5-2012

Techniques for Improving Parameter Estimation in ESR

Troy R. Broderick
University at Albany, State University of New York

TECHNIQUES FOR IMPROVING PARAMETER ESTIMATION IN ESR

Troy R. Broderick
Physics Department, University at Albany
1400 Washington Ave, Albany New York 12222

Abstract

This paper is to address the need for better estimates of the ESR spectral parameters of biological systems. In this paper I focus on models of spectral lineshapes and will be applying these techniques to more complicated lineshapes. Multifrequency ESR is a powerful technique for identifying and characterizing dynamic modes in natively unfolded biopolymers in general and loop regions in particular. I will introduce the reader to properties of ESR, formalisms and common techniques in the field. A common result for a simple ESR line is a Lorentzian distribution that is characterized by a mode and width. Information about the structure of the sample is contained in the position of the Lorentzian distribution attained from ESR. The dynamics of the sample can determined from the width of the distribution. The Lorentzian distribution can be modelled using Student's T distribution by allowing the degrees of freedom in the T distribution to approach unity. Before applying Bayesian analysis to the T distribution I first verify that my proposed techniques will be effective by applying these new methods of estimation to the Gaussian distribution, which is the limiting case of the number of degrees of freedom going to infinity in the Student's t distribution.

INTRODUCTION

ESR is a widely used technique for probing the atomic structure and magnetic properties of many compounds. The experiment is based on the magnetic properties of the sample and can be used with any sample that has a non-vanishing magnetic moment due to spin; orbital angular momentum, or both. On the other hand ESR is an approximation for samples that have one or more unpaired electrons with a net electron spin. If we imagine the electrons with a magnetic moment of ½ and a spin quantum number of ±1/2.

$$s=\frac{1}{2},$$

$$m_s=\frac{\pm 1}{2}.$$

With the application of a magnetic field of strength B_0 we observe that the -1/2 spin electrons align parallel to the field corresponding to lower energy and the +1/2 spin electrons align antiparallel to the field corresponding to higher energy. The energy splitting of each electron having a specific energy is due to the Zeeman Effect which will be covered in greater detail later in this paper. The energy difference between the lowest and highest energy is directly proportional to

$$\Delta E=\mu_B g_e B_0 \text{ [10].}$$

Where g_e is the electrons g-factor and μ_B is the Bohr magneton. The g-factor will be explored in greater detail in the following pages. For now we note the equations which directly follow from above and we notice that we have found the fundamental equation allowing ESR to be so effective.

$$\Delta E=\varepsilon=h\nu=g_e\mu_B B_0 \text{ [10].}$$

Where ε is the energy of a photon emitted or absorbed by the sample in order for it to move between energy levels. By varying either the frequency of incoming photons or the magnetic field we can produce a spectra containing information about the samples structure and dynamics. In most applications and in this paper, I will only discuss the case of varying magnetic field with constant microwave background. When the magnetic field strength is varied, most commonly called sweeping the field, it produces a variable energy difference between the spin up and spin down levels. When this energy is equivalent to the energy of the applied microwaves we have exact resonance and

the electrons can move most freely between up and down states. This behaviour is what we see when plotting the signal vs. the magnetic field strength. Figure 1 is a theoretical graph the signal vs. magnetic field strength of an ESR experiment on a free electron.

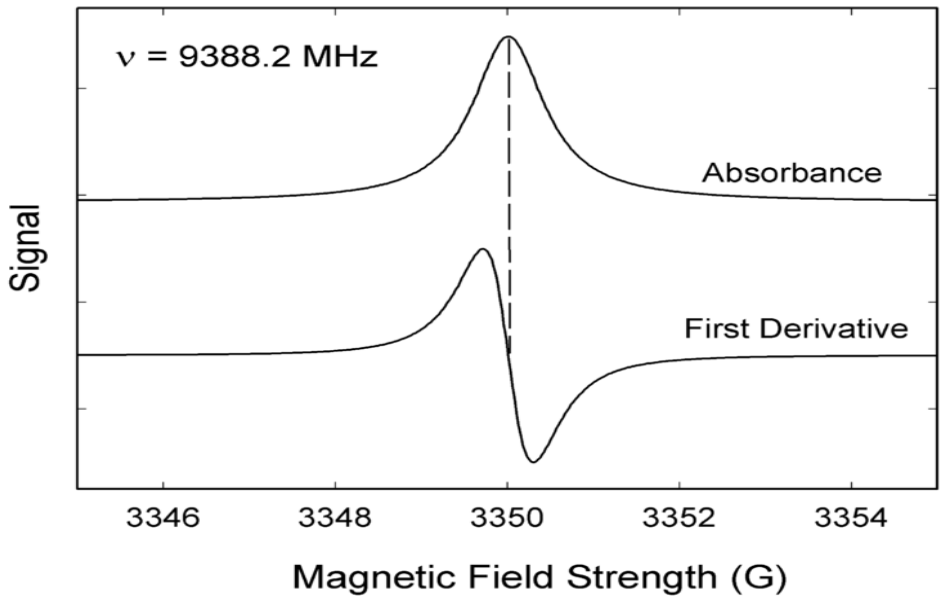

Fig. 1: Theoretical graph of the signal strength vs. the magnetic Field Strength.

The area under both halves of the first derivative graph is proportional to the number of unpaired electrons in a sample [8].

The width of the absorption curve is determined by the coupling of the sample to its surroundings and can be used to study relaxation phenomena. The amplitude is related to the number of ESR active spins in the sample. Sensitive spectroscopers have a sigh signal to noise ratio, which can aid in the detection of weak signals [10]. The sensitivity is reduced by noise which is always present in the system and reduction of which has been the focus of researchers since the beginning of ESR. There are three traditional means of reducing noise in ESR; reducing the bandwidth of the system, making more stable equipment, and by improving the amplifier. Commonly in ESR and throughout this paper noise will be specified by the signal-to-noise ratio.

