The Bedlewo Conference

Control, Constraints and Quanta

Noise Models and Controlling for Minimal Decoherence

Roger Brockett

October 11-16, 2007

Engineering and Applied Sciences

Harvard University

Outline of the talk

1. Some aspects of quantum statistical mechanics.

2. The density matrix and its equation of evolution.

3. Quantum spin and the Lindblad operators[1] of a dissipative quantum process. Coherence and decoherence in NMR.

4. NMR spectroscopy and why control is important.

5. Two dimensional NMR and the Cramer-Rao bound.

1. I was kindly informed at this meeting that although in the NMR literature these are usually refers to as Lindblad operators there are good reasons for calling them Kossakowski-Lindblad operators.

Basic example: Harmonic oscillator

$$h(x,p) = \frac{1}{2}(p^2/m + kx^2)$$

$$i\frac{\partial\psi}{\partial t} = \left(\frac{1}{2m}\left(\frac{i\partial}{\partial x}\right)^2 + \frac{1}{2}kx^2\right)\psi$$

or

$$\frac{\partial\psi}{\partial t} = i\frac{\partial^2\psi}{\partial x^2} - ikx^2\psi$$

We might seek to control this by manipulating the potential such as in the case

$$\frac{\partial\psi}{\partial t} = i\frac{\partial^2\psi}{\partial x^2} - ki(x-u)^2\psi$$

The u^2 term produces no measurable effect

This is, abstractly, $\dot{z} = Az + uBz + u^2z)$

The hamiltonian represents the total energy of a classical system, kinetic plus potential. Consistent with the equation

$$i\frac{\partial \psi}{\partial t} = h(p, x)\psi$$

we see that the eigenvalues of the operator h have the units of energy. Thus the eigenplanes of the previous slide can be labeled by using energy ($E = h\nu$) and a control that shifts the system from one eigenplane to another has a certain energy requirement associated with it. That is, the change of energy level in the system must be accounted for by the energy supplied by the control.

Moreover, by the assignment of eigenvalues and eignespaces, h defines a map

$$H : \mathbb{G}(2, \mathcal{H}) \to \mathbb{R}$$

which assigns an energy to each eigenplane.

The Hilbert Space

The Hilbert space which occurs in quantum mechanics is a space of square integrable functions mapping the set of possible configurations into the complex numbers. For the harmonic oscillator the Hilbert space is infinite dimensional. But in many interesting cases it is effectively finite dimensional.

John von Neumann

Paul Dirac

Off Topic but Interesting….

Compare with the conditional density equation for the system

$$dx = dw \;\; ; \;\; dy = xdt + d\nu$$

from linear estimation theory

Diffusion-like
Not wave-like

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\left(\frac{\partial^2}{\partial x^2} - x^2\right)\rho + (\dot{y} - x)^2\rho$$

Because the $\dot{\nu}^2$ term only produces a renormalization it can be ignored to get

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\left(\frac{\partial^2}{\partial x^2} - x^2\right)\rho + (\dot{y} - x)^2\rho + \dot{y}x\rho$$

What makes estimation easy makes control impossible.

The Density Matrix

For multi-particle systems or single particles which exist in a superposition of eigenstates, the density matrix provides a useful alternative to the Schrödinger equation. Think $\rho = \psi\psi^*$ and

$$\rho = i[H, \rho]$$

For multi-particle systems the off-diagonal elements represent the degree of coherence between different particles.

ρ is Hermitean and nonnegative definite

$$\rho = \sum \psi_i \psi_i^*$$

with the ψ_i being eigenstates.

Controlling an Ensemble with a Single Control

The actual problem involves many copies with almost the same dynamics

$$dx_1/dt = A(u)x_1+Bw_1$$

$$dx_2/dt = A(u)x_2+Bw_2$$

$$\ldots\ldots\ldots\ldots$$

$$dx_n/dt = A(u)x_n+Bw_n$$

$$y=(cx_1 + cx_{2+} \ldots+x_n) + n$$

Ensemble description
With noise terms

The system is not controllable or observable. There are something like 10^{23} copies of the same, or nearly the same, system. We can write an equation for the sample mean of the x's, for the sample covariance, etc. Multiplicative control is qualitative different from additive.

Given that $H = H_0 + uH_1$ we see that abstractly

$$\dot{\rho} = i[H_0 + uH_1, \rho]$$

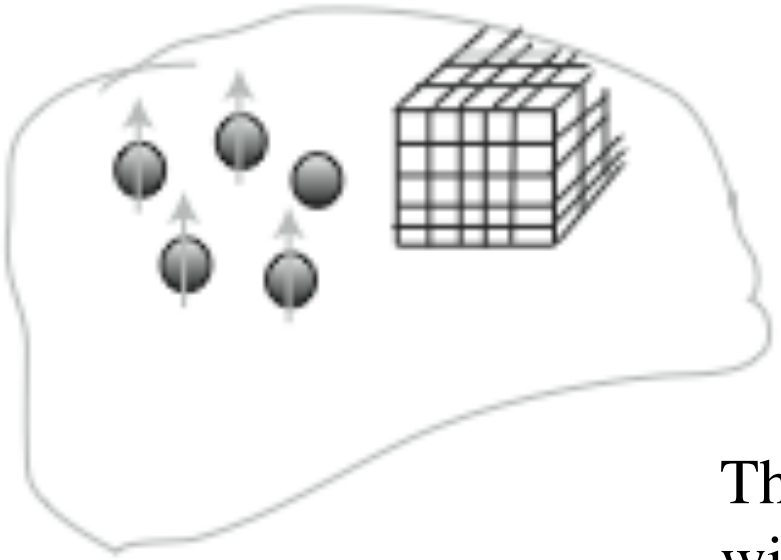
Isospectral flow!

is of the form

$$\dot{x} = (A + uB)x$$

Of course A and B can be expressed in terms of Kronecker products. In this notation the term "super operator" is often used in the physics literature. This form is convenient when the effects of noise are taken into account.

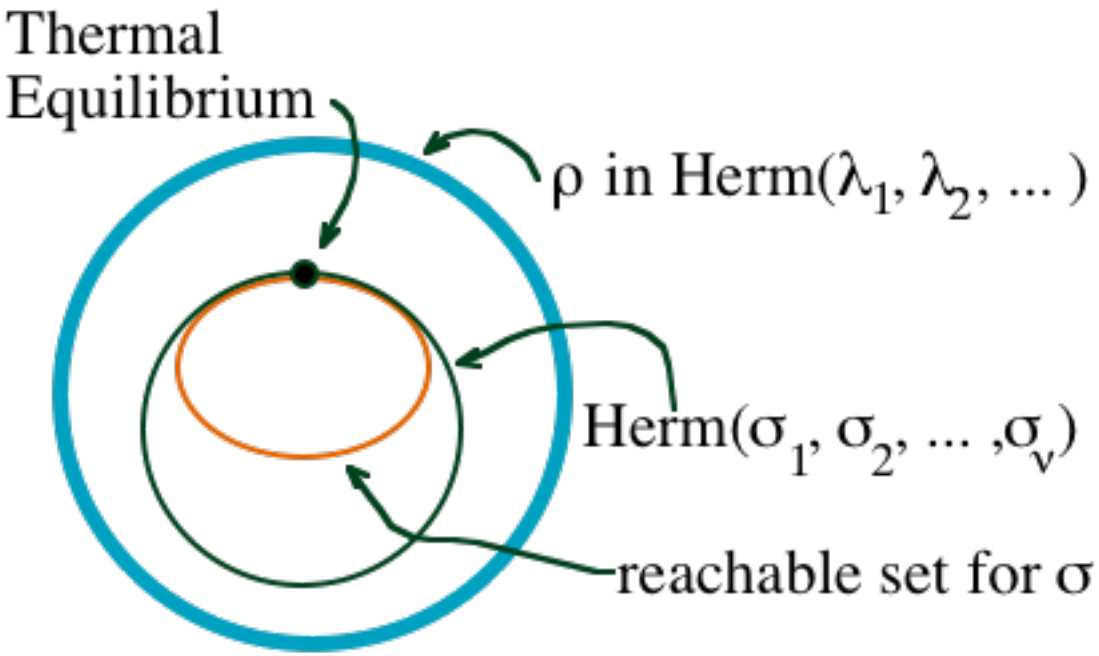
Wanted and Unwanted Dynamics



The lattice dynamics interact with the spin dynamics. We only hope to understand the spin Dynamics.

