\), \(K = 4\) equals the number of nearest neighbor \(^{12}\text{C}\) (non-magnetic) atoms, \(k\) equals the number of nearest neighbor \(^{13}\text{C}\) atoms (magnetic atoms), and \(p^{(13\text{C})}\) equals the probability of \(^{13}\text{C} = 1.1\%\). The probability of various combinations of magnetic and non-magnetic nuclei in the vicinity of a Si vacancy are listed in table 6.2. As illustrated in figure 6.6, the vacancy spectrum has a strong central line due to defect sites at which no magnetic carbon nuclei are present at either the four carbon nearest neighbor atomic sites. Also, buried within this center line are two strong inner side peaks separated by 2.96G due to the presence of only one \(^{29}\text{Si}\) magnetic nucleus and a weaker set of side peaks separated by 5.92G due to the presence of two \(^{29}\text{Si}\) magnetic nuclei. The five lines just described buried in the center line account for about 95\% of the total integrated intensity of the spectrum. (The inability to resolve these inner peaks is likely due to broadening that is typically encountered in EDMR or the disorder in the Si/SiC interface region.) Nearly all the remaining spectrum comes from sites with one magnetic \(^{13}\text{C}\) nucleus in the inner core of the vacancy.

<table>
<thead>
<tr>
<th># of NN (^{13}\text{C})</th>
<th># of NN (^{29}\text{Si})</th>
<th>Total Probability</th>
<th>Number of Pairings</th>
<th>Peak Amplitude</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.539</td>
<td>1</td>
<td>0.539</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0.317</td>
<td>2</td>
<td>0.159</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0.085</td>
<td>4</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0.014</td>
<td>8</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.024</td>
<td>2</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.014</td>
<td>4</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.003</td>
<td>8</td>
<td>&lt;&lt;1%</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.2: Table illustrating the more probable combinations of magnetic nuclei at the site of a silicon vacancy.
These magnetic carbon nuclei sites are responsible for two pairs of weak but far more widely separated side peaks, two with a separation of about 13G and a second pair, with one third the amplitude of the first, separated by about 28G. (The integrated intensities for the two carbon peaks include contributions from sites with zero, one, or two magnetic silicon nuclei.) As illustrated in figure 6.6 (b), the magnetic $^{13}$C nucleus can reside at any of the four nearest neighbor carbon sites. Two peaks are observed because sites $B$, $C$, and $D$ are equivalent and yield a resonance pattern different than when the magnetic $^{13}$C nucleus is positioned at site A. Note that this pattern is only detectable when the magnetic field is aligned with the crystalline c-axis. As the device is rotated, the possibility of detecting four separate sets of side peaks becomes possible because the carbon sites are no longer equivalent with respect to the applied magnetic field. However, these peaks are hard to resolve due to the broadening encountered in EDMR.

Figure 6.6: Illustration of the silicon vacancy in 4H SiC and the results of the simulation using previously reported hyperfine parameters.
Figure 6.7 compares the model illustrated in figure 6.6 and the $2^{nd}$ harmonic detected fast passage spectrum obtained at X-band [3]. Note that the model and experimental results are in close agreement. The very close correspondence observed between the experimentally observed SDR response and the hyperfine parameters previously reported for the well characterized negative silicon vacancy spectrum in large volume cubic centimeter samples via EPR provide very strong evidence linking the electrically active SiC MOSFET defect to this spectrum. Since this SDR spectrum has been observed in a wide variety of 4H SiC MOSFETs prepared under a wide variety of processing parameters, this study indicates that the silicon vacancy center is a trapping center in the SiC which is of widespread importance in SiC MOS technology.

Figure 6.7: This figure illustrates the simulated and measured EDMR response using the $2^{nd}$ harmonic detection of the fast passage signal. The close correspondence between the simulation and the data indicate that the observed spectrum is most likely dominated by the silicon vacancy.
CHAPTER 7 – RESULTS ON ZERO-FIELD DETECTION OF SDR AND SDT

We observe a magnetic field induced change in current in various microelectronic devices in the absence of electromagnetic radiation at zero and near-zero magnetic field. We show that by utilizing low frequency (audio) magnetic field modulation, one can clearly resolve electron-nuclear hyperfine interactions at low magnetic fields in 4H SiC bipolar junction transistors (BJTs), 4H SiC metal-oxide semiconducting field effect transistors (MOSFETs), amorphous SiC dielectrics, and Si power MOSFETs. This sections is intended to illustrate and explain these results.

7.1 – Bipolar Junction Transistors

The first device that we investigated at zero-field was the 4H SiC BJT. More specifically, the base collector junction was forward biased in order to investigate spin dependent recombination (SDR) in a 4H SiC diode. The process of SDR in a forward biased diode is illustrated in figure 7.1 along with the photo and layout of the BJT device that was investigated. The BJT devices were used in this study because they produce large SNR signals, the recombination current within either diode of the device can easily be simulated, and the EDMR response is well documented at X-band and exhibits many hyperfine interactions.

![Figure 7.1: Process of SDR in a pn junction via paramagnetic deep level defects through the formation of spin pairs. (a), (1) conduction electron and defect electron couple to form a triplet pair, (2) pair dissociates because recombination is forbidden. (b), (1) conduction electron and defect electron couple to form a singlet pair, (2) because angular momentum is conserved, recombination of electron and hole is now possible. (c) Top view photograph and (d) simplified illustration of the cross sectional view of the device used in this study.](image)

148
7.1.1 – BJT Zero-Field Response

A mesh plot of representative low-field EDMR responses is illustrated in figure 7.2 (a) while subjecting the device to an oscillating RF magnetic field at 200MHz at various $B_1$ amplitudes. Three distinct signals are apparent. The line centered at zero Gauss was not expected; however, the signals centered at $\pm 71 \, G$ are consistent with the conventional EPR resonance conditions. That is, with an RF oscillating fields of frequency $\nu = 200 \, MHz$, the resonance condition $\hbar \nu = g \mu_B B$ yields $B = \pm 71 \, G$. The amplitudes of the zero-field and resonant signals are plotted in figure 7.2 (b) as a function of $B_1$. Here, the fully saturated resonant signal is about 2.35 times larger than that of the zero-field signal. Figure 7.2 (c) compares the spectra obtained with and without the $B_1$ oscillating magnetic field.

Figure 7.2: Low-field detection of SDR via EDMR compared to zero-field detection of SDR in the base collector junction of a 4H SiC BJT. (a) 3D mesh of representative EDMR scans on a SiC diode for a series of $B_1$ amplitudes. (b) Amplitude of the zero-field and low-field resonant signals as a function of $B_1$. The EDMR amplitudes increase monotonically with increasing $B_1$. Note that the increase in the $B_1$ field saturates the resonant SDR response but doesn’t affect the amplitude of the zero-field response. (c) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied. The arrows in the resonant trace indicate the Breit-Rabi shift of the normally (high-field) isotropic hyperfine peaks at low-fields. (d) Comparison of the derivative of a spectrum acquired without any modulation (blue, top) and the spectrum obtained from a 1kHz, 2 Gauss modulation (red, bottom).
(Note that because of the low-field condition, the normally (high-field) isotropic hyperfine peaks, indicated by the arrows in the figure are shifted towards zero Gauss. This is an observation consistent with the Breit-Rabi correction. [52] [97]) As expected, the resonant signals are no longer present in the trace without application of the RF radiation; however, the zero-field signal remains. At exactly zero-field, the axis of quantization is defined by the field the electron experiences from its local magnetic environment while at relatively high magnetic fields (3 kG), it is defined primarily by the externally applied field. At low fields (< 100 G) however, the field the electron experiences is the vector sum of the electron’s local magnetic surroundings and externally applied magnetic field. As a result, the distribution of the spin pair’s relative orientation will gradually be modified as zero-field is approached. Modifying the distribution of the spin pair’s orientation essentially modifies the singlet-to-triplet ratio and hence, the recombination current. The zero-field spin pair distribution is not necessarily random as is the case for resonance. If they were, one would expect the amplitude of the saturated resonant and zero-field SDR responses to be equal which are clearly not as illustrated in figure 7.2 (a) and 7.2 (b). This idea of a gradually changing spin pair orientation distribution also explains why line broadening is observed in the ZFSDR response. This SDR detection mechanism also applies to the defect sites which experience a hyperfine induced magnetic field due to neighboring magnetic $^{29}\text{Si}$, $^{13}\text{C}$, and $^{14}\text{N}$ nuclei. When a defect site is coupled to one or more magnetic nuclei, it experiences a hyperfine field that can be “cancelled” by the slowly varying $B_0$ field. As a result, a response due to this interaction is observed away from 0 Gauss, at a field essentially corresponding to the magnitude of the hyperfine field. Figure 7.2 (d) compares the zero-field response detected with and without magnetic field modulation. As illustrated, utilizing lock-in amplification for demodulation greatly enhances the resolution of the spectrum. This technique essentially makes SDR sensitive to a single frequency, phase, and field direction. Magnetic field modulation allows the removal of flicker noise, current drift, background artifacts, and non-magnetic field induced artifacts from the SDR measurement, thereby sharpening SDR response and allowing more sensitive detection of hyperfine interactions. The peak-to-peak width of the spectrum without magnetic field modulation is approximately 20 Gauss and reveals only slight hints of hyperfine side peaks when the derivative is taken. The spectrum acquired with magnetic field modulation of 2 Gauss at 1 kHz has a central peak-to-peak width of only about 6 Gauss and quite significantly more visible hyperfine side peaks.

7.1.2 – BJT Hyperfine Interactions

To further investigate the hyperfine interactions of the zero-field response, we compared the results to that of the high-field resonant response. Figure 7.3 (a) compares the zero-field and at X-band (9.5GHz, 3394G) spectra and figure 7.3 (b) compares their derivatives. (The derivatives are used to enhance the observation of the hyperfine interactions.) For each side peak present in the X-band spectrum, a similar
corresponding peak is present in the ZFSDR spectrum. (Although a convincing identification of the defect (or defects) responsible for the EDMR spectrum is not yet determined, a plausible identification would be a defect aggregate, likely a divacancy or vacancy/antisite pair [8] [3] as discussed in chapter 6) Recall, that although the hyperfine interaction peaks in the X-band and ZFSDR traces are not in precisely the same location with respect to the center line, the overall patterns are closely correlated. An exact correspondence should not be expected for several reasons. Among them, at extremely low external fields, the nuclear magnetic moment quantization axis is likely different than what it would be at high field because the local magnetic field experienced by the nuclei is not dominated by the applied field but by the field due to the nearby unpaired electron. We speculate that at very low fields, the nuclei will tend to reorient themselves along the field lines produced by the electron, thereby modifying the field that the magnetic nucleus exerts on the electron. As a result, the hyperfine peaks will shift at low externally applied fields.

Figure 7.3: (a) First derivative of spectra observed at zero field (blue) and at X band (red) taken at 9.5 GHz, 3394 Gauss. (b) Second derivative of the spectra illustrated in (a). Note that each peak observed at X band is also seen in the zero field spectrum. (c) Zero-field and (d) derivative of spectrum acquired with low modulation amplitude (~ 0.25 Gauss). Note the deviation in slope at precisely zero Gauss illustrated by inset of (c) which shows up as a double peak in (d).
Figure 7.3 (c) and 7.3 (d) illustrate the spectrum and its derivative obtained when reducing the modulation amplitude to 0.25 Gauss, respectively. Notice the inflection point precisely at zero Gauss which is significantly more obvious when the derivative is taken as illustrated by the double peak in figure 7.3 (d). This split peak is precisely due to the exchange and or dipolar interactions discussed earlier in chapter 3.

The ideas expressed here with regard to hyperfine interactions and zero-field spin dependent phenomena are somewhat similar to those expressed by others to explain magnetic field effects on tunneling in double quantum dots [20] [21] [23] [24] [25] and magneto-resistive effects in organic semiconductors [26] [27]. In both of those cases however, the hyperfine interactions invoked involved a multitude of unequal, essentially random, distributions of interactions which failed to yield a distinguishable pattern of hyperfine spectra or an unambiguous interpretation of the details of the spin dependent response. Therefore, because these studies did not involve any direct evidence hyperfine interactions, their conclusions were only speculative. Because our measurements utilize single crystal devices, the random distribution of hyperfine fields is replaced by one which is not random and therefore allows direct correlation with a more conventionally observed pattern of hyperfine interactions.