BASIC THEORY OF MAGNETIC RESONANCE

To begin deriving the theories of magnetic resonance a simple model is used. The most basic model is the isolated spin in an external magnetic field. Allow the magnetic field to vary in time and write the torque induced on the magnetic moment by the field as,

$$\mathrm{T} = \mu \times B \ .$$

This torque will cause the magnetic moment to tend to point along the direction of the magnetic field. However in a frictionless model the direction of the moment will precess around the direction of the magnetic field. If the object with a magnetic moment also has angular momentum then the system will need to conserve the angular momentum and process circularly around the direction of the magnetic field. Again if this is frictionless then the angle between the moment and the magnetic field will remain constant over time. When friction is added the moment falls toward the direction of the magnetic field, corresponding to a relaxation process. Making the following substitutions we have the equation of motion for the magnetic moment.

$$\frac{dJ}{dt} = \mu \times B \ ,$$

with, $\mu = \gamma J$ giving,

$$\frac{d\mu}{dt} = \mu \times (\gamma B) \quad [4].$$

It is at this point in most ESR texts that we change to a rotating reference frame. The system is processing around the magnetic field as seen by its equation of motion, so make the reference frame process at the same rate. Then the magnetic moment, by defining the rate of change of the reference frame as equal to the rate of change of the magnetic moment, will appear to be constant in the absence of friction. To perform this reference frame transformation we take vectors i, j, and k defined as an orthonormal basis. A function F(t) is made up of its x, y and z components in the directions of i, j, and k respectively and is time dependent so that,

$$F(t) = F_x(t)\hat{i} + F_y(t)\hat{j} + F_z(t)\hat{k} \ ,$$

now we can investigate a rotating frame on our test function F(t) by letting i, j, and k rotate with fixed unit lengths. If they rotate with angular velocity Ω then there equation of motion can be described in the following differential equation,

$$\frac{d\hat{i}}{dt} = \Omega \times \hat{i} \ ,$$

which describes a rotation. Then when we take the time derivative of F(t) we get,

$$\frac{d\vec{F}}{dt} = \hat{i}\frac{df_x}{dt} + f_x\frac{d\hat{i}}{dt} + \hat{j}\frac{df_y}{dt} + f_y\frac{d\hat{j}}{dt} + \hat{k}\frac{df_z}{dt} + f_z\frac{d\hat{k}}{dt}$$

$$\frac{d\vec{F}}{dt} = \frac{\delta\vec{F}}{\delta t} + \Omega \times \vec{F}$$

,

with $\frac{\delta\vec{F}}{\delta t}$ defined as the time rate of change with respect to the coordinate vectors i, j and k of F(t) [4]. Now with this new rotating coordinate system we can write the equation of motion of μ.

$$\frac{d\mu}{dt} = \frac{\delta\mu}{\delta t} + \Omega \times \mu = \mu \times \gamma B \ ,$$

or,

$$\frac{\delta\mu}{\delta t} = \mu \times (\gamma B + \Omega) \ \text{[4].}$$

Note that this is the same equation of motion as before except with $B_0 = B + \frac{\Omega}{\gamma}$ giving an apparent B field that is different from its observed value in a lab frame. This is an instance of Larmors theorem. To understand our new frame of reference we can define the Larmor frequency γB_0 which is the angular frequency that the magnetic field would process at so that $\frac{\delta\mu}{\delta t} = 0$ meaning the magnetic moment is fixed with respect to the rotating field i, j and k.

Bloch Equations

As previously mentioned, when friction is added to the time evolution of the magnetic moment system in a magnetic field B, the magnetic moment will align parallel to magnetic field after some time called the relaxation time. The new equations of motion which include this relaxation factor are called the Bloch equations, named after Felix Bloch who proposed the equations in 1946. To find the Bloch equations we will define M_0 as the equilibrium

magnetization along the direction of B and M_z as the magnetization in the z direction. We can make the assumption that M_z approaches M_0 in the following manner,

$$\frac{d M_z}{dt} = \frac{M_z - M_0}{T_1} .$$

Where T_1 is the spin-lattice relaxation time of the sample. It is called the spin-lattice relaxation time because the energy given off by the rotating sample so that it can reach equilibrium is given to its surroundings or lattice. Experiments show that the spin-lattice relaxation times for electrons are much shorter than for nuclei, due to the strong coupling of the electrons to their surroundings, the nuclei.

Similarly there is a relaxation time associated with the x and y components of M. to understand this relaxation time we can imagine a magnetization M tilted slightly away from the z-axis toward the x-axis. When Precession is started each component of the sample has its own spin value due to small differences in the total magnetic field arising from the randomly oriented magnetic moment in the sample as well as the applied field and will precess at a slightly different rate, leading to an ensemble of spins initially in phase and progressively falling out of phase, then moving to equilibrium once again. T_2 is the characteristic time for the spins that start in phase to fall out of phase and is called the transverse relaxation time. Therefore we can write equations for the x and y components of M to tend to equilibrium as,

$$\frac{d M_x}{dt} = \frac{-M_x}{T_2}$$
$$\frac{d M_y}{dt} = \frac{-M_y}{T_2} .$$

In general the observed transverse relaxation time is made up partly of the spin-lattice relaxation time and a dephasing relaxation time, however the dephasing relaxation time is often so fast that the spin-lattice effect is nearly negligible. Using Redfield theory we can show that in many cases of interest $\frac{1}{T_2} = \frac{1}{T_2'} + \frac{1}{2T_1}$ [14].

We must begin with the Liouville equation for the density operator,

$$\frac{\delta \rho(t)}{\delta t} = -2\pi i [H, \rho(t)] ,$$

with h=1 and c=1 as is common practice. We can write the Hamiltonian as the combination of its unperturbed and perturbed parts $H = H_0 + \sum H_n$ where,

$$H_n = \sum A_k E_k(t) \ .$$

In the interaction picture we can make the following substitutions,

$$\rho_I(t) = \mathrm{e}^{2\pi i H_0 t} \rho(t) \mathrm{e}^{-2\pi i H_0 t}$$

$$\frac{\delta \rho_I(t)}{\delta t} = -2\pi i [H_{n,I}(t), \rho_I(t)] \ ,$$

which has the perturbation assumed to make small changes in $\rho_I(t)$. The fluctuating field operators $A_{I,K}$ now have a time dependence giving the perturbed hamiltonian a time dependence from two factors. We then integrate both sides and treat this solution as the beginning of a series expansion to give:

$$\rho_I(t) = \rho_I(0) - 2\pi i \int_0^t d\,\mathrm{T} [\mathrm{H}_{n,I}(\mathrm{T}), \rho_I(\mathrm{T})]$$

$$\rho_I(t) = \rho_I(0) - 2\pi i \int_0^t d\,\mathrm{T} [\mathrm{H}_{n,I}(0), \rho_I(0)] - 4\pi^2 \int_0^t d\,\mathrm{T} \int_0^{\mathrm{T}} d\,\mathrm{T}' [\mathrm{H}_{I,n}(\mathrm{T}), [\mathrm{H}_{I,n}(\mathrm{T}'), \rho_I(\mathrm{T})] \ ] + \ \cdots \quad \text{[14].}$$