The Reduced Density Equation

For tractability, separate the "lattice dynamics" from the spin dynamics, replacing the former by an effective random term. The resulting flow is no longer isospectral but is asymptotically stable to an equilibrium consistent with the Boltzmann distribution.



Think: blue is infinite dimensional and isospectral, green is finite dimensional (spin only Hilbert space) and isospectral. Orange is spin only, finite dimensional, not isospectral, the "master equation".

The Lindblad model for accounting for the lattice dynamics

$$\dot{\rho} = [iH, \rho] + [A, [A, \rho]]$$

Why this form? Recall that if we had a lossless linear system $\dot{x} = Sx$ and added a noise term $B\dot{w}$ we need to add a dissipation term at the same time to be consistent with the second law of thermodynamics. This term must take the form

$$dx = (S - BB^T)xdt + Bdw$$

$$\frac{d}{dt}\mathcal{E}x = (S - BB^T)\mathcal{E}x$$

In this way $\dot{\rho} = [iH + A\dot{w}, \rho]$ goes to the averaged equation above.

Consider the pair related by $z = e^{-At}x$.

$$\dot{x} = Ax + \sum u_i(t) B_i x$$

$$\dot{z} = \sum u_i(t) e^{-At} B_i e^{At} z$$

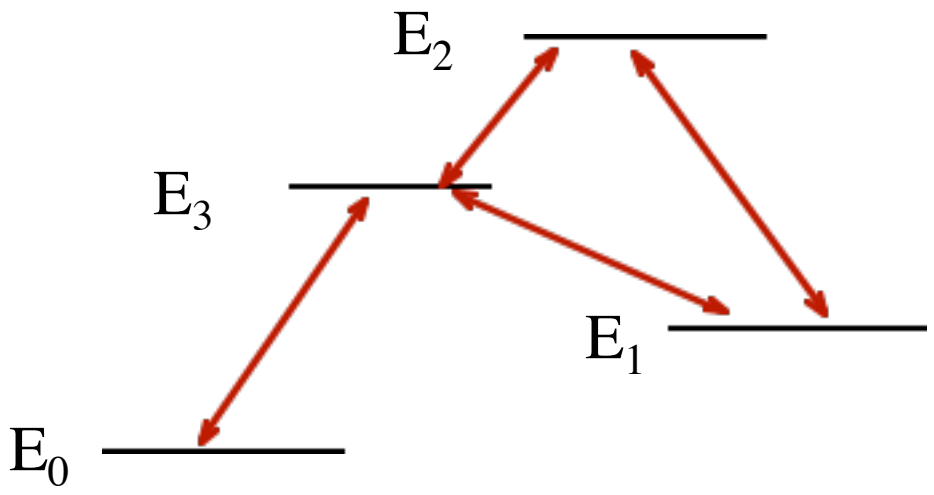
In the first case we would want to compute the Lie algebra $\{A, B_i\}_{LA}$. However, if A is skew adjoint and if it eigenvaslues are large compared with the size of the $u_i B_i$ then $e^{-At} B_i e^{At}$ is almost periodic and we can consider limiting ourselves to sinusoidial and "resonant" u's. Think **control by frequency selection leading to control by vector field selection** using

$$\int \sum u_i(t) e^{-At} B_i e^{At} z(t) dt \approx \sum \int u_i(t) e^{-At} B_i e^{At} dt z$$

Thinking like a Physicist and not a Geometer

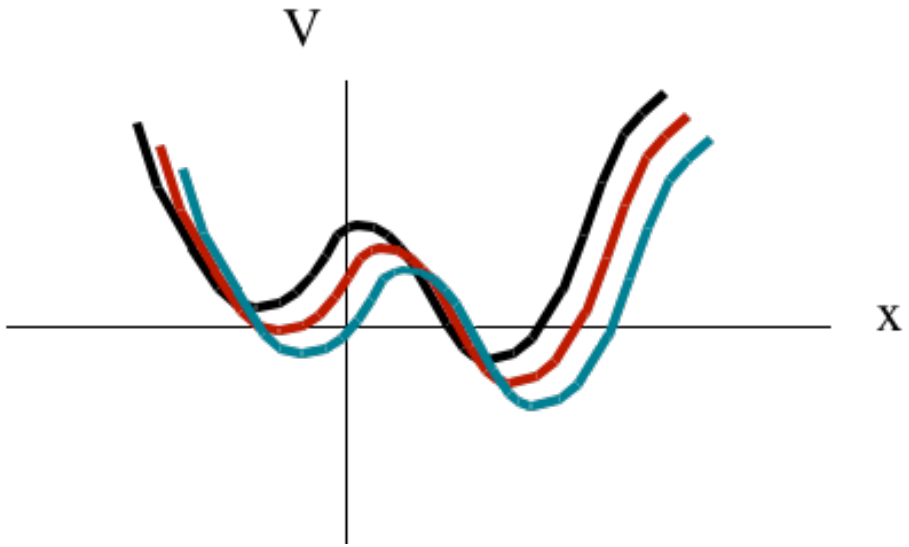
Often examples involve systems that take on only a few states so that the Schrodinger equation is thought of as being finite dimensional. These problems arise in the study of spin systems, but are also are good approximations for a number of systems in which electromagnetic radiation interacts with atomic systems. laser systems, spectroscopy, …

Instead of explaining behavior in terms of a selection of vector fields, explain it in terms of a selection of frequencies (which then determine vector fields through averaging).



Non Quadratics are more Controllable

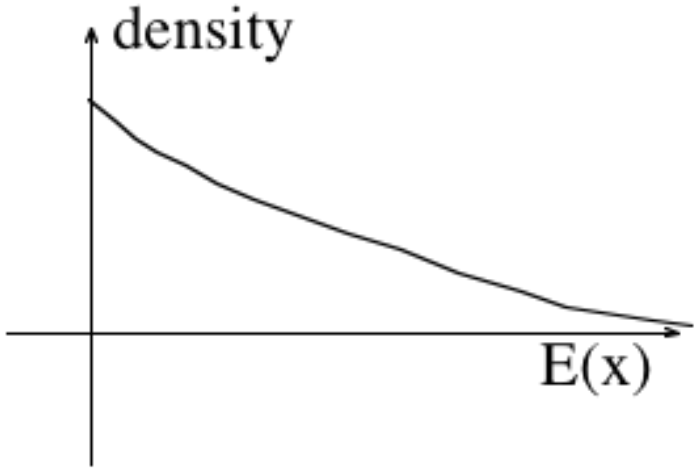
The harmonic oscillator describes much in quantum mechanics but control via shifting the minimum point of the potential is not effective because the Lie algebra is only four dimensional. Non-quadratic potentials are necessary to increase the dimension of the Lie algebra



Boltzmann Distribution for a Physical System in Equilibrium at Temperature T

$$\rho(x)=(1/Z)\exp-(E(x)/2kT)$$

Because magnetic moments that are aligned with the magnetic field have a little less energy than those opposing it, the Boltzmann distribution implies they are favored.