7.1.3 – Demonstration of Recombination

Because our zero-field response takes place in a well understood pn junction, we can directly demonstrate that its root cause is spin dependent recombination. Assuming a uniform distribution of trapping centers, to first order, the recombination current within the space charge region of a p-n junction can be described by [98]

\[ J_r = \frac{q n_i W}{2} \cdot v_{th} N_t \sigma \cdot exp \left[ \frac{q V_a}{2kT} \right] \]  

(7.1.1)

where \( v_{th} \) is the thermal velocity, \( N_t \) is the density of recombination defects, \( \sigma \) is the capture cross section of the defect, \( n_i \) is the intrinsic carrier concentration for 4H SiC at room temperature, \( V_a \) is the junction bias, and \( W \) is the width of the depletion region. For any forward bias up to the built-in voltage \( V_{bi} \), the depletion width is given by [98],

\[ W = \left[ \frac{2\epsilon (N_a + N_d)(V_{bi} - V_a)}{qN_a N_d} \right]^{1/2} \text{ for } 0 \leq V_a \leq V_{bi} \]  

(7.1.2)

where \( \epsilon \) is the permittivity of the semiconductor, \( N_a \) is the density of ionized impurity acceptor atoms, and \( N_d \) is the density of ionized impurity donor atoms. Here the built-in voltage \( V_{bi} \) is defined by [98],
\[ V_{bi} = \frac{kT}{q} \cdot \ln \left( \frac{N_a N_d}{n_t^2} \right) \] (7.1.3)

The relative amplitude of the recombination current versus applied bias can be calculated with moderate accuracy using the following,

\[ j_r = C \cdot \sqrt{V_{bi} - V_a} \cdot \exp \left[ \frac{qV_a}{2kT} \right], \quad \text{for } 0 \leq V_a \leq V_{bi} \] (7.1.4)

where the built-in voltage \( V_{bi} \) is accurately known and \( C \) is a device dependent constant which is immaterial in our case. Once the applied junction voltage exceeds the built-in voltage, the depletion region ceases to exist. This response is plotted in Fig. 7.4 (a). We measured the peak-to-peak amplitude of the zero-field response as a function of applied bias and plotted it against the recombination current calculation illustrated in Fig 7.4 (b). The correspondence is extremely close demonstrating that the zero-field phenomenon is certainly due to SDR in the diode’s space charge region. Plotted in Fig. 7.4 (c) is the percent change in SDR current versus applied bias which is simply calculated by dividing the zero-field peak-to-peak amplitude by DC diode current (\( \Delta I/I \)). The maximum response is almost 1% at a bias of 2.35 V. This is an exceedingly large effect at such small magnetic fields, but it is consistent with SDR phenomena involving triplet/singlet transitions as proposed by Kaplan et al. [13] Figure 7.4 (d) illustrates the integrated data that was acquired used to plot the data points in figs 7.4 (b) and 7.4 (c). Note that at significantly high junction biases (\( V_a \geq 2.65 \) V), the SDR signal begins to reduce in amplitude. Also, the inflection discussed earlier and illustrated in figs. 7.3 (c) and 7.3 (d) becomes more prominent and can be observed in the integrated intensity spectra. The inflection point is always observed in the SDR response of the BJT, however it becomes more prominent as the applied junction bias is increased. We speculate that this inflection is caused by the mixing of the three triplet states with the singlet state due to the lowered Zeeman energy at zero magnetic field. The amount of inflection increases as the applied junction bias is increased; a result suggesting that the applied bias affects the magnetic fields (or Zeeman energies) at which the triplet states mix with the singlet state. We repeated this measurement in a 3-layer magnetic shielded Mu-metal chamber to eliminate the possibility of Earth’s low DC magnetic field (\( \approx 0.5 \) Gauss) affecting the measurement. The spectrum obtained was identical to that obtained outside the chamber, a result expected considering the use of our magnetic field modulation technique. A somewhat similar phenomenon was observed [23] and simulated [21] in the tunneling of electrons in double quantum dots. These measurements showed a split peak precisely at zero-field due to singlet-triplet mixing where the amount splitting was shown to increase with quantum dot coupling.
Figure 7.4: This set of figures illustrates the behavior of the zero-field response as a function of applied bias. (a) Simulation of the recombination current for a 4H SiC diode at room temperature. (b) Illustration of the peak-to-peak amplitude of the zero-field signal as a function of applied bias plotted against the simulated recombination current displayed in (a) and the DC diode current. The close similarity of the plots demonstrates that this phenomenon is a recombination process. (c) The percent change in the zero-field SDR current acquired by dividing the peak-to-peak amplitude by the dc current as a function of bias voltage. Note that this response peaks at 2.35 volts with a relatively large change of almost 1%. (d) Integrated spectra acquired for the data plotted in Figs. 4(b) and 4(c). Note that as the bias is increased beyond the built-in voltage ($V_a \geq 2.65$ V), the SDR signal begins to decrease. Note also that beyond this voltage, a double peak appears to be forming. The four traces on the inset of this figure illustrate the transition of the single peak into two.

7.1.4 – Zero-Field Response vs. Temperature

As mentioned earlier, the SiC BJT would serve as a great sensing device for absolute magnetometry. However, the magnetometer would only be beneficial if its response was stable over a wide range of temperatures. As illustrated in figure 7.5, the response is in fact, independent of temperature. That is, although the overall amplitude of the response does change, the normalized response as a function of field is within experimental error, constant. In other words, the splitting of the hyperfine interactions remains the same as the temperature changes. The decrease in SNR at lower temperatures...
takes place because, as the temperature gets lower, the built in junction of the diode increases. As a result, more voltage is required to achieve the same current at lower temperatures. Two other factors can influence the overall amplitude but not the normalized amplitude versus field. (1) The spin lattice relaxation times are fairly strong as a function of temperature, and at really low temperatures, carrier freeze out becomes a problem.

Figure 7.5: BJT zero-field spectrum versus temperature. Note that the hyperfine lines are independent of temperature as indicated by the first and second derivatives.

7.2 – Zero- and Low-Field Transport in 4H SiC MOSFETs

In the previous section, it was demonstrated that detection of spin dependent transport at zero-field can provide some very useful information about deep level defects and electronic transport. More specifically, it was demonstrated that the detected zero-field response is due to recombination from singlet triplet mixing in the space charge region of the base collector junction of a 4H SiC BJT. It was also shown that electron nuclear hyperfine interactions can be detected at magnetic fields which cause singlet and triplet states to mix at defect sites in which a magnetic nucleus is present. It turns out that this effect can also be observed in 4H SiC MOSFETs. In this section, we demonstrate the zero-field mixing phenomenon can be detected with the biasing schemes of DCIV, charge pumping, and the bipolar amplification effect (BAE) [99].

7.2.1 – Zero-Field Detection of SDR using DCIV

The first MOSFET tested in the low-field spectrometer was an older generation 4H SiC n-channel MOSFET made by Cree in 2003. This device was characterized as having a mobility of 5.7 cm²/V·s, had a thermal ONO anneal, had a gate area of 200 x 200 μm², and the source and drain regions were doped with nitrogen. They suffered from many defects that were located precisely at the interface, one of those being the silicon vacancy just described. In this study, we apply the conventional DCIV biasing method
as a means to investigate the low-field transport. As illustrated in the inset of figure 6.11(c), the DCIV method consists of forward biasing the drain and source to substrate junctions while also applying a separate voltage to the gate [98] [100]. By varying the voltage at the gate, one may control the relative fraction of electrons to holes at the interface. At a particular voltage, the number of holes and electrons will be equal at the interface, thereby creating the greatest chance of recombination if recombination centers are concentrated at that interface. The results when applying -2.5V to the source and drain and 0V to the gate are illustrated in figure 7.6. As illustrated in figure 6.11, similar to the BJT, the hyperfine interactions associated with the zero field response can be detected at very low magnetic fields. Unlike the BJT, the amplitude of the zero-field response is slightly dependent upon the amplitude of the oscillating $B_1$ magnetic field applied during experiment. However, the zero-field SDR signal in the MOSFET can still be detected when there is no application of the RF radiation as illustrated in figure 7.6 (c) and (d). Coincidently, the ratio of the zero-field SDR signal to the fully saturated resonant signal is almost exactly the same as it was for the BJT. This value is approximately 2.35.

These results illustrated in figure 7.6 demonstrate the analytical power of the combination of zero- and low-field SDR spectroscopy. As mentioned earlier, performing magnetic resonance experiments at low magnetic fields may be beneficial for many reasons. Not only is the hardware cheaper, lighter, and requires a significantly smaller footprint, but there are many physical effects that occur at low-fields that may be advantageous. More specifically, second order corrections, first calculated by Breit and Rabi, demonstrate that the doublet center from a spin $I = 1/2$ nucleus will maintain its hyperfine peak splitting at low fields, however, it will be shifted in the direction of lower field [52] [97]. The corrected center field $B_C$ of the splitting can be shown to be,

$$B_C \approx B_0 - \delta \approx B_0 - \frac{\Delta B^2}{4B_0} \quad (7.2.1)$$

where $B_0$ is the unperturbed resonant field and $\Delta B$ is the magnetic field splitting of the hyperfine side peaks. Being able to observe these shifts is advantageous because it allows one to sometimes better resolve side peaks, particularly if they are obscured by the center line. For example, the shift of the inner side peaks of the silicon vacancy, separated by 12 G, allows one to better resolve the peak at the lower field as illustrated in figure 7.7 (a) and (b). The other advantage of using low-field resonance is that it is sometimes able to reduce g strain, that is broadening due to the disorder effects on the g tensor components. This effectively narrows the line width of the dominant signal, thereby allowing buried hyperfine interactions to become better resolved.
Figure 7.6: Low-field detection of SDR via EDMR compared to zero-field detection of SDR. (a) 3D mesh and (b) 2D traces of representative EDMR scans on a SiC MOSFET for a series of $B_1$ amplitudes. (c) Amplitude of the zero-field and low-field resonant signals as a function of $B_1$. The SDR amplitudes increase monotonically with increasing $B_1$. Note that the increase in the $B_1$ field saturates the resonant SDR response and also affects the amplitude of the zero-field response unlike the case for the BJT. (d) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied.

Figure 7.7 (c) illustrates the zero-, low-, and high-field (shifted from 3395G to 0G) spectra and figure 7.7 (d) illustrates the comparison of all three normalized responses centered about 0G. As illustrated in the latter figure, the line width of the low-field spectrum is significantly narrower than that obtained at X-band frequencies, despite utilizing the same modulation amplitude for each measurement. The reduction in line width allows the inner hyperfine interactions to be more clearly resolved and therefore can more accurately be analyzed. This also makes the task of finding the relative integrated intensities of the side peaks and center line more manageable. As illustrated in the figure, the amplitude of the dominating hyperfine side peaks only appears to be smaller in the spectra acquired at 215 MHz, when in fact, they are actually just less buried by the center line.
Figure 7.7: Advantages of low-field SDR. (a) Low- and zero-field trace of SDR in 4H SiC MOSFET and (b) zoomed-in plot of the resonant signal. Note that at low fields, the isotropic side peaks maintain their splitting, however, they are shifted toward 0G. Note that the apparent hyperfine peaks that are separated by 68 G are not consistent with the low field shift predicted by the Breit Rabi correction. Because the lower peak is located at a field at half the resonant field, it may actually correspond to a forbidden transition that is sometimes observed at low magnetic fields when dipolar interactions are present [101]. (c) Comparison of zero-, low-, and high-field (shifted from 76G and 3395G respectively) spectra and (d) comparison of all three normalized responses centered about 0G.

The amplitudes of these low-field hyperfine side peaks are more consistent to the amplitude of the inner side peaks of the silicon vacancy discussed in the previous section. (Note that the X-band spectrum illustrated here was not utilized to identify the silicon vacancy defect as discussed in the previous section. It was actually the 2nd harmonic detection of the fast passage signal that enabled this discovery. As mentioned previously, the fast passage detection method is capable of removing artifacts and background signals which enables the inner side peaks to become better resolved.) It is clear from the previous figure that the line width of the SDR spectrum obtained is significantly dependent on the radiation frequency applied during experiment.
Figure 7.8: Comparison of SDR via EDMR and zero-field detection. Here, (a) illustrates the first derivatives of these spectra while (b) and (c) represent the derivative of the spectra. Note that the inflections indicated by the green arrows suggest the presence of exchange or dipolar interactions of the electrons in the spin pair.