Now if we take the time derivative,

$$\frac{d\,\rho_I(t)}{dt} = -2\pi i [H_{I,n}(t), \rho_I(0)] - 4\pi^2 \int_0^t d\,\mathrm{T} [H_{I,n}(t), [H_{I,n}(\mathrm{T}), \rho_I(0)] \ ] + \ \cdots \ .$$

If we take the ensemble average of this with the assumption that $\langle E_k(t) \rangle = 0$, then the first term will drop off and we are left with,

$$\frac{d\,\rho_I(t)}{dt} = -4\pi^2 \int_0^t d\,\mathrm{T} \langle [H_{I,n}(t), [H_{I,n}(\mathrm{T}), \rho_I(0)] \ ] \rangle \ .$$

and separation yields,

$$\frac{d\,\rho_I(t)}{dt} = -4\pi^2 \sum_{l,k} \int_0^t \langle E_k(t) E_l(\mathrm{T}) \rangle d\,\mathrm{T} \langle [A_k(t), [A_l(\mathrm{T}), \rho_I(0)] \ ] \rangle \quad \text{[14].}$$

If the ensemble we average over is stationary then,

$$\frac{d\,\rho_I(t)}{dt} = -4\pi^2 \sum_{l,k} \int_0^t \langle E_k E_l(t - \mathrm{T}) \rangle d\,\mathrm{T} \langle [A_k(t), [A_l(\mathrm{T}), \rho_I(0)] \ ] \rangle$$

$$\frac{d\,\rho_I(t)}{dt} = -4\pi^2 \sum_{l,k} \int_0^t \langle E_k E_l(\mathrm{T}) \rangle d\,\mathrm{T} \langle [A_k(t), [A_l(t - \mathrm{T}), \rho_I(0)] \ ] \rangle \ .$$

Changing the upper limit of the integral to infinity, we make the assumption that the correlation functions $\langle E_k E_l(\mathrm{T}) \rangle$ decay rapidly,

$$\frac{d\,\rho_I(t)}{dt} = -4\pi^2 \sum_{l,k} \int_0^\infty \langle E_k E_l(\mathrm{T}) \rangle d\,\mathrm{T} \langle [A_k(t), [A_l(t - \mathrm{T}), \rho_I(0)] \ ] \rangle \ .$$

More specifically for the correlation functions we make another assumption that the time they take to decay is T_C, which allows the replacement of $\rho_I(0)$ with $\rho_I(t)$, and again have a differential equation for $\rho_I(t)$. Now I must work out the commutator terms for which I use the eigenfunctions of the Hamiltonian. We get,

$$\frac{d(\rho_I)_{\alpha\alpha'}}{dt} = -4\pi^2 \sum_{k,l} \int_0^\infty d\,\mathrm{T} \langle E_k E_l \rangle \langle \alpha | [A_l(t), [A_k(t-\mathrm{T}), \rho_I]] | \alpha' \rangle .$$

The commutator has four terms,

$$\langle \alpha | A_l(t) A_k(t-\mathrm{T}) \rho_I | \alpha' \rangle = -4\pi^2 \sum_{\beta,\beta} \langle \alpha | A_l(t) | \beta \rangle \langle \beta | A_k(t-\mathrm{T}) | \beta' \rangle \langle \beta' | \rho_I | \alpha' \rangle .$$

Which by the relation stated above can be shown to be equal to,

$$\sum_{\beta,\beta'} e^{2\pi i(\alpha-\beta)t} e^{2\pi i(\beta-\beta')(t-\mathrm{T})} (A_l)_{\alpha\beta} (A_k)_{\beta\beta'} (\rho_I)_{\beta'\alpha'} \quad [14].$$

After we crawl through the broken glass of substitutions and the use of a dummy variable we can get a results of the form,

$$\begin{aligned}
\langle \alpha | A_l(t) A_k(t-\mathrm{T}) \rho_I | \alpha' \rangle &= \sum_{\beta,\beta'} [\delta_{\alpha'\beta'} \sum_\gamma e^{2\pi i(\beta-\gamma)\mathrm{T}} e^{2\pi i(\alpha-\beta)t} (A_k)_{\gamma\beta} (A_k)_{\alpha\gamma}] (\rho_I)_{\beta\beta'} \\
-\langle \alpha | A_l(t) \rho_I A_k(t-\mathrm{T}) | \alpha' \rangle &= -\sum_{\beta,\beta'} e^{2\pi i(\alpha'-\beta')\mathrm{T}} e^{2\pi i(\alpha-\beta+\beta'-\alpha')t} (A_k)_{\beta'\alpha'} (A_k)_{\alpha\beta} (\rho_I)_{\beta\beta'} \\
-\langle \alpha | A_k(t-\mathrm{T}) \rho_I A_l(t) | \alpha' \rangle &= -\sum_{\beta,\beta'} e^{2\pi i(\beta-\alpha)\mathrm{T}} e^{2\pi i(\alpha-\beta+\beta'-\alpha')t} (A_k)_{\alpha\beta} (A_k)_{\beta'\alpha'} (\rho_I)_{\beta\beta'} \\
\langle \alpha | \rho_I A_l(t-\mathrm{T}) A_k(t) | \alpha' \rangle &= \sum_{\beta,\beta'} [\delta_{\alpha\beta} \sum_\gamma e^{2\pi i(\gamma-\beta')\mathrm{T}} e^{2\pi i(\beta'-\alpha)t} (A_k)_{\beta'\gamma} (A_k)_{\gamma\alpha'}] (\rho_I)_{\beta\beta'}
\end{aligned} ,$$

Going back to the Heisenberg picture we get the Redfield equation in its most general form.