Stochastic Equations Supporting Boltzman

$$dx = Axdt + \sum_{i=1}^{m} B_i x dw_i$$

Fact 1: If x satisfies a linear sample path equation then

$$\frac{d}{dt}\mathcal{E}x(t) = A\mathcal{E}x(t)$$

Proof: Take expectations of both sides of (1). Use the fact that $\mathcal{E}g(x)dw = 0$.

Fact 2: If x satisfies a linear sample path equation then for each positive integer p the set of moments of order p satisfy a linear homogeneous differential equation.

Introduce the notation $x^{[p]}$, where x is an n-vector and p is a positive integer, to denote the $\binom{n+p-1}{p}$-component vector whose entries are the independent monomials homogeneous of degree p. That is,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad ; \quad x^{[p]} = \begin{bmatrix} x_1^p \\ x_1^{p-1} x_2 \\ \vdots \\ x_n^p \end{bmatrix}$$

If $\dot{x} = Ax$, then it is easy to see that $x^{[p]}$ also satisfies a linear equation so that we have, for $A_{[p]}$ suitably defined,

$$\frac{d}{dt} x^{[p]} = A_{[p]} x^{[p]}$$

The Relevant Calculus for Stochastic Systems

We take this as a definition of $A_{[p]}$. In applying the Itô rule to (3.1) we get

$$dx^{[p]} = A_{[p]}x^{[p]}dt + \sum B_{i[p]}x^{[p]}dw_i + \text{“Itô term”}$$

the Itô term denoting the contributions due to

$$\left\langle \frac{\partial^2 x^{[p]}}{\partial x_i \partial x_j} g_i \ , \ g_i \right\rangle dt$$

in using the Itô rule.

The Density Equation from Statistical Mechanics

Because the entries of $x^{[p]}$ are homogeneous of degree p, their second derivatives are homogeneous of degree $p-2$. Since $B_i x dw_i$ is linear in x, we see that the Itô term is homogeneous of degree p in $(x_1, x_2, \ldots, x_n)$ or, what is the same, linear in $x^{[p]}$. Thus after taking expectation, we get

$$\frac{d}{dt}\mathcal{E}x^{[p]} = L_{[p]}\mathcal{E}x^{[p]}$$

for some $L_{[p]}$. One can show that $L_{[p]}$ is given by

$$L_p = (A - \frac{1}{2}B^2)_{(p)} + \frac{1}{2}(B_{(p)})^2$$

Proof: We introduce the notation $x^{[p]}$, where x is an n-vector and p is a positive integer, to denote the $\binom{n+p-1}{p}$-component vector whose entries are the independent monomials homogeneous of degree p. That is,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad ; \quad x^{[p]} = \begin{bmatrix} x_1^p \\ x_1^{p-1}x_2 \\ \vdots \\ x_n^p \end{bmatrix}$$

If $\dot{x} = Ax$, then it is easy to see that $x^{[p]}$ also satisfies a linear equation so that we have, for $A_{[p]}$ suitably defined,

$$\frac{d}{dt}x^{[p]} = A_{[p]}x^{[p]}$$

We take this as a definition of $A_{[p]}$. In applying the Itô rule to (3.1) we get

$$dx^{[p]} = A_{[p]}x^{[p]}dt + \sum B_{i[p]}x^{[p]}dw_i + \text{“Itô term”}$$

the Itô term denoting the contributions due to

$$\left\langle \frac{\partial^2 x^{[p]}}{\partial x_i \partial x_j} g_i \ , \ g_i \right\rangle dt$$

in using the Itô rule.

The Lindblad model for accounting for the lattice dynamics

$$\dot{\rho} = [iH, \rho] + [A, [A, \rho]]$$

Why this form? Recall that if we had a lossless linear system $\dot{x} = Sx$ and added a noise term $B\dot{w}$ we need to add a dissipation term at the same time to be consistent with the second law of thermodynamics. This term must take the form

$$dx = (S - BB^T)xdt + Bdw$$

$$\frac{d}{dt}\mathcal{E}x = (S - BB^T)\mathcal{E}x$$

In this way $\dot{\rho} = [iH + A\dot{w}, \rho]$ goes to the averaged equation above.

Decoherence-free subspaces in Lindblad model

The structure of A reflects the dynamics,resonances, etc. of the lattice which will structure the effective power spectrum of the the interaction terms. It can happen that it is possible to structure the interaction so that $[A, [A, \cdot]]$ has a null space. The kernel of this operator then defines a subspace that is "decoherence free" such supspaces play a roll in some discussions of quantum information. Problems involving the transfer of states which take advantage of such a subspace have been studied.

This could be the subject of a talk in itself.

The Concept of Quantum Mechanical Spin

First postulated as property of the electron for the purpose of explaining aspects of fine structure of spectroscopic lines, (Uhlenbeck-Gouldsmit). Electron spin was first incorporated into a Schrodinger-like description of physics by Pauli and then treated in a definitive way by Dirac. Spin itself is measured in units of angular momentum as is Plank's constant. The gyromagnetic ratio links the angular momentum to an associated magnetic moment which, in turn, accounts for most measurable aspects of spin. Protons were discovered to have spin in 1927 (Dennison) and in 1932 Heisenberg wrote a paper on nuclear structure in which the recently discovered neutron was postulated to have spin and a magnetic moment.

In this talk it is the net spin of the **nucleus** that plays the basic role, although there is an important indirect effect (the chemical shift) involving electronic spin which also plays a role.

Angular Momentum and Magnetic Moment

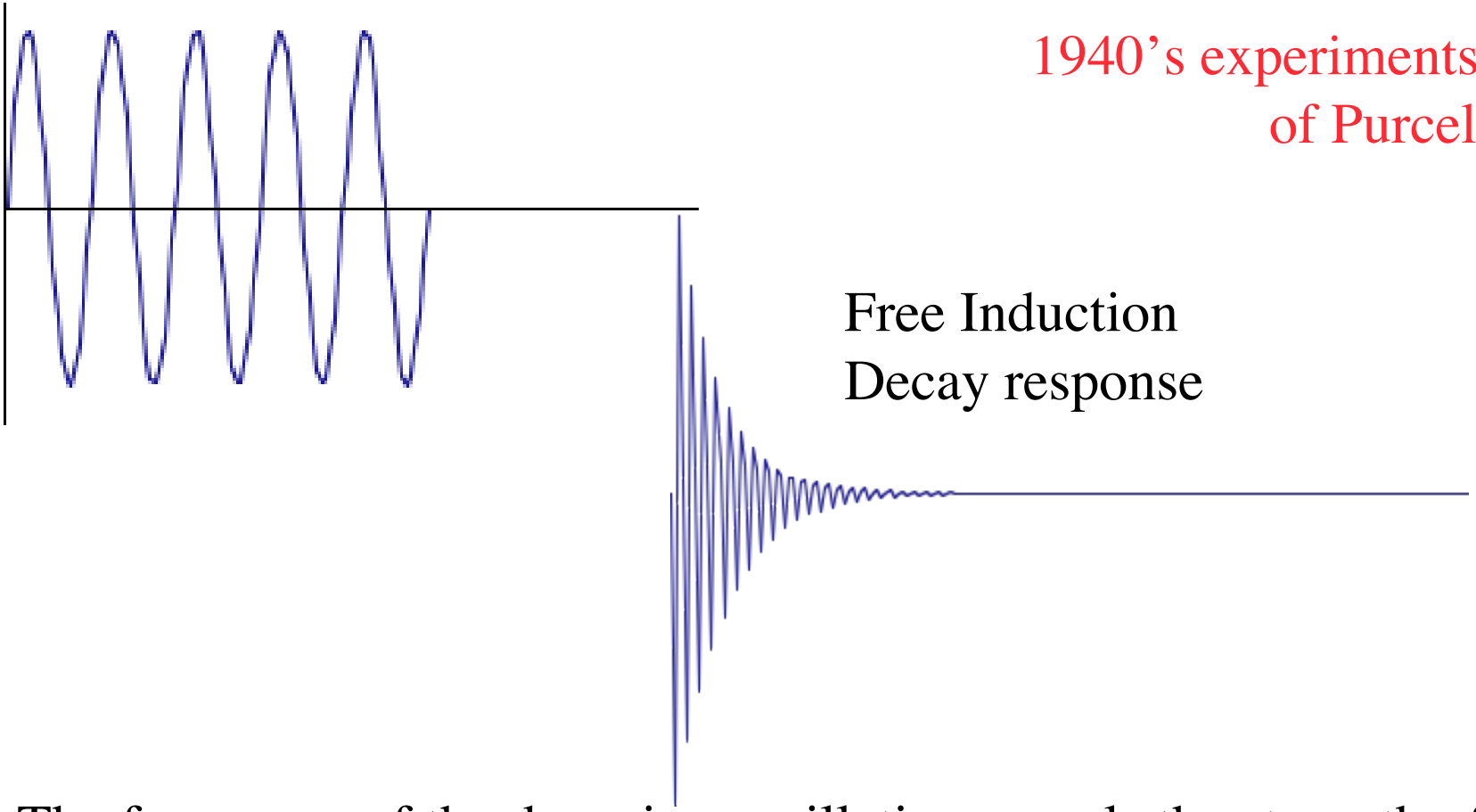
Spin (angular momentum) relative to a fixed direction in space is quantized. The number of possible quantization levels depends on the total momentum. In the simplest cases (spin one-half systems) the total momentum is such that the spin can be only plus or minus 1/2 in suitable units. Systems that consist of a collection of n such states give rise to a Hermitean density matrix of dimension 2^n .by 2^n.