Though the zero-field response is significantly broader than that of either the low- or high-field traces, it possesses much information about the hyperfine interactions. Consider the comparison of the zero-field and X-band spectra illustrated in figure 7.8. Similar to the BJT, the hyperfine interactions in the zero-field spectrum appear to be shifted to higher fields. For example, in 7.8 (a), the inner side peaks separated by 12G in the X-band spectrum appear to be shifted to 20G in the zero-field spectrum. As mentioned earlier, we speculate that at very low fields, the nuclei will tend reorient themselves along the field lines produced by the electron, thereby modifying the field that the magnetic nucleus exerts on the electron. As a result, the hyperfine peaks will shift at low externally applied fields. Figures 7.8 (b) and (c) illustrates the derivative of this spectra. Note the inflections just off of the center line indicated by the green arrows. These inflections are most likely due to precise mixing of singlet and triplet states off of zero magnetic field from the exchange and dipolar interactions of the electrons in the spin pair.
7.2.2 – Zero-Field Detection of SDR using SDCP

Charge pumping is a standard MOSFET electronic characterization technique which has a well-established capability to measure oxide/semiconductor interface traps [102] [103] [104]. Recently, the power of electrically detected magnetic resonance and charge pumping were combined to become what is known as spin-dependent charge pumping (SDCP) [105]. The spin dependence of charge pumping can be observed in a transistor by performing a continuous wave EDMR measurement with a modulated bias applied to the gate of a MOSFET as illustrated in the inset of figure 7.9 (b). It is shown in this section that zero-field SDCP can be observed without application of external electromagnetic radiation.

We performed SDCP on a device that was identical to the one used in the previous study. Once again, This device was characterized as having a mobility of $5.7 \text{ cm}^2/\text{V} \cdot \text{s}$, had a thermal ONO anneal, had a gate area of $200 \times 200 \ \mu\text{m}^2$, and the source and drain regions were doped with nitrogen. Figure 7.9 (a) illustrates the response we observe when the external magnetic field is swept across zero Gauss without application of electromagnetic radiation and figure 7.9 (b) illustrates its derivative. The inset of the latter figure illustrates the biasing scheme used for the measurement; the drain and source to substrate junctions were grounded and a trapezoidal waveform (250 kHz frequency, $20V_{pp}$ amplitude, -10V base level, and 500 ns rise/fall time) was applied to the gate. The figure illustrates a dominant signal at zero Gauss with broad side peaks separated by about 25 Gauss almost identical to the zero-field response obtained with the DCIV measurement in the previous study. The difference between the two plots is that there is a dramatic inflection at the zero crossing in the zero-field response. A plausible explanation for this observation is that the pattern is due to the exchange and dipolar coupling of the electrons in the spin pair. This effect is observed in the charge pumping measurement and not the DCIV measurement because the oscillating
voltage applied to the gate in the SDCP experiment was large enough to increasing the coupling of the electrons in the spin pair. This effect may be observed however in the DCIV measurements if the two bias voltages are increased. As mentioned previously, increasing the bias voltage brings out this effect in the BJTs.

7.2.3 – Zero Field Detection of SDR using BAE

EDMR based studies in MOSFETs have typically been performed using the biasing technique of DCIV. In the earlier SiC studies, this biasing technique was sufficient for defect detection of defects at the SiC/SiO₂ interface because in these earlier devices, interface defects dominated the current. However, there have been significant improvements to this interface which have resulted in a far lower interface defect density without a corresponding improvement in the bulk defect density, that is defects which extend more than a few nanometers below the interface. As this interface becomes better understood, advancements in processing have led to the reduction of interfacial defects. As a result, in the higher quality devices, the conventional biasing technique of DCIV has become ineffective in detecting defects at the interface. By utilizing a modified biasing technique, which is based upon the bipolar amplification effect (BAE), it is possible to detect only interface and very near interface defects with EDMR. In this section, it is demonstrated that the zero-field response can also be detected using the biasing scheme of BAE. In the BAE scheme, a forward bias is applied to the drain-to-substrate (or source-to-substrate) junction while applying a sub-threshold bias to the gate and measuring the current at the source (or drain if biasing the source). Under these biasing conditions, the response is only sensitive to the immediate vicinity of the SiC/SiO₂ interface. A more detailed discussion can be found elsewhere [106]. For each of the measurements, we biased the drain-to-substrate junction and gate at voltages which provided the maximum EDMR and zero-field responses.

We applied the BAE technique to a better quality 4H SiC MOSFET that was fabricated by Cree in 2007. This device was a lateral 4H SiC n-channel MOSFETs with gate areas of 100 x 5 μm². A device dependent voltage (1.5 volts) is applied to the gate and the drain-to-substrate junction is forward biased (V_d = -3V) while monitoring current at the source. Figure 7.10 illustrates the low-field SDR response as the external magnetic field was swept from -125G to +125G. The figure illustrates the two resonant signals (220MHz, ±79G) along with a zero-field response. Similar to the study of the older MOSFET using the biasing technique of DCIV, the amplitude of the zero-field response is dependent upon the amplitude of the oscillating magnetic field. The ratio of the fully saturated resonant signal to the zero-field signal (without application of RF radiation) is about 1.78.

Figure 7.11 compares this response to the X-band spectrum obtained on the same device. As illustrated in the figure and all of the previous studies, the zero-field response is much broader than that of
the X-band spectrum. Also, as in the other two cases, the large hyperfine interactions appear to exhibit a larger splitting; a result that we attribute to the reorientation of electron spin and nuclear spin quantization axes. However, as was the case in the other measurements, the interactions are quite similar. Note that when the derivative of the zero-field spectrum is taken, additional structure can be observed as illustrated in figure 7.11. Although all of these interactions cannot all be accounted for, they are repeatable and therefore real interactions that provide additional information in which the X-band data cannot relay. Note also the inflections just off of the center line indicated by the green arrows. These inflections were observed in the zero-field response in the older generation device and are most likely due to precise mixing of singlet and triplet states at very small magnetic fields due to the exchange and dipolar interactions of the electrons in the spin pair.

Figure 7.10: Low-field detection of SDR using the biasing scheme of BAE via EDMR and zero-field detection of SDR. (a) 3D mesh and (b) 2D traces of representative EDMR scans on a SiC MOSFET for a series of $B_1$ amplitudes. (c) Amplitude of the zero-field and low-field resonant signals as a function of $B_1$. The EDMR amplitudes increase monotonically with increasing $B_1$. Note that the increase in the $B_1$ field saturates the resonant SDR response and also affects the amplitude of the zero-field response like the case for the other MOSFET investigated. The inset of this figure illustrates the technique of biasing scheme of BAE. (d) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied.
So, this work demonstrates that this zero-field mixing phenomenon could be used to study SDR in fully processed MOSFETs, BJTs, and diodes. It was also demonstrated that the phenomenon could be detected with the biasing techniques of DCIV, BAE, and charge pumping. In the next section, we demonstrate that this effect can also be observed for spin dependent tunneling in capacitor type structures.

![Graph showing comparison of zero-field and X-band spectra on a 4H SiC MOSFET using the bipolar amplification effect. The derivative of this spectrum is illustrated in figure (b) and (c).](image)

**Figure 7.11:** (a) Comparison of zero-field and X-band spectra on a 4H SiC MOSFET using the bipolar amplification effect. The derivative of this spectrum is illustrated in figure (b) and (c).

### 7.3 – Zero and Low-Field Transport in Amorphous SiC Capacitors

Back end of line dielectrics with low dielectric constants are needed for current and future integrated circuit technology. However, an understanding of the defects that cause leakage currents and limit reliability in these films is not yet developed. In this study, zero-field spin dependent tunnelling (SDT) is investigated in SiC:H dielectrics which have great potential use for back end of line dielectrics. More specifically, we demonstrate that zero-field mixing of
singlet and triplet pair states that form from neighbouring defects in a dielectric, alters the rate at which electrons tunnel from defect to defect through the dielectric.

SDT via EDMR is a very useful measurement which helps to identify defects in dielectrics of fully processed devices and has also recently been shown to be capable of resolving defect energy levels [11]. In this study, it is demonstrated that zero-field detection of SDT in a dielectric can also provide much information about defects involved in trap assisted tunnelling. SDT makes it possible to directly detect defects involved in the trap assisted tunnelling which cause dielectric leakage current and are likely related to important reliability problems such as time dependent dielectric breakdown (TDDB). SDT takes place in these films because the tunnelling process involves defects with unpaired electrons. Tunneling transport between such defects is spin dependent, thus allowing magnetic resonance at the defects to influence the currents. An electron is able to tunnel through a large bandgap dielectric material by hopping from one defect to another if the tunneling electron conserves angular momentum. Similar to SDR, angular momentum is conserved if singlet pairs are formed between the electrons in neighboring orbitals. The application of electromagnetic radiation flips the spin of the unpaired electrons within the dielectric which increases the chance of a singlet pairing. As a result, the trap assisted tunneling current that is measured in an EDMR measurement is spin dependent. This process is illustrated in figure 7.12.

In this section, we demonstrate that SDT can be detected via the singlet triplet mixing phenomenon in a 50Å amorphous hydrogenated SiC (SiC:H) capacitor with a titanium gate and a silicon substrate. (A similar zero-field effect has been observed in tunneling of electrons in double quantum dot

Figure 7.12: Because tunneling also obeys the law of conservation of angular momentum, the formation of singlet pairs in neighboring defects can be used describe spin dependent tunneling of electrons through a dielectric. (a) An electron will not tunnel if the two coupled electrons do not form a singlet pair. (b) With the application of electromagnetic radiation, the electrons are able to flip their spin state which enhances the chance of a singlet pairing and allows the conduction electron to tunnel into a trap. (c) The trapped electron will only be allowed to tunnel to neighboring traps if the tunneling process continues to conserve angular momentum.

In this section, we demonstrate that SDT can be detected via the singlet triplet mixing phenomenon in a 50Å amorphous hydrogenated SiC (SiC:H) capacitor with a titanium gate and a silicon substrate. (A similar zero-field effect has been observed in tunneling of electrons in double quantum dot
structures [23].) Figure 7.13 illustrates the zero- and low-field SDT results obtained when the capacitor structure was subjected to RF radiation, of frequency $\nu = 220 \text{ MHz}$, the magnetic field that was swept from -125G to +125G, and the modulation amplitude was approximately 6 G. From the figure, it is clear that SDT can be observed at the resonant fields ($\pm 79G, 220\text{MHz}$) and at zero-magnetic field. (Note that because the devices we used to observe SDT were composed of an amorphous dielectric, a broad featureless line shape is observed since the randomly oriented defects yield a broad range of hyperfine responses.) As illustrated in the figure, the response of the zero-field signal is barely affected by the application of the RF radiation. This effect is similar to the response of the BJT, however not in the MOSFETs. Unlike the SDR cases, the zero-field SDT signal is quite a bit larger (by a factor of about 1.5) than the fully saturated resonant signals as illustrated in figure 7.13 (c).

Figure 7.13: Low-field and zero-field detection of SDT. (a) 3D mesh and (b) 2D traces of representative EDMR scans on a SiC MOSFET for a series of $B_1$ amplitudes. (c) Amplitude of the zero-field and low-field resonant signals as a function of $B_1$. The EDMR amplitudes increase monotonically with increasing $B_1$. Note that the increase in the $B_1$ field saturates the resonant SDT response but hardly affects the amplitude of the zero-field response. (d) Comparison of low-field scans acquired with (top, blue) and without (red, bottom) RF radiation applied.
Also note the significantly larger currents (100's of pA) that are measured in the SDT as compared to the SDR measurements (10's of pA). In the SDR case, the amplitude is representative of the number of defects responsible for recombination whereas the amplitude of the tunneling spectrum is related to the number of electrons that are able to tunnel through the dielectric through defect assistance. The latter measure will not necessarily provide information regarding the number of defects, but rather the number of defect paths that the electrons can use to spin dependently hop through the dielectric.