$$\begin{aligned}
\frac{\delta \rho_{\alpha\alpha'}(t)}{\delta t} &= -2\pi i(\alpha-\alpha')\rho_{\alpha\alpha'}(t) - 4\pi^2 \sum_{l,k} \sum_{\beta,\beta'} d\,\mathrm{T} \langle E_k E_l(\mathrm{T}) \rangle e^{2\pi i(\beta-\gamma)\mathrm{T}} (A_k)_{\gamma\beta} (A_l)_{\alpha\gamma} \\
&- \int_0^\infty d\,\mathrm{T} \langle E_k E_l(\mathrm{T}) \rangle e^{2\pi i(\alpha'-\beta')\mathrm{T}} (A_k)_{\beta'\alpha'} (A_l)_{\alpha\beta} - \int_0^\infty d\,\mathrm{T} \langle E_k E_l(\mathrm{T}) \rangle e^{2\pi i(\beta-\alpha)\mathrm{T}} (A_k)_{\alpha\beta} (A_l)_{\beta'\alpha'} \quad [14]. \\
&+ \delta_{\alpha\beta} \sum_\gamma \int_0^\infty d\,\mathrm{T} \langle E_k E_l(\mathrm{T}) \rangle e^{2\pi i(\gamma-\beta')\mathrm{T}} (A_k)_{\beta'\gamma} (A_l)_{\gamma\alpha'} \rho_{\beta\beta'}(t)
\end{aligned}$$

Finally for simple exponentially decaying correlation functions we can simplify the integrals,

$$\langle E_k E_l(\mathrm{T}) \rangle = \delta_{kl} \langle E^2 \rangle e^{\frac{-t}{T_0}} \equiv \delta_{kl} \langle E^2 \rangle e^{2\pi\zeta t} ,$$

$$\int_0^\infty d\,\mathrm{T} \langle E_k E_l(\mathrm{T}) \rangle e^{2\pi i\omega\mathrm{T}} = \delta_{kl} \frac{\langle E^2 \rangle}{2\pi(\zeta - i\omega)} \equiv \delta_{kl} \frac{1}{2\pi} J(\omega) .$$

Which gives our final Redfield equation as,

$$\frac{\delta \rho_{\alpha\alpha'}(t)}{dt} = -2\pi i(\alpha-\alpha')\rho_{\alpha\alpha'}(t)$$

$$-2\pi \sum_{\beta\beta'}[\delta_{\alpha'\beta'}\sum_{\gamma} J(\beta-\beta')A_{\gamma\beta}\cdot A_{\alpha\gamma} - [J(\alpha'-\beta')+J(\beta-\alpha)]A_{\beta'\alpha'}\cdot A_{\alpha\beta} + \delta_{\alpha\beta}\sum_{gamma} J(\gamma-\beta')A_{\beta'\gamma}\cdot A_{\gamma\alpha'}]\rho_{\beta\beta'}(t)$$

where,

$$A_{\beta'\alpha'}\cdot A_{\alpha\beta} = \sum_{l}(A_l)_{\beta'\alpha'}(A_l)_{\alpha\beta}\ .$$

The next step is to connect the Redfield equations to the spin relaxation in a fluctuating magnetic field. Specifically we will focus on the spin-1/2 corresponding to an electron, in a magnetic field in the z-direction and write the unperturbed Hamiltonian,

$$H_0 = -\epsilon I_z\ .$$

and fluctuating perturbations,

$$H_r = -\vec{H}\cdot\vec{I}\ ,$$

with $\vec{H}$ containing the gyromagnetic ratio and other constants and is proportional to the magnetic field [14]. So with this new form we can replace E^2 with H^2 and all of our A operators are now Spin operators. The energy in the ground state, $|0\rangle$, is equal to -1/2ϵ and the energy of the first excited state, $|1\rangle$, is 1/2ϵ and we see that the trace of the ρ matrix is conserved and is equal to 1, that is, $\rho_{00}+\rho_{11}=1$. Let us compute the derivative of the 00 component of ρ,

$$\frac{\delta \rho_{00}}{\delta t} = -2\pi\sum_{\beta,\beta'}[J(\beta-\gamma)A_{\gamma\beta}\cdot A_{0\gamma} - [J(-\beta')+J(\beta)]A_{\beta'0}\cdot A_{0\beta} + \delta_{0\beta}\sum_{\gamma} J(\gamma-\beta')A_{\beta'\gamma}\cdot A_{\gamma 0}]\rho_{\beta\beta'}$$ [14].

Using the Pauli spin matrices, assuming ħ=1, we can show,

$$(\vec{I})_{\gamma\beta}\cdot(\vec{I})_{0\gamma} = \frac{1}{2}\delta_{\beta 0}\delta_{\gamma 1} + \frac{1}{4}\delta_{\beta 0}\delta_{\gamma 0}$$

$$(\vec{I})_{\beta'0}\cdot(\vec{I})_{0\beta} = \frac{1}{2}\delta_{\beta 1}\delta_{\beta' 1} + \frac{1}{4}\delta_{\beta 0}\delta_{\beta' 0}\ .$$

$$(\vec{I})_{\beta'\gamma}\cdot(\vec{I})_{\gamma 0} = \frac{1}{2}\delta_{\beta' 0}\delta_{\gamma 1} + \frac{1}{4}\delta_{\beta' 0}\delta_{\gamma 0}$$

Now,

$$\frac{\delta \rho_{00}}{\delta t} = -\pi[J(\epsilon)+J(-\epsilon)]\sum_{\beta,\beta'}[\delta_{\beta 0}\delta_{0\beta'} - \delta_{\beta 1}\delta_{\beta' 1}]\rho_{\beta\beta'}\ ,$$

so that the time dependent form is,

$$\frac{\delta \rho_{00}(t)}{\delta t} = -\pi [J(\epsilon) + J(-\epsilon)][\rho_{00}(t) - \rho_{11}(t)] .$$

Now we put our expression for J into this equation and simplify. After the dust settles we get,

$$\frac{\delta \rho_{00}(t)}{\delta t} = \frac{-2\pi\zeta \langle H^2 \rangle}{\epsilon^2 + \zeta^2}[\rho_{00}(t) - \rho_{11}(t)] .$$

It is important at this point to note that the time derivative of the 00 component is equal to the negative dime derivative of the 11 component [14].