Wolfgang Pauli

Werner Heisenberg

Measured Response: If QM is Unitary, why the Decay?

Radio Frequency Pulse input



1940's experiments
of Purcell

Free Induction
Decay response

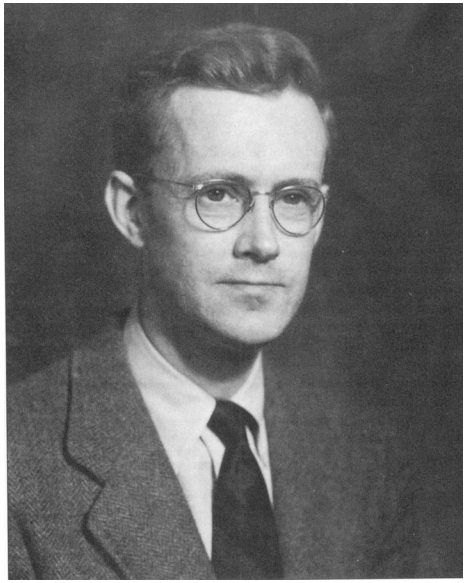
The frequency of the decaying oscillation reveals the strength of the effective magnetic field felt by the spinning nucleus. The decay is caused by decoherence.

The Block Equations

Bloch
Nuclear Induction

Bloch constructed this important phenomenological equation, valid in a rotating coordinate system, which applies to a particular type of time varying magnetic field.

$$dx_r/dt = (A+uB)x+ub$$



Purcell
Absorption

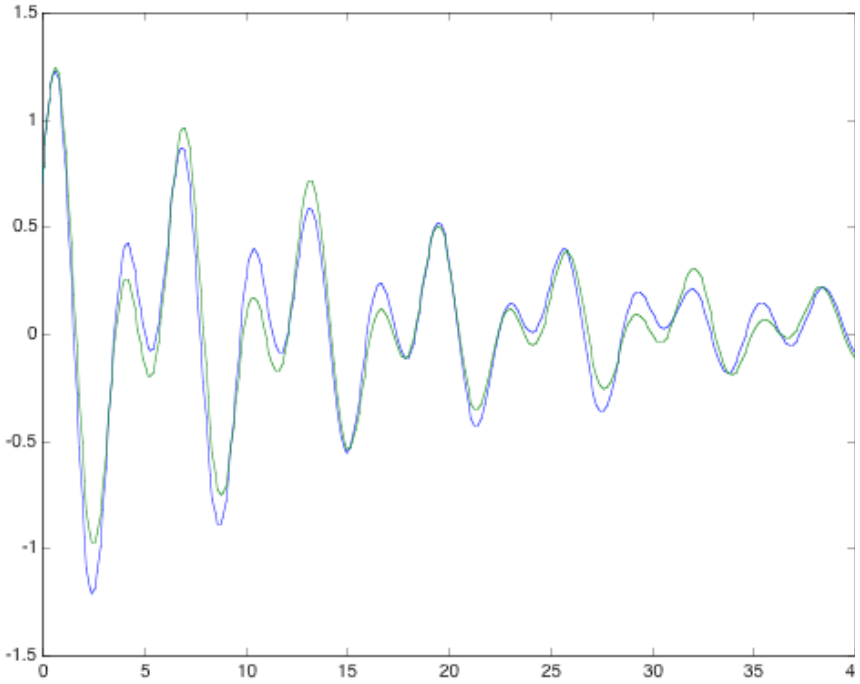
$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 & u & 0 \\ -u & -1 & f \\ 0 & -f & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$y = x_2 + n$$

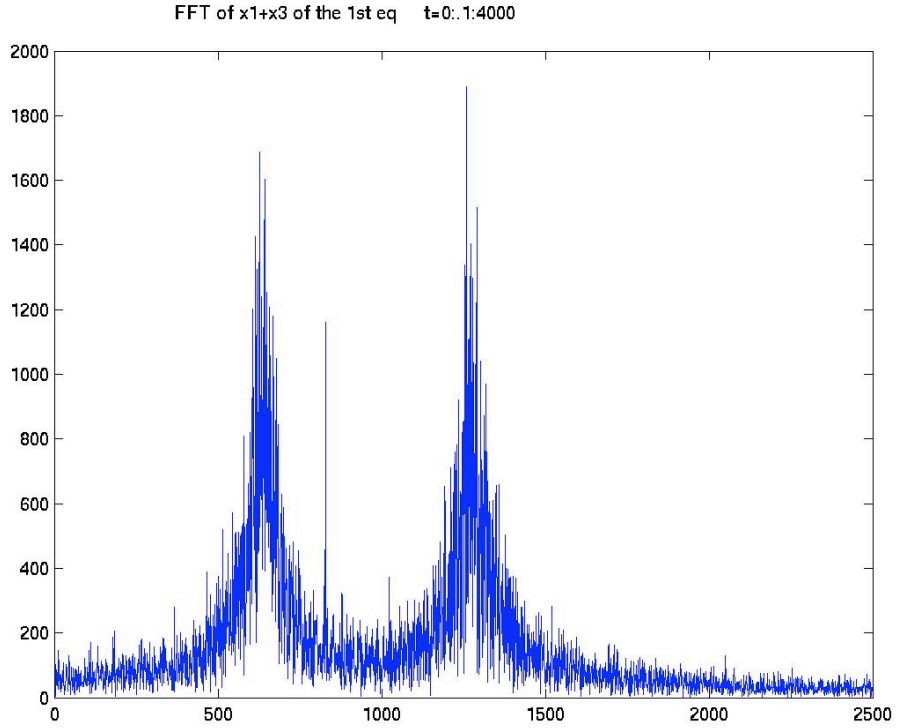
The Linearization Dilemma

Small input makes linearization valid but gives small signal-to-noise ratio. Large input give higher signal-to-noise ratio but makes nonlinear analysis necessary.