Another interesting feature observed in the zero-field signal that is not apparent in the resonant signals is the presence of the inflection precisely at the zero crossing. In the next set of experiments, the modulation amplitude was reduced to in order to further investigate this phenomenon. Figure 7.14 compares (a) the zero-field SDT response, the X-band response, and (b) their derivatives when utilizing a modulation amplitude of 1G. Similar to the SDR measurements, it is apparent that the zero-field response is significantly broader than the spectrum acquired at X-band. This is likely due to the gradual modification of the singlet to triplet ratio resulting from the gradual reorientation of the electron’s axis of quantization. The dramatic inflection at the zero crossing is likely due to the strong exchange and dipolar interactions of electrons within neighboring defects. This strong coupling suggests that the neighboring defects must be relatively close to each other such that there may be significant orbital overlap [46]. This suggests that this technique may be beneficial for studying the coupling and distance between neighboring defects.

Figure 7.14: Comparison of the SDT results obtained at zero-field and X-band. (a) Zero-field signal plotted against the X-band spectrum centered at zero Gauss and (b) its derivative. Note that the large inflection point at zero-field appears as a double peak when the derivative is taken indicates that there is very strong electron-electron coupling. These interactions are most likely due to the exchange or dipolar interactions of electrons in neighboring defects.
In the next set of measurements, the bias applied to the gate of the capacitor was increased to 2, 2.5, 3, 3.5, and 4 volts. Figure 7.15 (a) illustrates the DC leakage current $I$ and amplitude of the zero-field response $\Delta I_{ZF}$ as a function of applied bias and figure 6.20 (b) illustrates the DC leakage current $I$ and amplitude of the resonant field response $\Delta I_{RF}$ as a function of applied bias. Figures 7.15 (c) and (d) illustrate the change in current over the resonant and zero fields. (As before, the change in current was simply calculated by taking the ratio of zero-field and resonant amplitudes and the DC tunneling current). Note that the percent change in tunneling current for both the resonant and zero-field signals are not as large as the SDR case; however, they are still relatively large. In the structure tested, there is a chance that recombination may be playing a role in the observed response. However, the bias dependence illustrated in figure 7.15 are inconsistent with recombination and therefore can most likely be ruled out.

Figure 7.15: Illustration of the size of the resonant and zero-field phenomena. (a) Amplitude of zero-field SDT $\Delta I_{ZF}$ and DC leakage current $I$ as a function of applied bias (b) Amplitude of resonant field SDT $\Delta I_{RF}$ and DC leakage current $I$ as a function of applied bias. (c) Change in current ($\Delta I_{ZF}/I$) over zero-field as a function bias. (d) Change in current over a resonant field ($\Delta I_{RF}/I$) as a function of bias. Note that these curves are inconsistent with recombination which suggests that the zero- and low-field responses are due tunneling.
7.4 – Zero-Field SDR and SDT in Microelectronic Reliability

It has already been demonstrated that the zero-field detection of SDR and SDT can provide almost as much information about defects and spin dependent transport as the magnetic resonant detection. This is so because the electron nuclear hyperfine interactions are the most useful parameter for defect identification which can be detected, in most cases, with the zero-field measurement. Therefore, the tool of zero-field spectroscopy offers an opportunity to extend the power of magnetic resonance to many researchers in microelectronic reliability. For example, it may be used to study dependent dielectric breakdown (TDDB) in dielectrics and bias temperature instabilities in MOSFETs.

7.4.1 – Time Dependent Dielectric Breakdown in Dielectrics

Figures 7.16 and 7.17 illustrate the zero-/low-field SDT results obtained when performing an in-situ bias stress on a 50Å amorphous hydrogenated SiC (SiC:H) capacitor at room temperature. This experiment consisted of nearly 2000 scans performed in the dark that spanned a 40 hour period. The measurement utilized a “fresh” device which was biased with a 4.5V digital power source applied to the gate. (The digital power source was used to prevent the applied voltage from fluctuating during the experiment.) Figure 7.16 illustrates (a) the final average after the experiment was performed and (b), (c) illustrate the derivative of this spectrum. Note the small peaks located at ±40 G (exactly half of the resonant field). These are likely the EPR half-field transitions that are observed from the dipolar interactions of neighboring defects at low fields [101]. The mere presence of this signal suggests the use of this physics to provide a rough estimate of the average distance between defects within a dielectric [101]. This can be useful for those studying spin dependent hopping in dielectrics. However, the most interesting result that came from this experiment was the realization that the amplitude of both the resonant and zero-field SDT signals were increasing as a function of time. This implies that the number of defects was increasing as a function of stress time; a result that is commonly attributed to TDDB.

Figure 7.17 illustrates the measured amplitude of the zero-field and resonant responses as a function of time when subjected to a high voltage bias (4.5V) at room temperature. Here, the running average in plots (b)-(c) is simply the average from the 1st scan to the last scan whereas the windowed average represents the average of the five most recent scans as a function of time. Note that the amplitudes of the zero-field and resonant signals behave similarly and are monotonically increasing as a function of time. This first suggests that the zero-field and resonant signals are related to the same phenomenon. It also suggests that more defects are being created during stress time. This allows more percolation paths to be formed for the electrons to tunnel through the dielectric spin dependently. It is important that one can track the amplitude as a function of time because one is able to follow the time dependent current changes likely involved in TDDB.
Figure 7.16: (a) Resonant (79G, 220MHz) spectrum obtained using a high stressing bias of 4.5 V and a
modulation of 6G. Figure (b) illustrates the derivative and (c) illustrates the zoomed in version of the
derivative. The two symmetrically spaced peaks indicated by the arrows are most likely attributed to the
EPR half-field transitions that arise in the presence of dipolar interactions of neighboring defects.

This simple experiment only touches on the many possibilities that electrical detected zero-field
spectroscopy has in the world of microelectronic reliability. It was mentioned earlier that SDT via
resonance can be used to map out energy levels in thin film dielectrics [11]. Similar results can be
obtained at zero-field via the singlet triplet mixing because it was shown that it behaves similarly to the
resonant response. This is advantageous because significantly less expensive hardware is needed to
acquire similar data. It also may provide some insight in regard to the separation defects within the
dielectric that form the percolation path in which the electrons can spin dependently tunnel. Because we
observed hints of exchange and dipolar interactions in the zero-field responses, this tool may allow one to
extract the number and spacing of defects in thin film dielectrics. And finally, this technique can also be
used to study processing variations in various devices.
Figure 7.17: Results of high bias stress on an amorphous SiC capacitor while performing in-situ EDMR and zero-field detection. This figure illustrates (a) final spectrum obtained after experiment, (b) amplitude of zero-field SDT signal as a function of time, (c) amplitude of -80G resonant signal as a function of time, and (d) amplitude of +80G resonant signal as a function of time. (Note the running average in plots (b)-(c) is simply the average from the 1st scan to the last scan whereas the windowed average represents the average of the five most recent scans as a function of time.)

7.4.2 – Bias Temperature Instability in Silicon MOSFETs

Probably the most recent and interesting realization that came of this work is that zero-field SDR can be detected in Si MOSFETs. This is important because silicon makes up the majority of solid state devices used today. Even though Si MOSFETs are fabricated with very few atomic scale defects, many can be created during the lifetime of a device due to the increased biases and temperatures it experiences during operation. This phenomenon is commonly known as the bias temperature instability (BTI) and was first observed in 1967 by Deal et al [107]. Although many have tried to explain it, the detailed mechanisms of this process are still unknown [5] [108] [109]. At this time, it is clear that the elevated bias and temperature stressing generates dangling bond defects at the Si/SiO₂ interface, commonly referred to as the \( P_b^0 \) and \( P_b^1 \) defects. As illustrated in figure 7.18, these centers involve a dangling bond on a Si
atom which is back-bonded to three other silicon atoms precisely at the interface [110] [111] [112].
Fitzgerald and Grove showed that the interface state defect density $D_{it}$ can be measured by configuring a MOSFET as a gated controlled diode. When the source/drain to substrate junction of the gated diode is forward biased, the interface state density is directly proportional to the change in the substrate current [98] [100] [113]. Therefore, when the gate voltage is swept, recombination will be maximized when equal number of holes and electrons are present at the interface. This results in a peak in the substrate current $\Delta I_{SUB}$. This change in current can be used to calculate the interface defect density $D_{it}$ [98] [100] [113],

$$D_{it} = \frac{2\Delta I_{SUB}}{A_G q n_i \sigma v_{th} V_{DS}} \exp \left( -\frac{q V_F}{2kT} \right) \left( \frac{1}{eV \cdot cm^2} \right)$$

(7.4.1)

Here, $A_G$ is the gate area ($1 \times 10^{-2} \text{ cm}^2$), $n_i$ is the intrinsic number of carriers ($1.0 \times 10^{10} \text{ cm}^3$), $\sigma$ mean capture cross section of the defect ($4 \times 10^{16} \text{ cm}^2$), $v_{th}$ is the thermal velocity of the charge carriers ($1.08 \times 10^7 \text{ cm/s}$), and $V_{DS}$ is the source/drain to substrate forward biased junction voltage. We applied the zero-field technique to 48 nm Si power pMOSFETs with gate areas of $1 \times 10^6 \mu \text{m}^2$ (more information on these devices is provided in the experimental section). The zero-field effect (magnetoresistance) is not commonly observed in Si MOSFETs because there are very few defects present in unstressed devices. In this study, we show once again that the zero field response is directly due to defects responsible for recombination. We demonstrate that with elevated bias temperature stressing, the zero-field SDR can be readily be observed and can be used to study negative BTI (NBTI) phenomenon in Si pMOSFETs.

Figure 7.18: Illustrations of $P_b_0$ and $P_b_1$ defects. These defects are created at the Si/SiO$_2$ interface as a result of elevated bias and temperature stressing of a MOSFET.
We stressed the Si device by applying -25V to the gate for 10,000 seconds at temperature of 150°C. Figure 7.19 illustrates the DCIV and EDMR measurements for both pre and post stressed devices for $V_{DS} = 0.33V$. Figure 7.19(a) illustrates the DCIV measurement for both pre and post stressed devices with $V_{DS} = +0.33V$. In both DCIV measurements, two peaks are observed because the geometry of the device allows for equal number of holes and electrons to gather at the interface for different gate voltages. Because the gate not only overhangs the n-type body, but also slightly overhangs the p-type source/drain regions, two different gate voltages will be required for the collection of equal number of holes and electrons at each interface region due to the different doping densities involved.

Using (7.4.1), the $D_p$ for the unstressed device was calculated to be $0.6 \times 10^9 \text{eV}^{-1} \text{cm}^{-2}$ for the peak at $V_G = -0.3V$ and $2.6 \times 10^9 \text{eV}^{-1} \text{cm}^{-2}$ for the peak at $V_G = +1.8V$. The $D_p$ for the stressed device was calculated to be $7.37 \times 10^{10} \text{eV}^{-1} \text{cm}^{-2}$ for the peak at $V_G = -0.3V$ and $7.11 \times 10^{10} \text{eV}^{-1} \text{cm}^{-2}$ for the peak at $V_G = 1.5V$. Figure 7.19(b) illustrates the low-field EDMR ($\nu = 209 \text{MHz}$) measurements made on this device pre and post stress. It is evident from the signals present in the figure that the stressing conditions have generated defects at the Si/SiO$_2$ interface. The two signals that appear at $B \approx 75 \text{G}$ are associated with the expected resonance detection of SDR and the signal detected at $B = 0 \text{G}$ is associated with the zero-field detection of SDR described in the text above. The spectra obtained in figure 7.19(b) were both taken with the device biased for maximum recombination corresponding to the left peak ($V_G = -0.3 \text{V}$) in the DCIV measurement.

![Figure 7.19](image)

Figure 7.19: (a) Comparison of the DCIV measurements (bottom, red) pre and (top, blue) post stress. The mild stressing conditions involved applying a -25V to the gate of the MOSFET for 10,000 seconds at 150°C. The increase in the amplitudes of the two peaks indicates that a significant number of defects were created at the Si/SiO$_2$ interface. (b) Comparison of low-field measurements made (bottom, red) pre and (top, blue) post stress. (The pre stress spectrum is offset by 4pA for clarity.) As illustrated, the spectrum obtained from the stressed device reveals a zero-field and two resonant signals which are present due to the generation of defects at the Si/SiO$_2$ interface as a result of the negative bias temperature stress.
The inflection point that is observed at precisely zero Gauss disappears when the gate is biased for maximum recombination corresponding to the right peak \((V_g = +1.5V)\). This bias dependent phenomenon suggests the presence of exchange interactions. Figures 7.20(a) and 7.20(b) illustrate the zero- and low-field responses of the Si MOSFET biased with a gate voltage of +1.5V and -0.3V (voltages which correspond to the peaks in the DCIV measurement) respectively, with and without application of a 305MHz oscillating EM field. As illustrated in the both figures, the zero-field signal is independent of the applied EM radiation. Also, the zero-field signals are significantly broader than the resonant signals, a feature observed before and most likely attributed to the gradually changing singlet triplet ratio that occurs when the magnetic field swept towards zero. The difference between the two spectra is that the one obtained with a gate voltage of -0.3V has a significant inflection point located at precisely at zero Gauss. This electric field induced phenomenon is likely attributed to the exchange interactions of the two electrons involved in the precursor pair formation prior to recombination. The utilization of magnetic field modulation allows for the observation of this small interaction.