Now we can define the z-magnetization and the equation for coherences.

$$\frac{\delta M_z(t)}{\delta t} = \frac{-2\pi}{T_1} M_z(t) ,$$

and

$$\frac{\delta \rho_{01}(t)}{\delta t} = 2\pi i \epsilon \rho_{01}(t)$$

$$-2\pi \sum_{\beta\beta'} [\delta_{\beta' 1} \sum_{\gamma} J(\beta - \gamma) A_{\gamma\beta} \cdot A_{0\gamma} - [J(\frac{1}{2}\epsilon\beta') + J(\beta + \frac{1}{2}\epsilon)] A_{\beta' 1} \cdot A_{0\beta} + \delta_{\beta 0} \sum_{\gamma} J(\gamma - \beta') A_{\beta'\gamma} \cdot A_{\gamma 1}] \rho_{\beta\beta'}(t)$$

Since we already calculated the first operator term, we can calculate the second and third,

$$(\vec{I})_{\beta' 1} \cdot (\vec{I})_{0\beta} = \frac{-1}{4} \delta_{\beta' 1} \delta_{\beta 0}$$

$$(\vec{I})_{\beta'\gamma} \cdot (\vec{I})_{\gamma 1} = \frac{1}{2} \delta_{\beta' 1} \delta_{\gamma 0} + \frac{1}{4} \delta_{\beta' 1} \delta_{\gamma 1}$$

To complete the derivation, plug these into our equation for the time derivative of the 01 component of the density matrix and simplify. At this point notice that the complex conjugate of J*(ϵ) = J(-ϵ), and take the complex conjugate of the derivative of 01 to get the derivative of the 10 density matrix component. Now calculate the magnetization in the x and y directions,

$$M_x(t) = Tr[I_x \rho(t)] = \frac{1}{2}(\rho_{01} + \rho_{10}) ,$$

and,

$$\frac{\delta M_x(t)}{\delta t} = 2\pi(\epsilon + \Im[J(\epsilon)]) M_y(t) - 2\pi(J(0) + \Re[J(\epsilon)]) M_x(t) .$$

Finally we note that $M_y(t) = \frac{i}{2}(\rho_{01} - \rho_{10})$, and that the imaginary part of the spectral density will cause a frequency shift which we can ignore. Now the relaxation time has been proven to be,

$$\frac{1}{T_2} = J(0) + \Re[J(\epsilon)] = \frac{1}{2T_1} + \frac{1}{T_2'} \quad [14].$$

Now with these definitions of relaxation times we are ready to derive the Bloch equations. First we note that from above,

$$\frac{d\vec{M}}{dt} = \gamma \vec{B} \times \vec{M} - \vec{i}\frac{M_x}{T_2} - \vec{j}\frac{M_y}{T_2} - \vec{k}\frac{M_z - M_0}{T_1}$$

the B field in ESR is composed of the static field of the magnet in the z direction and the magnetic field of the photons. If we define the static field as B_0 and the magnetic field of the photons as B_1 in the x-y plane. Then I will define the total Magnetic field as,

$$\vec{B} = \vec{i}\, B_1 \cos(\omega t) + \vec{j}\, B_1 \sin(\omega t) + \vec{k}\, B_0 \quad [8].$$

Which gives the cross product,

$$\vec{B} \times \vec{M} = -(B_0 M_y - B_0 M_z \sin(\omega t))\vec{i} + (B_0 M_x - B_0 M_z \cos(\omega t))\vec{j} + (M_y \cos(\omega t) - M_x \sin(\omega t))\vec{k}$$

Finding the time derivatives of the x, y and z components of M then just requires substitution resulting in the following equations,

$$\frac{dM_x}{dt} = -\gamma B_0 M_y + \gamma B_1 M_z \sin(\omega t) - \frac{M_x}{T_2}$$

$$\frac{dM_y}{dt} = \gamma B_0 M_x - \gamma B_1 M_z \cos(\omega t) - \frac{M_y}{T_2} \quad .$$

$$\frac{dM_z}{dt} = \gamma B_1 (M_y \cos(\omega t) - M_x \sin(\omega t)) - \frac{M_z - M_0}{T_1}$$

Which are the Bloch equations [8].

Zeeman Splitting

The Zeeman splitting is named after Physicist Pieter Zeeman, who shared the 1902 Nobel Prize for discovering it [10]. The Zeeman effect applies to many spectral problems in nature and in the ESR case it is called the anomalous Zeeman effect. It was named thus because electron spin hadn't yet been discovered when the spectral line splitting of the electron sub-levels was found. This effect is essential to ESR as it is unique for every substance and defines the absorption energies for the sample. We are interested in the case that the Zeeman splitting is large compared to the fine-structure splitting. In this high-field case, high-field meaning it is a result of a large static magnetic field, we see

disruption in coupling between the orbital spins and the spin angular momenta, an effect which is called the Paschen-Back effect.

We shall start by defining the magnetic moment of an electron as the electric current multiplied by a vector **A** whose magnitude is equal to the surface area enclosed by the orbiting electron and a direction perpendicular to its surface. We also define the area classically,

$$\vec{\mu} = I\vec{A}$$

$$\vec{A} = \frac{1}{2}\frac{\vec{l}}{m_e}T \quad .$$

Above **l** is the orbital angular momentum of the electron, T is the period and m is the mass of the electron. Classically the current is just charge over period, by exploiting this we can write our magnetic moment again as follows.

$$\vec{\mu} = \frac{-e}{2\mathrm{m}_e}\vec{l} \quad [8].$$

Now that we have an explicit equation for the magnetic moment we can substitute it into the equation for the energy of a magnetic dipole in a magnetic field giving,

$$W = -\vec{\mu}\cdot\vec{B}$$

$$W = \frac{e}{2\mathrm{m}_e}\vec{l}\cdot\vec{B} \quad [8].$$

Quantum mechanically $\vec{l}\cdot\vec{B} \equiv \hbar MB$ where M is the magnetic quantum number. Making substitutions once again we get,

$$W = \frac{e\hbar}{2\mathrm{m}_e}MB \quad .$$

But the Bohr magneton is just $e\hbar/2\mathrm{m}_e$ so finally we have,

$$W = \mu_B MB \quad .$$

This result is in agreement with an exact quantum mechanical calculation when spin, and therefore fine-structure, are neglected [8]. Since M is the angular momentum component in the direction of **B** we have some selection rules to follow. When polarization is perpendicular to B which is commonly called σ-polarization in literature in the field, then ΔM can take the values ±1. When the polarization is parallel to B, called π-polarization, ΔM=0. With these selection rules we see we have a spectral line split to three components by the magnetic field. This type of three

component split by the Zeeman effect is called a Lorentz triplet. If ω_o is the angular frequency of the line without a magnetic field perturbation then the frequencies of the triplet can be expressed as follows,

$$\Delta M = +1, \qquad \omega_- = \omega_0 - \frac{\mu_B}{\hbar}\vec{B} \qquad \textit{circluar polarization}$$

$$\Delta M = 0, \qquad \omega_0 \qquad \textit{linear polarization} \quad [8].$$

$$\Delta M = -1, \qquad \omega_+ = \omega_0 + \frac{\mu_B}{\hbar}\vec{B} \qquad \textit{circluar polarization}$$