response

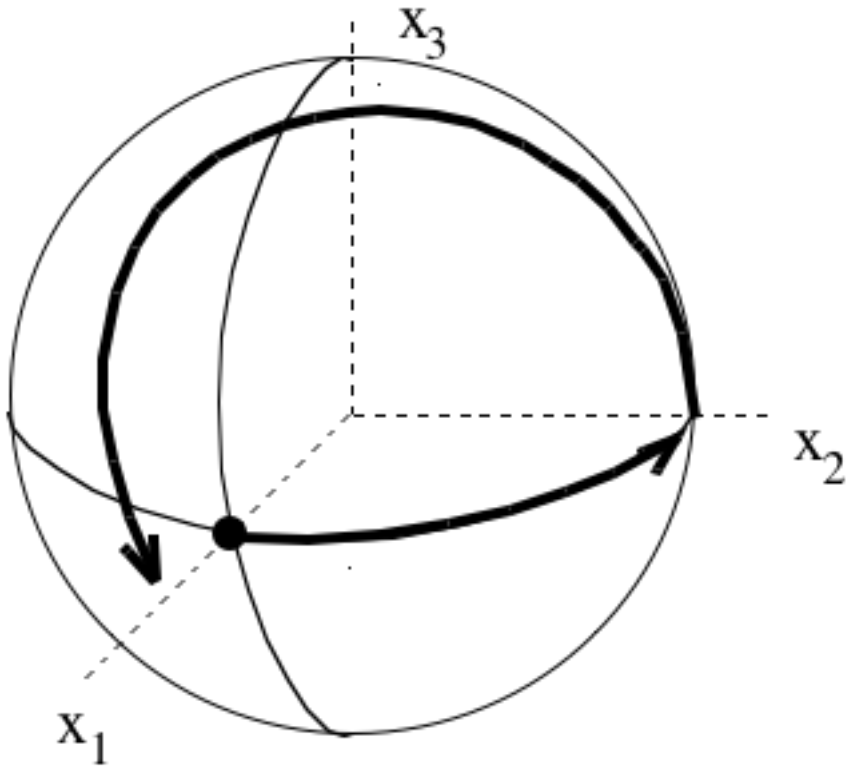


Fourier transform



Qualitative Analysis Based on the Mean

If we keep u at zero there is no signal. If we apply a pulse, rotating the equilibrium state from $x_1 = 1$, $x_2=0$,$x_3=0$ to $x_1 = 0$, $x_2=1$,$x_3=0$, Then we get a signal that reveals the size of f. The actual signal with noise present can be expected to have similar behavior.



Why are Radio Frequency Pulses Effective in moving x?

$$dx/dt = (A+u(t)B)x$$

Let z be exp(-At)x so that the equation for z takes the form

$$dz/dt = u(t)e^{-At}Be^{At}z(t)$$

If Ax(0)=0 and if the frequency of u is matched to the frequency of exp(At) there will be secular terms and the solution for z will be approximated by z(t) = exp(Ft)x(0). Thus x is nearly exp(At)exp (Ft)x(0).

Decoherence and the Role of the Density Matrix

Each ψ has a phase angle but only $|\psi|$ is related to probability, Thus for a single particle phase is not detectable. However for two weakly interacting particles the relative phase angle matters. The size of the off-diagonals in ρ measures the conherence of the relative phase angles.

Spin (angular momentum) relative to a fixed direction in space is quantized. The number of possible quantization levels depends on the total momentum. In the simplest cases the total momentum is such that the spin can be only plus or minus 1/2. Systems that consist of a collection of such states give rise to a density matrix of dimension 2^n .

The Density Equation from Statistical Mechanics

Typical systems consist of a part we care about and a "lattice" or "bath" that we would like to "average out". In some cases the lattice has structure in the form of lightly damped resonances. The Lindblad model replaces the lattice with a noise term of the form

$$dz = Azdt + Dzdw$$

As we have seen, this implies

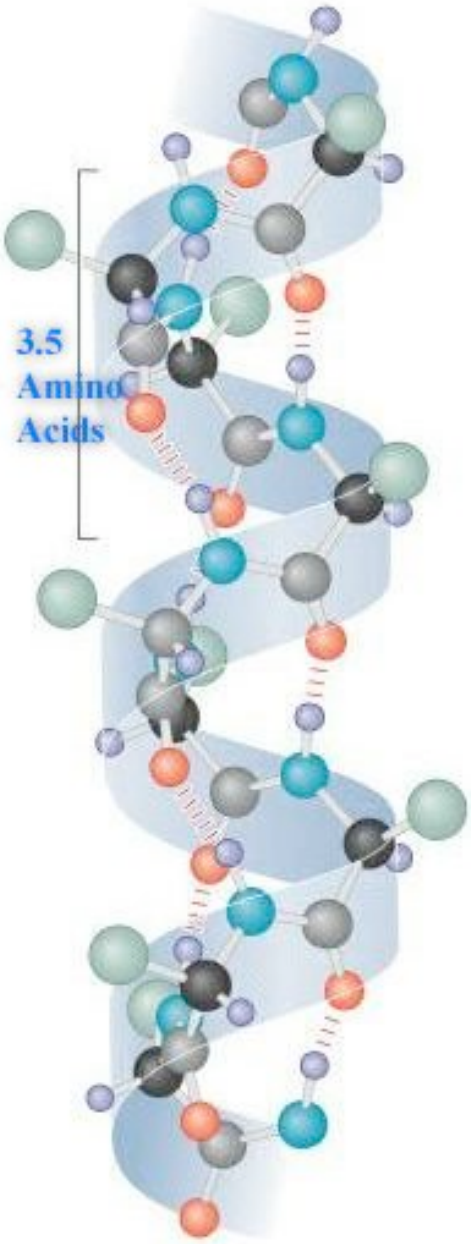
$$\frac{d}{dt}\mathcal{E}zz^T = A\mathcal{E}zz^T + \mathcal{E}zz^T A^T + D(\mathcal{E}zz^T)D$$

Why is the spectroscopy problem so hard and why do people keep working on it?

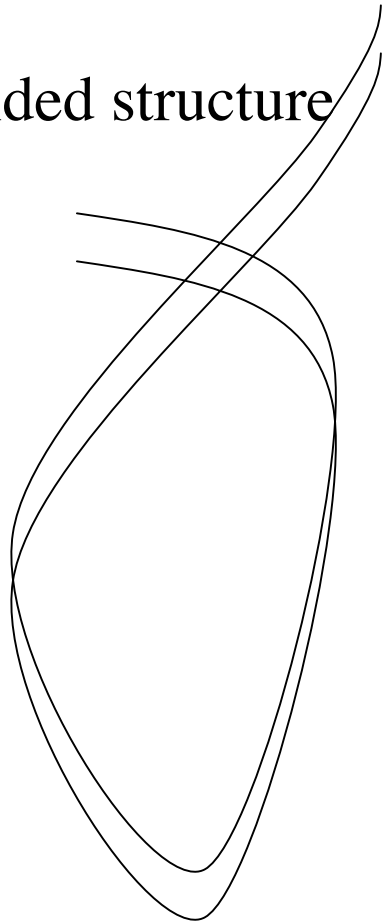
Gespard De Prony

The geometric structure of proteins

An unfolded alpha helix structure.



A folded structure



NMR can help determine the geometry!

Prony's Problem (Slightly Simplified)

Given

$$y(t) = \sum_{i=1}^{n} c_j e^{\lambda_j t} + n(t)$$

for $t \in [a, b]$, and some noise term n, estimate c_n, λ_j.

Suppose that A is an n by m matrix of rank m and that b is in the range space of A. If we observe $b+w$ with w "noise" and wish to estimate x such that $\mathcal{E}Ax = b$, the estimate $x = A^{\#}(b + w)$ is unbiased if $\mathcal{E}w = 0$ and $A^{\#} = (A^T A)^{-1} A^T$.