Figure 7.21 illustrates the zero-field and low-field \((B_0 \approx 75G, \nu = 209MHz)\) amplitude responses as a function of applied oscillating \(B_1\) magnetic field when the MOSFET was biased with \(V_g = +1.5V\) and \(V_{DS} = +0.32V\). As expected, the resonant SDR signal ceases to exist when the EM field is turned off and saturates when the oscillating \(B_1\) field is large enough to randomize the spin pair’s relative orientation. The zero-field SDR signal on the other hand is completely unaffected by this perturbing field suggesting it to be an inexpensive alternative to EDMR as a means to study spin dependent transport in microelectronics.

Figure 7.20: (a) Comparison of measurements taken (top, blue) with and (bottom, red) without electromagnetic radiation when biased with +1.50V. (b) Comparison of measurements taken (top, blue) with and (bottom, red) without electromagnetic radiation when biased with -0.30V. The significant inflection point observed precisely at zero Gauss is likely due to the exchange interaction the electrons experience when an appropriate electric field bias is applied to the gate of the MOSFET.
Figure 7.21: Comparison of the zero-field and low-field ($B_0 = 75G, \nu = 209MHz$) amplitude responses as a function of applied oscillating $B_1$ magnetic field when the MOSFET was biased with $V_g = +1.5V$ and $V_{DS} = +0.32V$. As expected, the resonant SDR signal saturates when maximum power is applied and ceases to exist when no power is applied. The zero-field SDR signal is completely unaffected by this perturbing field.

However, it remains to be shown that the zero-field phenomenon is due to recombination. Figure 7.22 demonstrates this proof. The figure illustrates the amplitude of the zero-field and low-field signals as a function of applied bias plotted against the current measured with the DCIV method. As expected, the amplitude of the resonant signal is well correlated with the response of the DCIV measurement. This is how one conventionally demonstrates the observation of SDR via EDMR in a MOSFET. The interesting result here is that the zero-field response is also highly correlated with the DCIV and resonant responses. This is direct evidence that the zero-field signal and resonant signals are attributed to the same mechanisms; that being defects that cause SDR. Figure 7.22(b) illustrates a three dimensional mesh plot of the integrated data used to plot the amplitude of the zero-field response in figure 7.22(a). The bias dependent inflection point that arises precisely at zero magnetic field (also illustrated in the previous figures) leads us to believe that the gate voltage plays a role in the coupling of electrons in the spin pair that form prior to recombination. The presence of this inflection at certain gate voltages may alter the dipolar or exchange coupling of the two electrons in the spin pair. As a result, zero-field detection of SDR is capable of providing additional information that of the conventional detection of SDR via EDMR.
Figure 7.22: (a) Illustration of the zero-field and low-field amplitude responses as a function of applied bias plotted against the current measured in the DCIV method. Note that the zero-field response is highly correlated with the low-field and DCIV responses. This is direct evidence that the zero-field signal is due to spin dependent recombination at the Si/SiO₂ interface of the MOSFET. (b) Three dimensional mesh plot of the integrated zero-field spectra used to plot the amplitude response in (a). Note that as the voltage is decreased from $V_G = +1.5 \text{ V}$, a double peak begins to form which is most likely attributed to electron exchange coupling.

We have performed EDMR on the same pMOSFET device at relatively high magnetic fields ($B_0 \approx 3285$, $\nu = 9.228 \text{ GHz}$) and compared the results to that of the zero- and low-field spectra as illustrated in figure 7.23. Using a 2G modulation, the high-field trace reveals two different overlapping signals which could not be detected in the zero- and low-field measurements because of the reduction in resolution that occurs when making measurements at lower fields. The $g = 2.0065$ and $g = 2.0033$ are due to the $Pb_0$ and $Pb_1$ defects that are almost always created when Si pMOSFETs are exposed to negative biases at elevated temperatures for a prolonged period of time. In the low-field signal, the three defects are separated in field by an amount which cannot be resolved. In the zero-field signal, all three defects collapse onto one another and are all located precisely at zero Gauss. Therefore, the amplitude of this signal is representative of the cumulative response from all defects present within a device making the technique not capable of being able to discern between defects with different $g$ values. However, the fact that it has the capability of detecting defects responsible for SDR with associated hyperfine interactions and electron exchange interactions makes it a significantly less expensive and viable alternative to defect detection technique of EDMR.

These initial studies of TDDB in thin film dielectrics and NBTI in Si pMOSFETs only hint at the potential uses of this zero-field spectroscopy method in microelectronic reliability. Not only can this emerging physics be used in microelectronic research, but zero-field transport has much potential in the design of magnetometers and miniature zero- and low-field spectrometers.
Figure 7.23: Comparison of (top) high-field ($B_0 \approx 3287 \, G, \nu = 9.228 \, GHz$), (middle), low-field ($B_0 \approx 110 \, G, \nu = 305 \, MHz$), and (bottom) zero-field signal (no applied radiation). Note that the two signals observed at high-field are unresolvable in the zero- and low-field measurements because of the reduction in resolution that occurs when making measurements at low magnetic fields.
CHAPTER 8 – APPLICATIONS OF ZERO-FIELD TRANSPORT

The change in recombination current at zero field for all the 4H SiC devices is so large that it has many potential applications. One particular application that stands out among others is a sensor that provides absolute magnetometry (with directional polarity) which utilizes lock-in based detection when used with a precision current controlled magnetic field sweep. It has already been argued that somewhat similar phenomenon involved in low-field EDMR in organic devices may be useful for absolute magnetometry [114]. However, utilizing the ZFSDR phenomenon in SiC devices has many advantages over organic devices. Unlike organic devices, SiC devices are inherently quite robust, and capable of very long term operation in challenging environments including high temperatures. Perhaps more importantly, SiC devices utilize a single crystal which allows for the detection of multiple sets of quite stable hyperfine interactions which can be used for magnetic field calibration without the requirement of RF components. For example, the 3 pairs of symmetrical peaks illustrated in Fig 3(d) would serve as stable magnetic field markers as they would not move as a result of changes in biasing condition and or temperatures. Essentially, the magnetometer would be self-calibrating. It may also be useful to note that it has been suggested that SiC also has great potential in quantum computation for several reasons: its large bandgap allows for multiple defect-induced states [115] [116]; SiC defects are known to have long spin relaxation times [92] [3]; and the defects observed in this study involve sites with multiple nuclear spins (almost certainly \(^{29}\)Si and \(^{14}\)N) that can act as a universal quantum gate [117] [118]. And finally, because this ZFSDR phenomenon can be observed in multiple solid state electronic components including MOSFETs, BJTs, diodes, and capacitors, it would be useful for semiconducting manufacturing companies to incorporate simple automated low-field/zero-field EDMR spectrometers into wafer fabrication/probing equipment to study the defects in solid-state electronics during fabrication. Because only very low fields are required, low-field EDMR and ZFSDR can be performed easily and inexpensively.

The technology developed herein is used for two inventions and is based upon the same underlying physics: an absolute magnetic field sensor and as a miniature low-field electrically detected magnetic resonance (EDMR) spectrometer for wafer probing stations. The inventions provide a robust measurement of magnetic fields and a means for extremely sensitive measurement of spin dependent transport mechanisms and the identification of defects within fully processed micro- and nano-electronic devices. As discussed previously in the thesis, We have discovered that certain spin dependent transport mechanisms including recombination and tunneling can be detected within micro- and nano-electronics in the absence of an oscillating magnetic field while sweeping a small external quasi-static magnetic field across zero field. This behavior is unexpected from the conventional theory of EDMR. There is quite significant potential for use of this new physics in many applications.
Conventional electron paramagnetic resonance (EPR) spectrometers range in cost from several hundred thousand dollars to about one million dollars for state of the art systems. Considerably less sensitive and much less versatile systems can be purchased for as little as $25,000. EDMR spectrometers can be built with modest modification of the EPR spectrometers. EDMR spectrometers offer multiple advantages over conventional EPR systems in applications to solid state electronics. They offer much higher sensitivity and a sensitivity limited exclusively to imperfections which play a role in the electronic behavior of the devices under study. The inexpensive EPR spectrometers would be very difficult to modify for EDMR. The space and power requirements for the more expensive EPR spectrometers are considerable, with typical systems utilizing power supplies of several kilowatts and chilled water heat exchangers typically requiring footprints of ten or more square feet.

Although conventional EPR and EDMR are quite powerful analytical tools for the evaluation of materials physics problems in solid-state electronic devices, the measurements are quite time consuming and generally require extensive sample preparation to allow for insertion of samples into specialized microwave resonant cavities. The invention described herein offers a much less expensive and much more convenient scheme for EDMR evaluation of semiconducting device technology. The high costs and complexity of the conventional EDMR measurements require very large magnetic fields (typically 0.35 Tesla or higher) which must be exceptionally stable and high frequency electromagnetic radiation (typically 9 GHz or higher). Our invention eliminates the need for the high well-characterized stable magnetic field, microwave radiation, and typically quite expensive (> $10,000 from Bruker) specialized microwave cavities. Without the need for the high field or the microwave resonator, far more straightforward measurements become possible.

The technology described herein can also provide a much more inexpensive, robust, and lightweight alternative to costly (> $10,000) low-field based magnetic resonance magnetometers. A drawback of our alternative may be lower precision, at least in comparison with the highest precision systems. However, the invention described herein would also provide absolute highly repeatable measurements due to the inherent self-calibrating capabilities of the invention and the invention could, at least in principle, offer higher sensitivity over currently available magnetic resonance based magnetometers.

8.1 – Magnetic Field Sensor

The ZFSDR response that we observe in our silicon carbide (SiC) based devices (BJTS, MOSFETs, and capacitors) provide the ideal physics for absolute magnetometry at very low cost with a highly robust device. (Other inorganic material systems other than silicon carbide could also quite likely provide comparable physics. Some of these material systems include silicon; germanium, and CdTe, and
CdS are other likely candidates.) The underlying idea behind the functionality of the sensor is what we call the ZFSDR signal which is strongest at zero field but the response extends with a very repeatable structure up from to zero to small magnetic fields. In the specific device that we use, we observe side peaks due to electron nuclear hyperfine interactions at particular magnetic fields slightly larger than zero. These extremely stable peaks allow us to calibrate the sensor and precisely measure the external magnetic field. As illustrated in figure 1, the sensor is comprised of a SiC based sensing device, a set of Helmholtz coils to provide a sweeping magnetic field (electromagnet), an additional set of Helmholtz coils to provide magnetic field modulation, a device biasing source, and a current-to-voltage preamplifier with analog signal conditioning capabilities. In addition, a precision current controller is implemented for the electromagnet so that the static magnetic field produced can be easily determined. (Note that by precisely controlling the current, the magnetic field produced by the electromagnet is indirectly, but precisely, known. Magnetic field control is not utilized because the controller would compensate for the magnetic field one is attempting to detect.) The remainder of the technology is implemented in software. More specifically, a digital PI controller, lock-in amplification, digital signal processing, and field prediction are implemented by a computer that is connected to the analog circuitry through a serial bus. Note that a digital signal processor / microcontroller could also be used instead of a computer to perform the required ADCs, DACs, control, and processing to make the unit quite portable.