These equations describe the normal Zeeman effect and give a difference between two adjacent angular frequencies as $\frac{\mu_B}{\hbar}\vec{B}$ or by substituting in wave numbers which are often used in spectroscopy it is shown,

$$\Delta\tilde{\nu} = \Delta\frac{1}{\lambda}$$
$$\Delta\vec{\nu} = \frac{e}{4\pi m_e c}\vec{B} \qquad [8].$$

The constant in the last equation is called the Lorentz factor and $\Delta\upsilon$ is the Lorentz unit of the Zeeman splitting. There are conditions in order for the Lorentz theory to be valid. We require $\Delta\upsilon$ to be much greater than the fine-structure splitting. We can expand this for the 1 electron system, or the hydrogen case.

$$\frac{e}{4\pi m_e c} > \frac{R\alpha^2 Z^4}{n^3 L(L+1)}.$$

Which for the 2p and 3p states B>.78T and B>.23T respectively, meaning that Zeeman effect in hydrogen is normal when B>1Tesla.

Now if we assume that the Zeeman splitting is much less than the fine-structure splitting we redefine μ in the energy equation as, $\vec{\mu} = -\mu_B g\vec{J}$, with g being the Lande g-factor and J is the total angular momentum vector [8].

CALCULATING LINESHAPES

With this basic introduction to the principles of ESR we can begin looking at the parameter estimation problem. To begin I will go over two different methods of calculating the ESR lineshape, the stochastic Liouville equation for slow motion and general descriptions of ESR spectra, and Redfield theory which applies to the fast tumbling limit. The slow motion limit is defined by $\Delta > 1/\tau$ where Δ is the magnitude of the magnetic anisotropies and τ is the

relaxation time. I will examine first the ultra-slow motion case involving solutions to the Liouville equation, then consider Redfield theory alternatively.

The g Tensor Problem

The axially symmetric g tensor when considered within the secular approximation is is one of the simplest cases for the asymptotic analysis of the stochastic Liouville operator. This is the slow procession limit where Redfield theory breaks down. To generate an expression for the spin spin correlation function we can start with the differential equation,

$$\frac{\delta\hat{\rho}}{\delta t}=\hat{\hat{\mathscr{L}}}\hat{\rho}\,,$$

which by inspection gives solutions,

$$\hat{\rho}(t)=\mathrm{e}^{\hat{\hat{\mathscr{L}}}t}\hat{\rho}(0)\,.$$

Now taking the fourier transformation of the column vector ρ,

$$\tilde{\rho}(t)=F(\hat{\rho})=\int \mathrm{e}^{i\omega t}\mathrm{e}^{\hat{\hat{\mathscr{L}}}t}\hat{\rho}(0)\,dt$$

$$\tilde{\rho}(t)=\frac{1}{\hat{\hat{\mathscr{L}}}+i\,\omega}\hat{\rho}(0)$$

Now multiplying rho transverse by the fourier transform of rho we have,

$$\hat{\rho}(0)^{\dagger}\tilde{\rho}(t)=\tilde{\rho}(0)^{\dagger}[\hat{\hat{\mathscr{L}}}+i\,\omega]^{-1}\hat{\rho}(0)\equiv\left\langle \nu\middle|[\hat{\hat{\mathscr{L}}}+i\,\omega]\middle|\nu\right\rangle$$

$$\hat{\mathscr{L}}=\mathrm{H}^{x}+\gamma$$

In this expression for the spin-spin correlation function H^{x}, can be represented by $-\omega^{\perp}-\Delta x^{2}$. This allows me to write the spectral function as a matrix element of an inverse matrix,

$$\vec{I}(\omega)=\Re[\int dx(i(\omega-\omega_{\perp}-\Delta x^{2})+\Gamma+\frac{1}{T_{2}^{0}})^{-1}]\ \text{[5].}$$

where ω is the microwave frequency, $1/T_{2}^{0}$ is the intrinsic linewidth, Γ is the stochastic time evolution operator and x=cos(β) is the stochastic variable associated with the angle β between the molecular axis of symmetry and the

direction of the static field B_0 [5]. Δ is the magnetic anisotropy and is equal to the difference in the resonance frequencies, $\omega_{\|}-\omega_{\perp}$.

$$\omega_{\|}=\frac{g_{\|}\beta_e B_0}{\hbar}$$
$$\omega_{\|}=\frac{g_{\perp}\beta_e B_0}{\hbar} \quad [5].$$

If you represent the integral in **I**(ω) in Dirac notation we can see that it represents a matrix element of the inverse operator,

$$\int dx\, O^{-1}(x)=\langle \nu|C^{-1}|\nu\rangle .$$

where $|\ \nu\ \rangle$ = const. Is the orientational distribution function for an isotropic medium. The derivative spectrum is given as follows,

$$I'(\omega)=\Re[\frac{\delta}{\delta\theta}\langle\nu|C^{-1}|\nu\rangle]$$
$$\frac{\delta}{\delta\theta}(CC^{-1})=\frac{\delta}{\delta\theta}1=0$$
$$[\frac{\delta}{\delta\theta}C]C^{-1}+C\frac{\delta}{\delta\theta}C^{-1}=0$$
$$\frac{\delta}{\delta\theta}C^{-1}=-C^{-1}\frac{\delta}{\delta\theta}CC^{-1} \quad .$$
$$\frac{\delta}{\delta\theta}\langle\nu|C^{-1}|\nu\rangle=\left\langle\nu\left|\frac{\delta}{\delta\theta}C^{-1}\right|\nu\right\rangle=-\left\langle\nu\left|C^{-1}\frac{\delta}{\delta\theta}CC^{-1}\right|\nu\right\rangle$$
$$define\,|U\rangle=C^{-1}|\ \nu\ \rangle$$
$$\frac{\delta}{\delta\theta}\langle\nu|C^{-1}|\nu\rangle=\left\langle U\left|\frac{\delta C}{\delta\theta}\right|U\right\rangle$$