Background: The Linearized Prony Problem

Now suppose that A is an operator mapping a finite dimensional inner product space into a Hilbert space and that b is in the range space of A. If we observe $b + w$ with w "noise" and wish to estimate x such that $Ax = b$, the estimate $x = A^{\#}(b + w)$ is unbiased if $\mathcal{E}w = 0$ and $A^{\#} = (A^T A)^{-1} A^T$.

If $\mathcal{E}ww^T = \Sigma$ then the variance of x is $(A^{\#})^T \Sigma A^{\#}$. Any other unbiased estimate has a higher variance.

In particular, if $A^\dagger A$ is invertible $x(f) = (A^\dagger A)^{-1} A^\dagger f$ is the minimum variance unbiased estimate. It results in a random variable x with variance $(A^\dagger A)^{-1} A^\dagger \Sigma A (A^\dagger A)^{-1}$, where Σ is the variance of f. If the equations are scaled so that $\Sigma = \kappa^2 I$ then the variance of x is just $\kappa^2 (A^\dagger A)^{-1}$. This result holds even when $A : \mathbb{C}^n \rightarrow \mathbb{H}$ with $\mathbb{H}$ being an infinite dimensional Hilbert space, provided that the operator $A^\dagger A$ is well defined and invertible. if we want an unbiased estimate of b and can observe $y(k) = a^k b + \kappa n(k)$ for $k = 0, 1, ...,$ with $n(k)$ being zero mean, unity variance Gaussians, then the estimate of b is

$$\hat{b} = \left(\sum_{k=0}^{\infty} a^{2k} \right)^{-1} \sum_{k=0}^{\infty} a^k y(k)$$

Let $\dagger$ denote conjugate transpose. It is a familiar fact that if A is an n by m complex matrix of rank m then the choice of x that minimizes $(Ax-b)^\dagger(Ax-b)$ is $x=(A^\dagger A)^{-1}A^\dagger b$. Thus $(A^\dagger A)^{-1}A^\dagger$ is an explicit expression for the Moore-Penrose inverse applicable when columns of A are linearly independent. Suppose f is a random variable taking on values in a complex inner product space $\mathcal{F}$ and that f has a finite mean and variance. If A is a linear operator defined on a second inner product space, $A:\mathcal{X}\rightarrow\mathcal{F}$ it often happens that it is necessary to find an estimator such that the expected value of $Ax(f)$ agrees with the expected value of f and the variance of $x(f)$ is minimal. If f has a Gaussian density, $\rho(f)=(1/N)e^{-\frac{1}{2}(f-m)^\dagger\Sigma^{-1}(f-m)}$ then the role of the Moore-Penrose inverse in solving this problem is well known.

Suppose we are given $y(t_k) + \kappa n(t_k)$ with the $n(t_k)$ being zero mean Gaussian random variables of unit variance and wish to find (possibly complex) vectors $b = [b_1, b_2, ...b_n]^T$ and $\lambda = [\lambda_1, \lambda_2, ..., \lambda_n]^T$ such that

$$y(t_k) = \sum_{i=1}^{n} b_i e^{\lambda_i t_k} \quad ; \quad k = 1, 2, 3, ...$$

As suggested by the Cramer-Rao bound, a key role is played by the linear term in the Taylor series expansion

$$\sum_{i=1}^{n} (b_i + \delta_i) e^{(\lambda_i + \gamma_i) t_k} = \sum_{i=1}^{n} b_i e^{\lambda_i t_k} + \sum_{i=1}^{n} \delta_i e^{\lambda_i t_k} + \sum_{i=1}^{n} b_i \gamma_i t_k e^{\lambda_i t_k} + ...$$

There is a structured matrix involved here.

It is convenient to adopt a vector-matrix notation. Let Λ and b be defined as

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \quad ; \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}$$

and let η denote a row vector whose entries are all ones. Then, in a notation in which the derivative is expressed as a row vector, we have

$$\begin{bmatrix} \frac{\partial \eta e^{\Lambda t} b}{\partial b} & \frac{\partial \eta e^{\Lambda t} b}{\partial \Lambda} \end{bmatrix} = \begin{bmatrix} \eta e^{\Lambda t} & t b^T e^{\Lambda t} \end{bmatrix}$$

If the real parts of λ_i and λ_j are negative then for all nonnegative integers k

$$\int_0^\infty t^k e^{\bar{\lambda}_i t} e^{\lambda_j t} dt = \frac{(-1)^{k+1} k!}{(\bar{\lambda}_i + \lambda_j)^{k+1}}$$

so that

$$\int_0^\infty t^k e^{\bar{\Lambda} t} e^{\Lambda t} dt = W_{k+1}(\Lambda)$$

with

$$W_k(\Lambda) = (-1)^k (k-1)! \begin{bmatrix} \frac{1}{(\bar{\lambda}_1+\lambda_1)^k} & \frac{1}{(\bar{\lambda}_1+\lambda_2)^k} & \cdots & \frac{1}{(\bar{\lambda}_1+\lambda_n)^k} \\ \frac{1}{(\bar{\lambda}_2+\lambda_1)^k} & \frac{1}{(\bar{\lambda}_2+\lambda_2)^k} & \cdots & \frac{1}{(\bar{\lambda}_2+\lambda_n)^k} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{(\bar{\lambda}_n+\lambda_1)^k} & \frac{1}{(\bar{\lambda}_n+\lambda_2)^k} & \cdots & \frac{1}{(\bar{\lambda}_n+\lambda_n)^k} \end{bmatrix}$$

The Fisher Information matrix tells why Prony's problem is difficult

The Jacobian of

$$\phi(b,\lambda)(t) = \eta e^{\Lambda t} b$$

is

$$J = \left[\begin{array}{cc} \frac{\partial \psi}{\partial b} & \frac{\partial \psi}{\partial \Lambda} \end{array} \right] = \left[\begin{array}{cc} \eta e^{\Lambda t} & t b^T e^{\Lambda t} \end{array} \right]$$

The Moore-Penrose inverse of J maps $L_2[0,\infty)$ to $\mathbb{C}^{2n}$ and is given by

$$J^{\#}(\cdot) = F^{-1} \int_0^{\infty} \left[\begin{array}{c} e^{\bar{\Lambda} t} \eta^T \\ t e^{\bar{\Lambda} t} \bar{b} \end{array} \right] (\cdot) dt$$

where

$$F = \left[\begin{array}{cc} W_1 & W_2 * (\eta^T b^T) \\ W_2 * (\bar{b}\eta) & W_3 * (\bar{b} b^T) \end{array} \right]$$

The matrix F is positive definite if and only if the λ_i are distinct and all the components of b are nonzero.