![Diagram of the sensor system](image)

**Figure 8.1:** Illustration of the technology used as a magnetic field sensor.

The sensor works as follows. The software sweeps the current (bidirectional) through the magnet coils which produces a linear varying magnetic field. A waveform generator continuously drives an additional set of Helmholtz coils with a sinusoid which allows for field modulation. The received modulated current signal first is conditioned (current-to-voltage conversion, amplification, and filtering) and then sampled at a high rate through an analog-to-digital converter. This signal is then demodulated in
software. Signal processing algorithms are utilized to precisely calculate the field of the observed response. A shift in the zero field response indicates that an external magnetic field is present as illustrated in figure 8.2. This sensor has many advantages over using a conventional Hall based sensor. Unlike Hall sensors, the measurement is completely independent of temperature and works especially well at lower fields ($\leq 0.1\text{G}$ or $\leq 10^{-2}\text{ mT}$). Also, this sensor is capable of indicating magnetic polarity. Therefore, this sensor can easily be applied in three dimensions which allows the technology to be used as a 3 dimensional magnetic field mapping sensor.

The 1% change in current of the BJT at zero field is a huge effect and has the potential to be incorporated into many other applications. Using a material system such as a SiC would allow for the use of the Si vacancy defect in quantum computing applications. This defect has neighboring magnetic nuclei which would allow multiple qubit states to be realized. Another potential application is a magneto-isolation circuit. This device would function similar to that of an opto-isolating circuit that uses optical components such as the LED and photodiode packages. This would allow control by means of a magnetic field without actually connecting to the circuit.

![Figure 8.2: Illustration of the response of the sensor in the presence and absence of a magnetic field. The magnetic field is easily measured by calculating the shift in the response from zero magnetic field, the ideal response. The measured field is the $B_{\text{offset}}$. In this schematic illustration, $B_{\text{offset}}$ is 15 Gauss. The measured field illustrated was chosen merely to make the underlying principle clear. However, much smaller local fields could and generally would be utilized in the measurement.](image-url)
8.2 – Miniature Low- / Zero-Field Spectrometer

A final application, the most important and also the most obvious, would be a cheap and easy way to study electron nuclear hyperfine interactions and the defect environment of certain material systems. One primary issue that has deterred scientists from purchasing X-band spectrometers is their extremely high costs. High field magnetic resonance requires large expensive magnets, cooling systems, and microwave generation and plumbing components which are typically really expensive. (A Bruker spectrometer is typically about $250,000 to $1,000,000 depending on hardware purchased.) By moving to low field and removing the electromagnetic field generation, one may possibly be able to extract the same information at zero field as can be done at X-band frequencies. Further studies will be needed to show the extent of this capability however.

Semiconductor companies evaluate performance and reliability of the solid-state devices they fabricate, generally utilizing characterization solely based upon electrical measurements. Electrical measurements provide absolutely no information about the physical and chemical nature of the performance limiting defects at the atomic level. Magnetic resonance measurements are capable of providing this information. If the manufacturer wanted to obtain this atomic scale information they could seek outside help from academic laboratories equipped for magnetic resonance. However, the possibly preferable approach of doing this work in house is discouraged by the aforementioned cost and complexity of the currently available apparatus as well as the cost and complexity of the measurements they can provide. Our invention will allow these companies to perform their own research on the devices during and just after fabrication. This will save the company time and money. Our invention will provide very straightforward measurements, though in some cases, at the cost of some loss in analytical power over the conventional non zero field EDMR and EPR.

The circuit previously described for the sensor is actually a miniature EDMR spectrometer without the radio frequency (RF) circuitry to provide the oscillating magnetic field. The apparatus could be used identify the physical and chemical nature of atomic scale defects which limit performance of solid-state devices utilizing low-field and zero-field spin dependent transport observations obtained from the technology. Because the technology is miniaturized, one could envision that it could serve as a standalone miniature spectrometer or easily be added to existing wafer probing stations. As a result, semiconductor manufacturing companies could study these important defects with spin dependent properties possibly during or just after fabrication of wafers and devices. The physical mechanism in which the data is gathered is identical to that of the magnetic field sensor. The changes required are illustrated in Figure 8.3. A slight modification to the electromagnet and modulation coils would allow it to be used with wafer probing stations as well. Figure 8.4 illustrates the current prototype which can be used as desktop zero- and low-field EDMR spectrometer.
Figure 8.3: Illustration as the technology used as a miniature spectrometer for wafer probing stations.

Figure 8.4: Photograph of a desktop zero- and low-field spectrometer.
CHAPTER 9 – SUMMARY AND CONCLUSIONS

In conclusion, this work demonstrated that the spin dependent transport phenomena of recombination and tunneling can be detected at zero magnetic field in Si and SiC based devices in the absence of electromagnetic radiation. This mechanism is most likely due to the mixing of singlet and triplet energy states of the electron spin pairs involved in the respective spin dependent processes. Theoretically, the technique is applicable to devices of all material systems in which defects play a role in spin dependent transport. Some of the up and coming materials systems in which this phenomenon will most likely be observed are CdTe and GaN. This makes the technique optimal for studying atomic scale defects in solar cells, optoelectronics, high power electronics, and nanoelectronics. This technique is capable of extracting hyperfine interactions and allows for possible detection of electron-electron dipolar and exchange interactions. It also has much promise in device reliability studies, as it is directly applicable to time dependent dielectric breakdown in thin film dielectrics and bias temperature instabilities in MOSFETs. Also, the zero-field detection of SDR and SDT have the potential for many applications which include the design of self-calibrating magnetometers, opto-isolation circuits, and miniature zero- and low-field electrically detected spin dependent transport spectrometers for wafer probing stations. We were also able to show that by demodulating higher order harmonics that arise from passage effects in high-field EDMR, one is able to extract the fast passage signal which allows for better resolved defect identification. This technique has allowed for the detection and identification of a very important recombination center in 4H SiC MOSFETs; the silicon vacancy. To the best of our knowledge, this was the first study that utilized passage effects for defect identification in EDMR. And finally, this work demonstrated that the technique of adaptive signal averaging is capable of reducing the noise variance of a single scan by a factor of 10 or more and reduction in acquisition time by the same amount. This technique is applicable to all methods in which signal averaging is utilized, some of which include medical imaging, electrocardiography, or electroencephalography.
APPENDIX A – Acronyms

ALP – Adaptive Linear Prediction
ANC – Adaptive Noise Cancellation
AR - Autoregressive
ASA – Adaptive Signal Averaging
BJT – Bipolar Junction Transistor
BTI – Bias Temperature Instability
CdTe – Cadmium Telluride
CTFT – Continuous Time Fourier Transform
CVD – Chemical Vapor Deposition
DC – Direct Current
DCIV – Direct Current Current Voltage
DFT – Discrete Fourier Transform
DTFT – Discrete Time Fourier Transform
DUT – Device Under Test
EDMR – Electrically Detected Magnetic Res.
EM – Electromagnetic
EMF – Electromotive Force
ENDOR – Electron Nuclear DOuble Resonance
EPR – Electron Paramagnetic Resonance
ESR – Electron Spin Resonance
EWRLS – Exponentially Weighted RLS
FFT – Fast Fourier Transform
FIR – Finite Impulse Response
FPGA – Field Programmable Gate Array
GaN – Gallium Nitride
IDTFT – Inverse Discrete Time Fourier Transform
IID – Independent and Identically Distributed
IIR – Infinite Impulse Response
ISC – Inter System Crossing
KSM – Kaplan, Solomon, and Mott
LIA – Lock-in Amplifier
LMS – Least Mean Squares
LS – Least Squares
MOSFET – Metal Oxide Semiconductor Field
MR – Magnetic Resonance
MSE – Mean Squared Error
NLMS – Normalized Least Mean Squares
NMR – Nuclear Magnetic Resonance
NN – Nearest Neighbor
NO – Nitric Oxide
PCB – Printed Circuit Board
PCI – Peripheral Component Interconnect
PID – Proportional, Integral, Derivative
PLL – Phase Lock Loop
RAM – Random Access Memory
RF – Radio Frequency
RLS – Recursive Least Squares
Si - Silicon
SiC – Silicon Carbide
SiO₂ – Silicon Dioxide
SDT – Spin Dependent Tunneling
SDR – Spin Dependent Recombination
SNR – Signal-to-Noise Ratio
SRH – Shockley-Read-Hall
SSE – Sum of Squared Errors
TDDDB – Time dependent dielectric breakdown
VLIA – Virtual Lock-in Amplifier
ZF – Zero-Field
ZFSDR – Zero-Field SDR
ZFSDT – Zero-Field SDT
### APPENDIX B – Spin Properties of Common Nuclei

<table>
<thead>
<tr>
<th>Z</th>
<th>A</th>
<th>Atom</th>
<th>Spin</th>
<th>NA %</th>
<th>Z</th>
<th>A</th>
<th>Atom</th>
<th>Spin</th>
<th>NA %</th>
<th>Z</th>
<th>A</th>
<th>Atom</th>
<th>Spin</th>
<th>NA %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>1</td>
<td>1/2</td>
<td>98.88</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>2</td>
<td>He</td>
<td>2</td>
<td>0</td>
<td>99.88</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>3</td>
<td>Li</td>
<td>3</td>
<td>1</td>
<td>99.88</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>4</td>
<td>Be</td>
<td>4</td>
<td>0</td>
<td>99.88</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>5</td>
<td>3/2</td>
<td>97.73</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>6</td>
<td>1</td>
<td>97.73</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>7</td>
<td>N</td>
<td>7</td>
<td>1</td>
<td>97.73</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>8</td>
<td>O</td>
<td>8</td>
<td>1</td>
<td>97.73</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>9</td>
<td>1</td>
<td>97.73</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>10</td>
<td>Ne</td>
<td>10</td>
<td>1</td>
<td>97.73</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
<tr>
<td>11</td>
<td>Na</td>
<td>11</td>
<td>1</td>
<td>97.73</td>
<td>32</td>
<td>Ge</td>
<td>36</td>
<td>2.56</td>
<td>53</td>
<td>Ca</td>
<td>40</td>
<td>2.56</td>
<td>74</td>
<td>Lu</td>
</tr>
</tbody>
</table>

---

Source: http://www.easypin.org/
APPENDIX C – Derivation of Projection Operators in Matrix Form

Consider an isolated electron that has spin \( s = \frac{1}{2} \). In the presence of a magnetic field, its angular momentum will be quantized along the axis in which the field is applied. As a result, the electron will align itself with or against this field. These orientations define the two states of the electron, where the projected spin angular momentum is \( m_s = \pm \frac{1}{2} \). In chapter 2, it was shown that the electron spin operator could be represented as a vector. In general, the electron spin operator is more accurately represented as a matrix. The goal of this section is to derive this spin matrix and the spin projection operator for a single electron and then the projection operators.

As noted before, the electron states can be represented by the notation, \( |s, m_s\rangle \), where
\[
|\alpha\rangle = \frac{1}{2}, \frac{1}{2} = |\uparrow\rangle \\
|\beta\rangle = \frac{1}{2}, -\frac{1}{2} = |\downarrow\rangle
\]

where \( s \) and \( m_s \) are the electron spin quantum numbers. From table 2.3, it was shown that for a single electron,
\[
\hat{S}_z |\alpha\rangle = m_s |\alpha\rangle = \frac{1}{2} |\alpha\rangle \\
\hat{S}_z |\beta\rangle = m_s |\beta\rangle = -\frac{1}{2} |\beta\rangle \\
\hat{S}^2 |\alpha\rangle = s(s + 1)|\alpha\rangle = \frac{3}{4} |\alpha\rangle \\
\hat{S}^2 |\beta\rangle = s(s + 1)|\beta\rangle = \frac{3}{4} |\beta\rangle
\]

The states \( |\alpha\rangle \) and \( |\beta\rangle \) are considered to be orthogonal and therefore can be represented as basis vectors,
\[
\alpha = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (C.7) \\
\beta = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (C.8)
\]

The electron can be in either the up or down spin states, or in any linear combination of both up and down states. Therefore, the electron spin wave function \( |\psi\rangle \) can be represented as,
\[ |\psi\rangle = c_1 |\alpha\rangle + c_2 |\beta\rangle \quad (C.9) \]

where the constants \( c_1 \) and \( c_2 \) provide information about the state probability, therefore, \( (c_1)^2 + (c_2)^2 = 1 \). (Orthonormal in this case implies that the quantum states, or basis functions, are orthogonal and of unit length. Once again, \(|\cdot\rangle\) is referred to as the ket which represents the wavefunction and \( <\cdot|\) is referred to as the bra which represents the complex conjugate of the wavefunction. This notation was used because it provided a convenient way for representing quantum states as a linear combination of orthogonal basis functions.)