which shows that both **I**′(ω) and **I**(ω) can be given analytically, so the ESR spectra can be easily calculated over the entire molecular motion rates range. There are a few properties of this lineshape to note at this point. First we have that the profile of the peak is preserved when we decrease the ratio of the diffusion coefficient to the magnitude of the magnetic anisotropy, that is D/|Δ|, if the frequency azis is properly magnified [5]. This suggests that the linshape is asymptotic independent of the rate of motion. Further, it is shown in Moro and Segre's paper on asymptotic lineshapes that when analysing the derivative spectrum the peak is associated with D/|Δ| while the broadening is due to the ratio of the diffusion process to the effective relaxation rate [5]. The conclusion of the Moro-Segre paper is an

important driving force for research into reducing noise in lineshapes to obtain better experimental results and therefore more accurate information about the sample.

INFORMATION GEOMETRY

One essential question we must ask is: what is information geometry and how can we use it to find probabilities for parameters of a distribution. The probability manifolds of the Riemann space are given by a metric called the fisher metric. To begin parameter estimation for the ESR lineshape, I calculated the Fisher Information Metric for the Gaussian distribution [2]. Then by computing the derivatives of the Fisher metric I can define the Hessian which contains information on critical points [9].

I have computed the Fisher metric for the Gaussian, and future research will include the Hessian and its applications. I also have estimated the parameters of the Gaussian using Bayesian analysis. What follows is the application of information geometry and Bayesian estimation. Since we are given data that is dependent on an equation, with two parameters we will initially want to estimate. This is a perfect example of a use for Bayes theorem.

Fisher Information Metric

The Fisher information matrix tells of the amount of information that an observable variable, X contains about an unknown parameter θ when the probability of X depends on θ. The fisher information is also known as the variance of the score and is defined in the following way,

$$I(\theta) = -E[(\frac{\delta^2}{\delta\theta^2}\log f(X;\theta))^2\theta] \text{ [8],}$$

For my applications I consider the Gaussian function dependent on parameters μ and σ. I define a model Gaussian in which I vary the parameters to test against a signal Gaussian. The signal has unknown defined parameters which I seek to determine.

$$S(\sigma_0, \mu_0|x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{\frac{-(x-\mu_0)^2}{2}\sigma_0^2} + nN(\mu_n = 0, \sigma_n)$$

.

$$M(\sigma, \mu|x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2}\sigma^2}$$

The noise is included in the signal and is scalable by some defined number n. The probability that a system exists in the i-th substate is defined by the partition function as follows,

$$P_i = \frac{1}{Z} e^{\frac{-(S_i - M_i)^2}{2\sigma_n}},$$

where sigma is the square root of the variance of the random noise and,

$$Z = \sum_i e^{\frac{-(S_i - M_i)^2}{2\sigma_n}} \quad [9].$$

In order to determine the Fisher matrix I need to calculate derivatives of the partition function. The following derivatives are presented without derivation,

$$\frac{\delta P_i}{\delta \mu} = \frac{1}{Z} \frac{(S_i - M_i)}{\sigma_n^2 \sqrt{2\pi}\sigma} \frac{(x-\mu)}{\sigma^2} e^{\frac{-(S_i - M_i)^2}{2\sigma_n^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} - \frac{1}{Z^2} e^{\frac{-(S_i - M_i)^2}{2\sigma_n^2}} \sum_i \frac{(S_i - M_i)}{\sigma_n^2 \sqrt{2\pi}\sigma^3} \frac{(x-\mu)}{\sigma^2} e^{\frac{-(x-\mu)^2}{2\sigma^2}} e^{\frac{-(S_i - M_i)^2}{2\sigma_n^2}},$$

$$\frac{\delta P_i}{\delta \sigma} = \frac{1}{Z} \frac{(S_i - M_i)}{\sigma_n \sqrt{2\pi}} e^{\frac{-(S_i - M_i)^2}{2\sigma_n^2}} \left[\frac{(x-\mu)^2}{\sigma^4} e^{\frac{-(x-\mu)^2}{2\sigma^2}} - \frac{1}{\sigma^2} fun\, e^{\frac{-(x-\mu)^2}{2\sigma^2}} \right]$$

$$; \quad -\frac{1}{Z^2} e^{\frac{-(S_i - M_i)^2}{2\sigma_n^2}} \left[\frac{\sum_i (S_i - M_i)^2}{\sigma_n^2} \sqrt{2\pi} e^{\frac{-(S_i - M_i)^2}{2\sigma_n^2}} \left[\frac{(x-\mu)^2}{\sigma^4} e^{\frac{-(x-\mu)^2}{2\sigma^2}} - \frac{1}{\sigma^2} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \right] \right]$$

and last,

$$\frac{\delta P_i}{\delta n} = \frac{(S_i - M_i)}{\sigma_n^2} n e^{\frac{-(S_i - M_i)^2}{2\sigma_n^2}}.$$

Now form of the fisher information matrix,

$$\begin{bmatrix} l_\sigma l_\sigma & l_\sigma l_\mu & l_\sigma l_n \\ l_\mu l_\sigma & l_\mu l_\mu & l_\mu l_n \\ l_n l_\sigma & l_n l_\mu & l_n l_n \end{bmatrix}.$$

if we define $l_\sigma(i) = \frac{\delta}{\delta \sigma} \ln P_i = \frac{1}{P_i} \frac{\delta}{\delta \sigma} P_i$ then our Fisher matrix becomes,

$$\frac{1}{P_i^2}\begin{bmatrix} \frac{\delta P_i}{\delta\sigma}\frac{\delta P_i}{\delta\sigma} & \frac{\delta P_i}{\delta\sigma}\frac{\delta P_i}{\delta\mu} & \frac{\delta P_i}{\delta\sigma}\frac{\delta P_i}{\delta n} \\ \frac{\delta P_i}{\delta\mu}\frac{\delta P_i}{\delta\sigma} & \frac{\delta P_i}{\delta\mu}\frac{\delta P_i}{\delta\mu} & \frac{\delta P_i}{\delta\mu}\frac{\delta P_i}{\delta n} \\ \frac{\delta P_i}{\delta n}\frac{\delta P_i}{\delta\sigma} & \frac{\delta P_i}{\delta n}\frac{\delta P_i}{\delta\mu} & \frac{\delta P_i}{\delta n}\frac{\delta P_i}{\delta n} \end{bmatrix} \quad [10].$$