It is nonsingular if b has nonzero entries and the eigenvalues are distinct

Taking the partial derivative of q with respect to b and interpreting the derivative as a column vector yields

$$\frac{1}{2}\frac{\partial q}{\partial b} = \int_0^\infty -e^{\bar{\Lambda}t}\eta^T(y(t) - \eta e^{\Lambda t}b)dt =$$

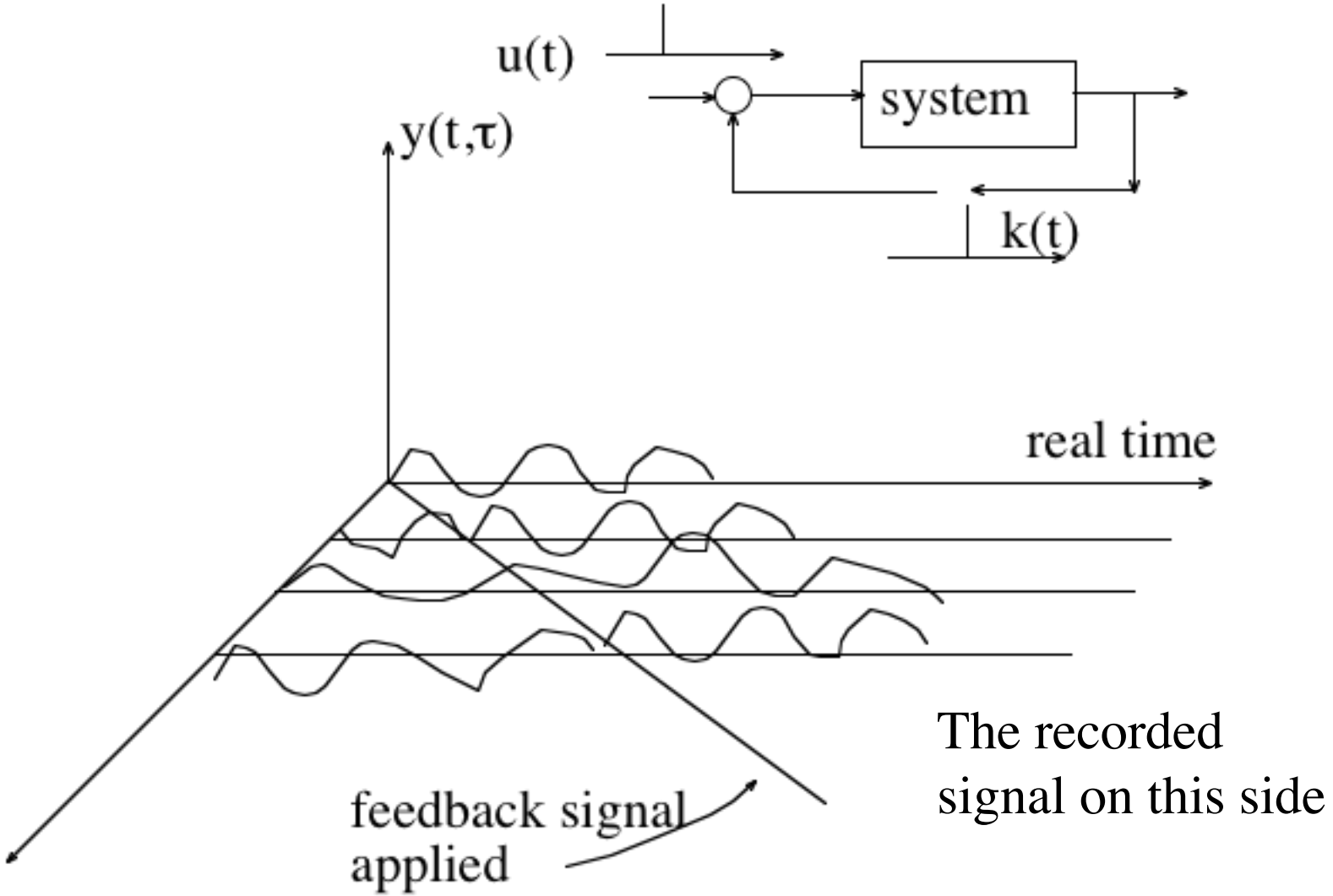
$$W_1 b - \int_0^\infty e^{\bar{\Lambda}t}\eta^T y(t)dt$$

The Fisher information matrix determines a lower bound on the error variance associated with identification. If this matrix is poorly conditioned in the sense that the ratio of its largest to its smallest eigenvalue is large, then some aspects of the system will be much less well characterized than others. For example, if the Gramian of $\psi(b, \lambda)(t) = b_1 e^{\lambda_1 t} + b_2 e^{\lambda_2 t}$ is evaluated at $\psi(t) = e^{-t} + e^{-2t}$ then F is the four-by-four matrix

$$F = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{9} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{9} & \frac{1}{16} \\ \frac{1}{4} & \frac{1}{9} & \frac{2}{8} & \frac{2}{27} \\ \frac{1}{9} & \frac{1}{16} & \frac{2}{27} & \frac{2}{64} \end{bmatrix}$$

The eigenvalues of this matrix range from about 5×10^{-5} to about .88. The smallest eigenvalue has an associated eigenvector $v^T \approx [.60, -.58, -.18, -.52]$. If we observe $\dot{y}(t) = b_1 e^{-\lambda_1 t} + b_2 e^{-\lambda_2 t} + \kappa \dot{w}$ on $[0, \infty)$ the rms error in determining $b_1 - b_2 - .5\lambda_1 - \lambda_2$ is about 244κ for the given values of the parameters. Clearly, even in such an unexceptional situation, reasonable identification of b and λ requires a very small value of κ. Marginalizing the probability distribution for the error by eliminating the dependence on b increases the size of the smallest eigenvalue, reducing the rms error associated with its eignevector to about 33κ. This is still a significant amplification of the observational noise level.

The 2-D technique: modal mixing for better identification



Two Dimensional NMR Improves the Cramer-Rao Bound

Postponing for now some details, this approach leads to a modified problem of the Prony type, now involving the fitting a function of two variables $y(t, \tau)$ with an exponential approximation. The exponents used in the expansion must be shared so the problem takes the form

$$y(t, \tau) \approx \sum_{i=1, j=1}^{n, n} b_{ij} e^{\lambda_i t} e^{\lambda_j \tau}$$

Moreover, there are constraints on the b_{ij} limiting the number of independent variables to $2n$ as above. This will become clear in the development below.

On to the Two-Dimensional Situation

Lemma 2: Let η, b and Λ be as in Lemma 1 and let T be a nonsingular matrix. Considered as a mapping from (b, λ) to $L_2[0,\infty) \times [0,\infty)$, the Jacobian of $\phi(b,\Lambda)(t,\tau) = \eta e^{\Lambda t} T e^{\Lambda \tau} b$ is

$$J(t,\tau) = \left[\begin{array}{cc} \frac{\partial \phi}{\partial b} & \frac{\partial \phi}{\partial \Lambda} \end{array} \right] =$$

$$\left[\begin{array}{cc} \eta e^{\Lambda t} T e^{\Lambda \tau}, & t b^T e^{\Lambda \tau} T^T e^{\Lambda t} + \tau \eta e^{\Lambda \tau} T e^{\Lambda t} * b^T \end{array} \right]$$

The Two-Dimensional Jacobian

The Moore-Penrose inverse of this Jacobian maps $L_2[0,\infty) \times [0,\infty)$ to $\mathbb{C}^{2n}$ and is given by