Consider now a spin operator \( \hat{S}_n \) that acts on this generalized electron spin wave function, where \( n \) represents a particular axis, ie: \( n \in \{x, y, z\} \). This results in a new wave function \(|\varphi\rangle\) which is also a linear combination of the two electron spin states with different coefficients,

\[ |\varphi\rangle = \hat{S}_n |\psi\rangle = \hat{S}_n (c_1 |\alpha\rangle + c_2 |\beta\rangle) = c_3 |\alpha\rangle + c_4 |\beta\rangle \quad (C.10) \]

Note that the operator simply modified the coefficients of each electron state which simply implies that the probabilities of the electron being in either the up or down state have changed. The goal now becomes to find these new probabilities \( c_3, c_4 \). To find \( c_3 \), multiply equation (C.10) through by the bra \( <\alpha| \). (This is done in order rid the dependence of the equation on the orthonormal basis functions \(|\alpha\rangle, |\beta\rangle\).) Because the inner products of the orthonormal basis functions are \( <\alpha|\alpha\rangle = 1 \) and \( <\alpha|\beta\rangle = 0 \), \( c_3 \) can be found by,

\[
c_1 <\alpha|\hat{S}_n|\alpha\rangle + c_2 <\alpha|\hat{S}_n|\beta\rangle = c_3 <\alpha|\alpha\rangle + c_4 <\alpha|\beta\rangle
\Rightarrow c_3 = c_1 <\alpha|\hat{S}_n|\alpha\rangle + c_2 <\alpha|\hat{S}_n|\beta\rangle \quad (C.11)
\]

Similarly, to find \( c_4 \), one must multiply through the bra \( <\beta| \) (Note that the inner products \( <\beta|\alpha\rangle = 0 \) and \( <\beta|\beta\rangle = 1 \) because these basis states are orthonormal) to get,

\[
c_1 <\beta|\hat{S}_n|\alpha\rangle + c_2 <\beta|\hat{S}_n|\beta\rangle = c_3 <\beta|\alpha\rangle + c_4 <\beta|\beta\rangle
\Rightarrow c_4 = c_1 <\beta|\hat{S}_n|\alpha\rangle + c_2 <\beta|\hat{S}_n|\beta\rangle \quad (C.12)
\]

These equations can easily be written in matrix format,

\[
\begin{bmatrix} c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} <\alpha|\hat{S}_n|\alpha\rangle & <\alpha|\hat{S}_n|\beta\rangle \\ <\beta|\hat{S}_n|\alpha\rangle & <\beta|\hat{S}_n|\beta\rangle \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad (C.13)
\]
Note that these equations are for an arbitrary axis $n$. Now assume that we define this axis of quantization to be the $z$ axis, i.e: $[0,0,1]$, the direction in which we apply our static magnetic field. Analyzing each component,

$$\langle \alpha | \hat{S}_z | \alpha \rangle = \frac{1}{2} \langle \alpha | \alpha \rangle = \frac{1}{2}$$ \hspace{1cm} (C.14)

$$\langle \alpha | \hat{S}_z | \beta \rangle = -\frac{1}{2} \langle \alpha | \beta \rangle = 0$$ \hspace{1cm} (C.15)

$$\langle \beta | \hat{S}_z | \alpha \rangle = \frac{1}{2} \langle \beta | \alpha \rangle = 0$$ \hspace{1cm} (C.16)

$$\langle \beta | \hat{S}_z | \beta \rangle = -\frac{1}{2} \langle \beta | \beta \rangle = -\frac{1}{2}$$ \hspace{1cm} (C.17)

Rewriting in matrix form yields,

$$\begin{bmatrix} c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & -1/2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{2} \sigma_z \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$ \hspace{1cm} (C.18)

Where $\sigma_z$ is known as the Pauli spin matrix for the spin projection in the $z$ axis. In order to find these matrices for the other two axes, one must use the raising and lowering operators. These operators essentially raise or lower the discrete values of spin angular momentum to the next quantum state. For this reason, these operators are also called ladder operators. The raising and lowering operators are defined respectively as [44],

$$\hat{S}_+ |s,m_s\rangle = \sqrt{(s(s+1) - m_s(m_s + 1))} |s,m_s + 1\rangle$$ \hspace{1cm} (C.19)

$$\hat{S}_- |s,m_s\rangle = \sqrt{(s(s+1) - m_s(m_s - 1))} |s,m_s - 1\rangle$$ \hspace{1cm} (C.20)

These definitions lead to the results below,

$$\hat{S}_+ |\alpha\rangle = 0$$ \hspace{1cm} (C.21)

$$\hat{S}_- |\beta\rangle = 0$$ \hspace{1cm} (C.22)

$$\hat{S}_+ |\beta\rangle = |\alpha\rangle$$ \hspace{1cm} (C.23)

$$\hat{S}_- |\alpha\rangle = |\beta\rangle$$ \hspace{1cm} (C.24)
Note that $\hat{S}_+|\alpha\rangle = 0$ and $\hat{S}_-|\beta\rangle = 0$ because the $\alpha$ state cannot be raised to a spin greater than $\frac{1}{2}$ and $\beta$ cannot be lowered to a spin less than $-\frac{1}{2}$. The ladder operators can also be used to reorient the spin angular momentum with respect to the two axes $x$ and $y$ [40]. The operators are also defined as,

\[
\hat{S}_+ = \hat{S}_x + i\hat{S}_y \\
\hat{S}_- = \hat{S}_x - i\hat{S}_y
\]

where $i$ is an imaginary number. Note that $\hat{S}_-$ is the complex conjugate of $\hat{S}_+$. The spin operators for the $x$ and $y$ directions can then be constructed [40],

\[
\hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-) \\
\hat{S}_y = -\frac{i}{2}(\hat{S}_+ - \hat{S}_-)
\]

Substituting the $\hat{S}_x$ operator in for $\hat{S}_n$ in equation (B.13) yields,

\[
\langle \alpha|\hat{S}_x|\alpha\rangle = \frac{1}{2}(\langle \alpha|\hat{S}_+|\alpha\rangle + \langle \alpha|\hat{S}_-|\alpha\rangle) = \frac{1}{2}\langle \alpha|\beta\rangle = 0 \\
\langle \alpha|\hat{S}_x|\beta\rangle = \frac{1}{2}(\langle \alpha|\hat{S}_+|\beta\rangle + \langle \alpha|\hat{S}_-|\beta\rangle) = \frac{1}{2}\langle \alpha|\alpha\rangle = \frac{1}{2} \\
\langle \beta|\hat{S}_x|\alpha\rangle = \frac{1}{2}(\langle \beta|\hat{S}_+|\alpha\rangle + \langle \beta|\hat{S}_-|\alpha\rangle) = \frac{1}{2}\langle \beta|\beta\rangle = \frac{1}{2} \\
\langle \beta|\hat{S}_x|\beta\rangle = \frac{1}{2}(\langle \beta|\hat{S}_+|\beta\rangle + \langle \beta|\hat{S}_-|\beta\rangle) = \frac{1}{2}\langle \beta|\alpha\rangle = 0
\]

Therefore,

\[
\begin{bmatrix}
  c_3 \\
  c_4
\end{bmatrix} = \begin{bmatrix}
  0 & 1/2 \\
  1/2 & 0
\end{bmatrix} \begin{bmatrix}
  c_1 \\
  c_2
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
  0 & 1 \\
  1 & 0
\end{bmatrix} \begin{bmatrix}
  c_1 \\
  c_2
\end{bmatrix} = \frac{1}{2} \sigma_x \begin{bmatrix}
  c_1 \\
  c_2
\end{bmatrix}
\]

And finally, substituting the $\hat{S}_y$ operator in for $\hat{S}_n$ in equation (B.13) yields,

\[
\langle \alpha|\hat{S}_y|\alpha\rangle = -\frac{i}{2}(\langle \alpha|\hat{S}_+|\alpha\rangle - \langle \alpha|\hat{S}_-|\alpha\rangle) = -\frac{i}{2} < \alpha|\beta >= 0 \\
\langle \alpha|\hat{S}_y|\beta\rangle = -\frac{i}{2}(\langle \alpha|\hat{S}_+|\beta\rangle - \langle \alpha|\hat{S}_-|\beta\rangle) = -\frac{i}{2} < \alpha|\alpha >= -\frac{i}{2}
\]
\begin{align}
\langle \beta | \hat{\mathbf{S}}_y | \alpha \rangle &= -\frac{i}{2} (\langle \beta | \hat{\mathbf{S}}_+ | \alpha \rangle - \langle \beta | \hat{\mathbf{S}}_- | \alpha \rangle) = \frac{i}{2} < \beta | \alpha > = \frac{i}{2} \\
\langle \beta | \hat{\mathbf{S}}_y | \beta \rangle &= -\frac{i}{2} (\langle \beta | \hat{\mathbf{S}}_+ | \beta \rangle - \langle \beta | \hat{\mathbf{S}}_- | \beta \rangle) = -\frac{i}{2} < \beta | \alpha > = 0
\end{align}
(C.36)

Therefore,
\[
\begin{bmatrix}
  c_3 \\
  c_4
\end{bmatrix} = \begin{bmatrix}
  0 & -i/2 \\
  i/2 & 0
\end{bmatrix} \begin{bmatrix}
  c_1 \\
  c_2
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
  0 & -i \\
  i & 0
\end{bmatrix} \begin{bmatrix}
  c_1 \\
  c_2
\end{bmatrix} = \frac{1}{2} \sigma_y \begin{bmatrix}
  c_1 \\
  c_2
\end{bmatrix} \tag{C.38}
\]

So, in summary, the Pauli matrices for an electron (or any spin \(\frac{1}{2}\) particle) are defined as,

\[
\sigma_z = \begin{bmatrix}
  1 & 0 \\
  0 & 1
\end{bmatrix} \tag{C.39}
\]

\[
\sigma_x = \begin{bmatrix}
  0 & 1 \\
  1 & 0
\end{bmatrix} \tag{C.40}
\]

\[
\sigma_y = \begin{bmatrix}
  0 & -i \\
  i & 0
\end{bmatrix} \tag{C.41}
\]

The Pauli matrices span a 2-dimensional complex Hilbert space (can be represented by a Bloch sphere) which is intended to map the probabilities \(c_1, c_2\) to \(c_3, c_4\). They represent the observable of spin in each spatial direction. Note that because we defined the \(z\) axis as our axis of quantization (axes of external magnetic field), the eigenvector of the \(\sigma_z\) matrix is simply the columns of the matrix (a property of a diagonal matrix), ie: the original basis vectors defined by \(|\alpha\rangle\) and \(|\beta\rangle\). Note that the Pauli matrices are related to the spin matrices by,

\[
\sigma_j = 2 \hat{S}_j \tag{C.42}
\]

where

\[
\sigma = \sigma_x + \sigma_y + \sigma_z \tag{C.43}
\]

and the spin operator along an arbitrary axis is defined by,

\[
\hat{\mathbf{S}} = \frac{1}{2} \sigma = \hat{\mathbf{S}}_x + \hat{\mathbf{S}}_y + \hat{\mathbf{S}}_z = \frac{1}{2} \sigma_x + \frac{1}{2} \sigma_y + \frac{1}{2} \sigma_z \tag{C.44}
\]
Note the difference between this electron spin operator and the one discussed in section 2. This newly derived operator is essentially the most general form of the electron spin operator and is of matrix form. Unlike the operator discussed in section 2, this operator takes into account that there is a certain probability that an electron can be in the up or down state. It is able to relay all the information about an electron spin in 3 Cartesian axes.