Using Matlab I created data using random parameters and noise. Near the optimum parameters the fisher matrix looks like the following,

$$\text{Fisher matrix} = \begin{bmatrix} 1067.5 & 0.2 & -13.8 \\ 0.2 & 0.1 & 0 \\ -13.8 & 0 & 1665.5 \end{bmatrix};$$

with eigenvectors and eigenvalues,

$$\text{Eigenvectors} = \begin{bmatrix} 0.9997 & -0.0230 & -0.0002 \\ 0.0002 & 0.0000 & 1.0000 \\ 0.0230 & 0.9997 & -0.0000 \end{bmatrix}$$

.

$$\text{Eigenvalues} = \begin{bmatrix} 1067.1 & 0 & 0 \\ 0 & 1665.9 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}$$

Finally near the optimum parameters the entropy, S, is equal to approximately 6.6214. Later research will include investigating this matrix in more detail. I will be continuing work on the fisher matrix and its derivative which is the Hessian, to determine information about the distribution.

Bayesian Approach

To use Baye's Theorem to estimate the parameters we need to find the partition function of the distribution. I intend to show how one can find the maximum of entropy of the Signal minus Model distribution. This corresponds to a maximization of the probability that the parameters in the model match the signal. By doing this I was able to get results with a small percent error. To perform this approximation I define the partition function as above. Then by creating a matrix of values for the partition function, each column of the matrix has a different value of mu in the model distribution, while sigma varies from row to row. By performing this calculation I was able to make a 3 dimensional plot of the probability surface with variables mu and sigma. The maximum corresponds to the model parameters matching the signal parameters.

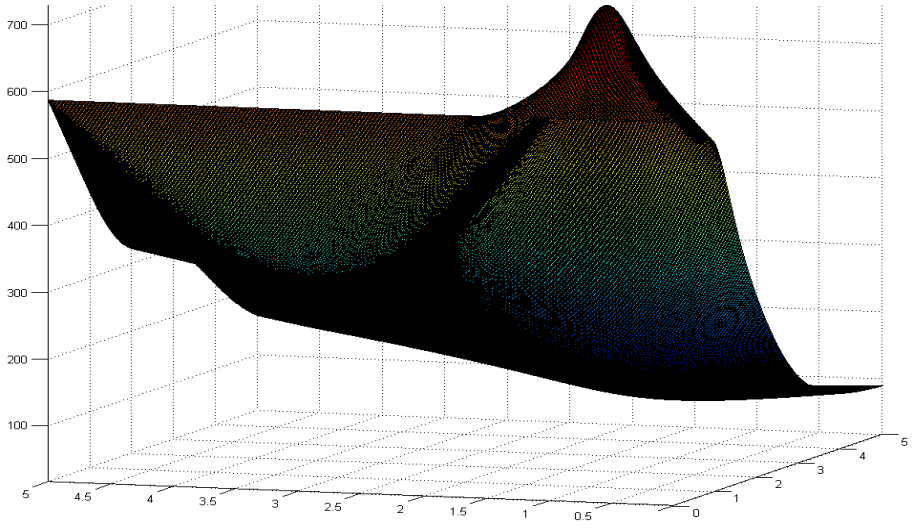

Fig. 2: This is the surface plot of the Partition function,

note the clear peak where the model parameters match the signal.

Thereby making accurate predictions of the parameters of the Signal Gaussian.

CONCLUSION

By beginning at the beginning and explaining the concepts of ESR, I then proceeded to procure the theory behind experimental results. I have shown how a lineshape is created through ESR in two different limits of relaxation time. The lineshape, we know, contains information about the sample, however by using information geometry I hope to show how much information can be taken from a lineshape. The Fisher information matrix that I have calculated shows analytically how much information is contained in each parameter. Then by applying Baye's theorem to the partition function of the signal-model system I was able to make accurate predictions of the parameters of the signal.

I will be continuing this research at the University at Albany with Professor Keith Earle. I will be extending my focus to the more complicated Student's t distribution, then deriving the Hessian matrix as well, which can be difficult to analytically find.

REFERENCES

[1] Shun'ichi Amari, Hiroshi Nagaoka. *Methods of information geometry*, Translations of mathematical monographs; v. 191, American Mathematical Society, 2000

[2] C. Cafaro. The Information Geometry of Chaos. Doctoral Dissertation SUNY Albany, 2008.

[3] Spin Dynamics, Malcolm H. Levitt, Wiley, 2001

[4] F. Morgan. Riemannian Geometry. A K Peters, 1998.

[5] G. Moro and U. Segre. Ultraslow Motions and Asymptotic Lineshapes in ESR. J. Magn. Reson., 83:65-78, 1988.

[6] K. Earle, Parameter Sensitivity of the Saturated Lineshape. 2010

[7] Frieden, B. Roy. *Science from Fisher Information: A Unification*. Cambridge Univ. Press, 2004.

[8] P. H. Heckmann and E. Trabert. Introduction to the Spectroscopy of Atoms, North-Holland Physics, 1989.

[9] Huang, Kerson, "Statistical Mechanics", John Wiley & Sons, New York, 1967.

[10] T. L. Squires. An Introduction to Electron Spin Resonance. Academic Press Inc. 1964.

[11] Wikipedia contributors. "Electron paramagnetic resonance." *Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia, 9 May. 2012. Web. 18 May. 2012.

[12] Wikipedia contributors. "Fisher information." *Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia, 18 Apr. 2012. Web. 18 May. 2012

[13] Wikipedia contributors. "Information geometry." *Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia, 26 Apr. 2012. Web. 18 May. 2012.

[14] G. Zwan. Derivation of the Redfield relaxation equation. 2003.