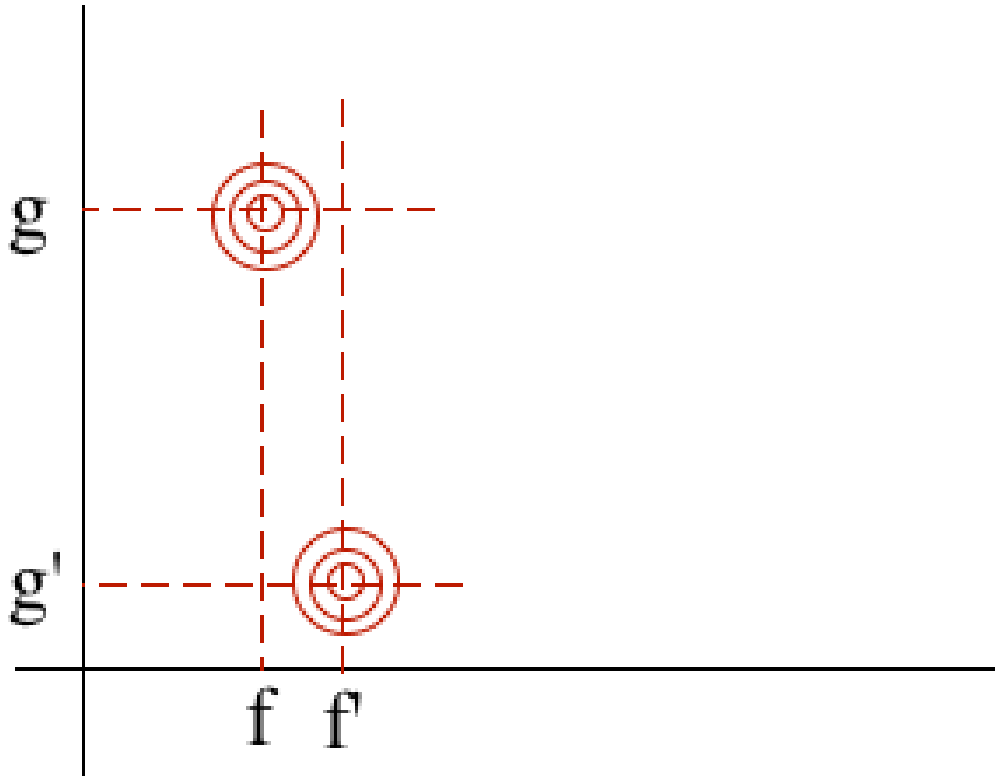
$$F^{-1} \int_0^\infty \int_0^\infty \begin{bmatrix} e^{\bar{\Lambda}\tau} T^\dagger e^{\bar{\Lambda}t} \eta^T \\ t e^{\bar{\Lambda}t} \bar{T} e^{\bar{\Lambda}\tau} \bar{b} + \tau e^{\bar{\Lambda}\tau} T^\dagger e^{\bar{\Lambda}t} \eta^T * \bar{b} \end{bmatrix} (\cdot) dt \, d\tau$$

where

$$F = \int_0^\infty \int_0^\infty J^\dagger(t,\tau) J(t,\tau) dt \, d\tau$$

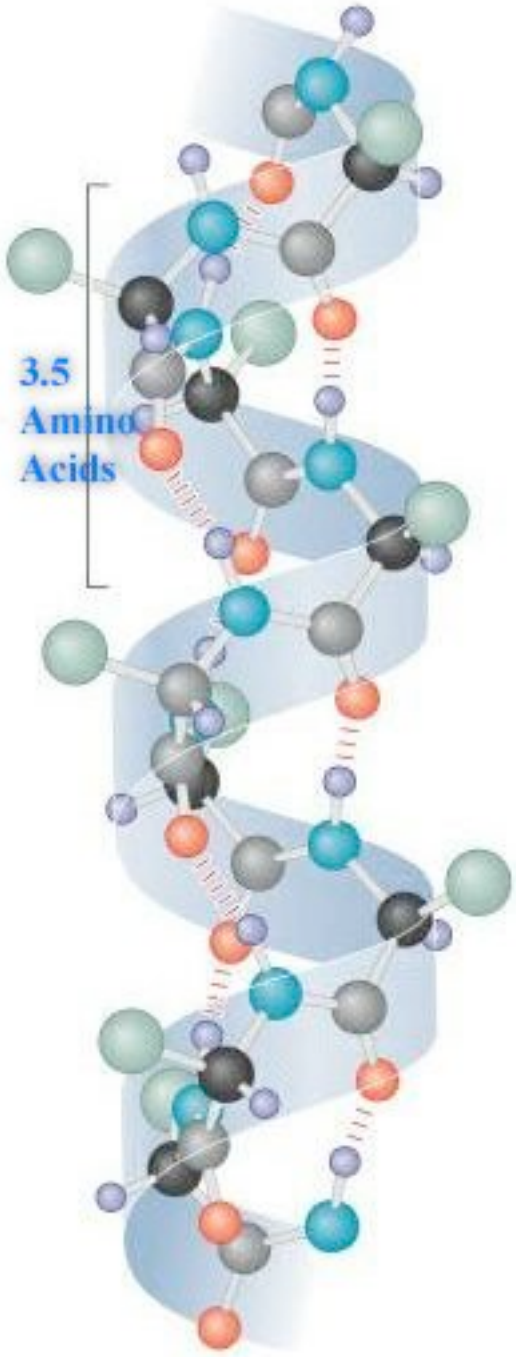
Getting to the Schur Product

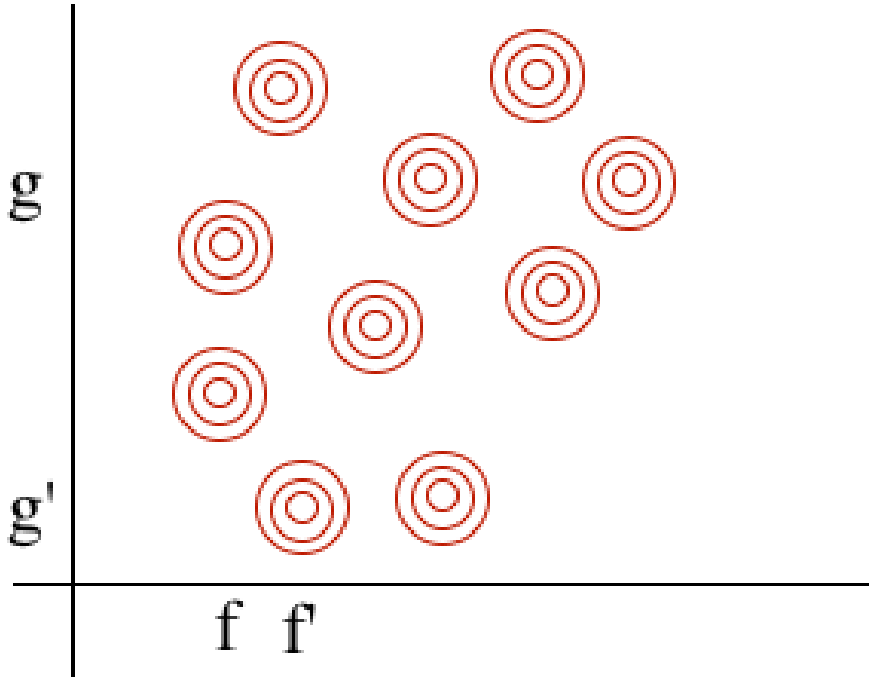
A measure of the improvement in the accurcy is reflected in the Fisher information matrix which is expressible using Schur products. As noted above, if Q and H are Hermitian and nonnegative definite their Schur product is also nonnegative definite. This is easily shown by expanding one of the factors, say H, as a sum of the form $\sum b_i b_i^\dagger$ and observing that for an arbitrary vector x, $x^\dagger(Q * b_1 b_1\dagger)x = (x * b_1)^T Q(x * b_1) \geq 0$ or by observing that $bb^\dagger * Q$ is actually congruent to to Q.



2D Fourier Space

In spectroscopic applications the spreading is caused by the shielding of the magnetic field by nearby electronic spins. The extent of the shielding is is measured by the frequency shift and this, in turn, is determined by the geometry.





2D Fourier Space

In Conclusion…

1. The density equation defines a flow on an adjoint orbit. measurement of its response to controlled inputs is an essential part of many processes of technological interest.

2. The Lindblad model is a description of the effect of unmodeled dynamics on the sub-system of interest.

3. Traditional NMR spectroscopy is limited by decoherence caused by interactions and modeled via Lindblad terms.

4. The Cramer-Rao bound provides a quantitative explanation of the effectiveness of what is called 2D NMR.

A reference on the use of Cramer-Rao bounds in this context:

R. W. Brockett, “Using Feedback to Improve System Identification” in, *Control of Uncertain Systems: Modelling, Approximation, and Design* (Edited by B.A. Francis and J.C. Willems) Springer Lecture Notes in Control and Information Sciences, 2006.