The goal now is construct what are known as the projection operators. Consider now the total spin pair and spin pair magnitude operators,

\[ \mathbf{S} = \mathbf{S}_A + \mathbf{S}_B = \frac{1}{2}(\sigma_A + \sigma_B) \]  
\[ \mathbf{S}^2 = \frac{1}{4}(\sigma_A + \sigma_B)^2 = \frac{1}{4}(\sigma_A^2 + \sigma_B^2 + 2\sigma_A\sigma_B) \]

(C.45)  
(C.46)

Note that,

\[ \sigma_A^2 = \sigma_{Ax}^2 + \sigma_{Ay}^2 + \sigma_{Az}^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 3I \]
\[ \sigma_B^2 = \sigma_{Bx}^2 + \sigma_{By}^2 + \sigma_{Bz}^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 3I \]

(C.47)

where \( I \) in this case is the 2x2 identity matrix (not to be confused with nuclear spin quantum number). Therefore, the total spin magnitude operator becomes,

\[ \mathbf{S}^2 = \frac{1}{4}(3I + 3I + 2\sigma_A\sigma_B) = \frac{1}{2}(3I + \sigma_A\sigma_B) \]

(C.48)

It is also postulated that the eigenvalues of the total spin magnitude operator \( \mathbf{S}^2 \) obey the same rules of discretized angular momentum as for an individual electron discussed earlier. For example,

\[ \mathbf{S}^2|s,m_s\rangle = s(s + 1)|s,m_s\rangle \]

so,

\[ \frac{1}{2}(3 + \sigma_A\sigma_B)|s,m_s\rangle = s(s + 1)|s,m_s\rangle \]
where,

\[ \sigma_A \sigma_B = 2s(s + 1) - 3 \quad (C.49) \]

One can use this result to find the projection operators of the singlet and triplet states. The projection operators essentially provide a means to find the probability of either state using the individual spin matrices \( \sigma_A \) and \( \sigma_B \). From equation (C.49),

- \( \sigma_A \sigma_B = -3 \quad \text{for a singlet, } s = 0 \) \quad (C.50)
- \( \sigma_A \sigma_B = +1 \quad \text{for a triplet, } s = 1 \) \quad (C.51)

It is desired that if \( \sigma \sigma_B = -3 \), the singlet projection operator \( \hat{P}_s \) map that value to 1 indicating the presence of a pure singlet state. It is also desired that if \( \sigma_A \sigma_B = 1 \), the singlet projection operator \( \hat{P}_s \) map that value to 0 indicating that a pure triplet is present. This operator is easily constructed by,

\[ \hat{P}_s = \frac{1}{4} (1 - \sigma_A \sigma_B) \quad (C.52) \]

Similarly, it is desired that if \( \sigma_A \sigma_B = 1 \), the triplet projection operator \( \hat{P}_t \) map that value to 1 indicating the presence of a pure triplet state. It is also desired that if \( \sigma_A \sigma_B = -3 \), the triplet projection operator \( \hat{P}_t \) map that value to 0 indicating the presence of a pure singlet state. This operator is also easily constructed,

\[ \hat{P}_t = \frac{1}{4} (3 + \sigma_A \sigma_B) \quad (C.53) \]

The projection operators provide a means to evaluate the probability of a pure singlet and pure triplet state given the spin vectors of \( \vec{S}_A \) and \( \vec{S}_B \). So essentially, \( \sigma_A \sigma_B \) defines whether the electron pair is in a pure singlet or pure triplet state and the projection operators just provide a means to map these values to 0 or 1. These operators can also be written in terms of the spin matrix operators. From before,

\[ \sigma_j = 2\vec{S}_j \quad (C.54) \]
which leads to,

\[
\hat{P}_s = \frac{1}{4} (1 - 2\hat{S}_A \hat{S}_B) = \frac{1}{4} - \hat{S}_A \hat{S}_B \\
\hat{P}_t = \frac{1}{4} (3 + 2\hat{S}_A \hat{S}_B) = \frac{3}{4} + \hat{S}_A \hat{S}_B
\]

\hspace{1cm} (C.55)

\hspace{1cm} (C.56)

Where, once again, \(\hat{S}_A\) is the spin angular momentum matrix operator for particle A and \(\hat{S}_B\) is the spin angular momentum matrix operator for particle B. This is the most general form of the projection operators which took into account that the electron spins can be represented as a linear combination of up and down spins. These equations were solved for the spin represented as vector in section 3.2.

**APPENDIX D – Introduction to the Density Matrix**

In the KSM model for recombination, it was assumed that the electron spins could be pointing in all directions. Because of this assumption, the model allowed spin pairs to be in a mixed state, rather than a pure singlet or pure triplet state. Also, because the triplet states and singlet states allow the electron spins to be entangled, a more accurate representation of the pair would be in a mixed state. Therefore, the more precise treatment would involve the use of the density matrix to represent a pair of spins. A density matrix allows one to represent a quantum state as a mixture of states rather than a pure state. The matrix is defined by,

\[
\rho = \sum_i c_i |\psi_i\rangle \langle \psi_i|
\]

\hspace{1cm} (D.1)

where \(\psi_i\) represents the orthonormal basis states, \(c_i\) represents the probability of those states, and \(i\) represents the number of states. \(|\psi_i\rangle \langle \psi_i|\) represents an outer product of quantum states which essentially forms a matrix representation of the eigenstates. The expectation value of operator \(A\) on the density matrix equals the average value of operator \(A\) on the individual quantum states,

\[
\langle A \rangle = tr(\rho A) = \sum_i c_i \langle \psi_i | A | \psi_i \rangle
\]

\hspace{1cm} (D.2)

Consider the case for an individual electron. Because the electron can be in either the up or down states \((i = 2)\), it is more generally represented as a mixture of these states, \(|s, m_s\rangle \in \{|\uparrow\rangle, |\downarrow\rangle\}\)
\[ |\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle = c_1 |\uparrow\rangle + c_2 |\downarrow\rangle \quad (D.3) \]
\[ |\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \langle \psi_1 | = [\langle \uparrow | 0] \]
\[ |\psi_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \langle \psi_2 | = [0 \langle \downarrow |] \]

Therefore, the density matrix can be represented by,
\[ \rho = \sum_{i=1}^{2} c_i |\psi_i\rangle \langle \psi_i | = c_1 |\psi_1\rangle \langle \psi_1 | + c_2 |\psi_2\rangle \langle \psi_2 | = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \quad (D.4) \]

(Note that this matrix is related to the Pauli spin matrix derived in appendix C.) Now consider the case for two electrons. The basis states used to represent the total state of the system could be the standard basis \(|s, m_s\rangle \in \{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}\). The problem with using this basis set is that one has to work with a density matrix that is not diagonal which can be very inconvenient. It is more convenient to represent it in the singlet triplet basis so that a diagonal matrix can be formed, \(|s, m_s\rangle \in \{T_{-1}, T_0, T_{+1}, S_0\}\). From before, the basis states are represented as,
\[ |s, m_s\rangle \in \{|\uparrow\uparrow\rangle, (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2}, (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}, |\downarrow\downarrow\rangle\}\]
\[ |\psi_1\rangle = \begin{pmatrix} |\uparrow\uparrow\rangle \\ 0 \\ 0 \end{pmatrix}, \quad \langle \psi_1 | = [\langle \uparrow\uparrow | 0,0,0] \]
\[ |\psi_2\rangle = \begin{pmatrix} 0 \\ (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2} \\ 0 \end{pmatrix}, \quad \langle \psi_2 | = [0, (\langle \uparrow\downarrow | + \langle \downarrow\uparrow |)/\sqrt{2}, 0,0] \]
\[ |\psi_3\rangle = \begin{pmatrix} 0 \\ (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2} \\ 0 \end{pmatrix}, \quad \langle \psi_3 | = [0, 0, (\langle \uparrow\downarrow | - \langle \downarrow\uparrow |)/\sqrt{2}, 0] \]
\[ |\psi_4\rangle = \begin{pmatrix} 0 \\ 0 \\ |\downarrow\downarrow\rangle \end{pmatrix}, \quad \langle \psi_4 | = [0,0,0, \langle \downarrow\downarrow |] \]

The density matrix can be represented as,
\[
\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| = c_1 |\psi_1\rangle\langle\psi_1| + c_2 |\psi_2\rangle\langle\psi_2| + c_3 |\psi_3\rangle\langle\psi_3| + c_4 |\psi_4\rangle\langle\psi_4|
\]

\[
= \begin{bmatrix}
c_1 & 0 & 0 & 0 \\
0 & c_2 & 0 & 0 \\
0 & 0 & c_3 & 0 \\
0 & 0 & 0 & c_4 \\
\end{bmatrix}
\]

Therefore, the diagonal coefficients represent the probability of two electron spins to be in either of the singlet or triplet basis states. The reason for using the density matrix is so that one can represent a quantum state as a superposition of orthonormal quantum states. Because the superposition of quantum states are essentially random, it is straightforward to use stochastic or random processes to describe the state of the quantum system as it evolves as a function of time when operated on by a Hamiltonian. This is typically performed with the use of the Liouville equation which is defined as,

\[
\text{i} \hbar \frac{d}{dt} \rho(t) = [\mathcal{H}, \rho(t)] + \mathcal{R}(t)
\]

where \(\rho(t)\) is the density matrix as a function of time, \(\mathcal{H}\) is the Hamiltonian operator, \(\mathcal{R}\) is spin relaxation, and the notation \([\cdot, \cdot]\) represents the commutation operation. This method was first used to investigate SDR by Haberkorn and Dietz in 1980 [16]. Following studies were performed by Barabanov and Tretiak in 1996 [17] and then in 2003 by Boehme and Lips [14]. Because it would require a great amount of time to explain the mathematics in these studies, it will not be discussed here. However, it should be noted that in the study of Haberkorn, it was shown that simply the addition of an external magnetic field strongly influences the recombination current within a device.
APPENDIX E – Biasing Box used for DCIV, SDCP, and BAE

The biasing box illustrated below was used to appropriately bias the various fully processed devices in the experiments performed in this study. By configuring the $S$, $B$, $D$, and $G$ switches, one is able to bias the DUT with the techniques of either DCIV, SDCP, and BAE. (The switches illustrated in the figure are configured for DCIV biasing mode.) Here, the internal power sources simply represent batteries. An external power supply may be used by appropriately configuring the DPDT $SW_1$ and connecting its leads to the positive and negative input banana plugs of the box. This mode is recommended for experiments that need very stable voltages to applied to a device for long periods of time. Note that an external waveform must be connected to the $ext$ BNC connector in order to perform a SDCP measurement.
APPENDIX F – Non-Technical Abstract

The semiconductor industry is currently pressing its technological limits in attempts to keep pace with Moore’s Law. Despite much skepticism, corporations still manage to mass produce micro- and nano-electronic circuits which enables the continuing trend of doubling the number of transistors that can be placed on a chip every 18 months. Only recently has it been realized by scientists and engineers that, even though the goal they are trying to meet is steadily increasing, the requirements for advances in the downscaling of technology are ever more challenging. Today, devices are being fabricated on an extremely small scale (oxide thicknesses of about 10 atoms across) and are being made with new and more complex material systems. At such extremely small scales, individual atomic scale defects can be responsible for the improper functioning of such devices. As a result, new and more innovative ways are needed to detect these atomic scale defects in fully processed devices. Using sophisticated digital signal processing methods, custom built spectrometers, and the utilization of undiscovered physics, we demonstrate new acquisition techniques that are capable of detecting the very same atomic scale defects that are responsible for device performance. Detection and identification of these defects helps scientists and engineers to fabricate more efficient devices. Conventional spectrometers used for this purpose typically weigh 1,000s of lbs., have footprints of dozens of square feet, and can cost more than $250,000. We demonstrate that our new detection techniques allows for spectrometers to be built that can weigh as low as 10 lbs, can be held in one hand, and be built with parts costing no more than $2,500. The work has also led to an increase in the sensitivity of the measurements made and allows one to acquire the data more rapidly. Not only are these techniques useful for device physics research, but also may find much use in magnetometry and magnetic isolation circuits.
Bibliography


VITA

Corey Jonathan Cochrane was born on January 27, 1982 in Bristol Pennsylvania to Diane Susan Fritsch and John Edward Cochrane. He graduated from Pennsbury High School in June 2000 and then enrolled in classes at Penn State University beginning in August. While at Penn State, he held internships during the summers of 2003 and 2004 with NASA’s Undergraduate Student Research Program (USRP) working on biologically inspired robotics in the BioVisualization Lab at NASA’s Ames Research Center. In December 2004, Corey received a B.S. degree in electrical engineering and enrolled in graduate school at Penn State University while performing research in the Semiconductor Spectroscopy Lab. During this time, his work involved utilizing digital signal processing as a means to reduce noise in continuous wave magnetic resonance based measurements. He earned an M.S. degree in electrical engineering from Penn State University in December 2007. From 2008 to 2010, he worked on satellite communication systems in El Segundo CA at Boeing Satellite Systems in the DSP algorithms group. Corey returned to The Pennsylvania State University in 2010 and earned his Ph.D. in Engineering Science in May 2013. During this time, Corey has published 21 papers, has been granted one patent with an additional one pending, made 15 oral presentations, and made 11 poster presentations.