

Quantum Computing: NMR and Otherwise

- The NMR paradigm
- The quantum mechanics of spin systems.
- The measurement process
- Berry's phase in a quantum setting

Outline of the Day

- 9:30-10:15 Part 1. Examples and Mathematical Background
- 10:45 - 11:15 Coffee break
- 11:15 - 12:30 Part 2. Principal components, Neural Nets, and Automata
- 12:30 - 14:30 Lunch
- 14:30 - 15:45 Part 3. Precise and Approximate Representation of Numbers
- 15:45 - 16:15 Coffee break
- 16:15-17-30 Part 4. Quantum Computation

Importance and Timeliness of Quantum Control and Measurement

1. NMR is the main tool for determining the structure of proteins, key to the utilization of gene sequencing results, and it is now known that the existing methods are far from optimal.
2. NMR is a widely used tool for noninvasive measurement of brain structure and function but higher resolution is needed.
3. Quantum control plays an essential role in any realistic plan for the implementation of a quantum computer.
4. There are beautiful things to be learned by studying methodologies developed by physicists and chemists working in these fields, especially in the area of nonlinear signal processing.

Rough Abstract Version of the NMR Problem

Consider a stochastic (via W and n) bilinear system of the form

$$dx/dt = (A + W + u(t)B(t))x + b + n(t) \quad y=cx$$

A given waveform u gives rise to an observation process y . Given a prior probability distribution on the matrices A and B there exists a conditional density for them. Find the input waveform $u(t)$ which makes the entropy of this conditional density as small as possible.

In NMR the matrix A will have complex and lightly damped eigenvalues often in the range 10^7 /sec. Some structural properties of the system will be known and y may have more than one component. A popular idea is to pick u to generate some kind of resonance and get information on the system from the resonant frequency. Compare with optical spectroscopy in which identification is done by frequency.

An Example to Fix Ideas

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

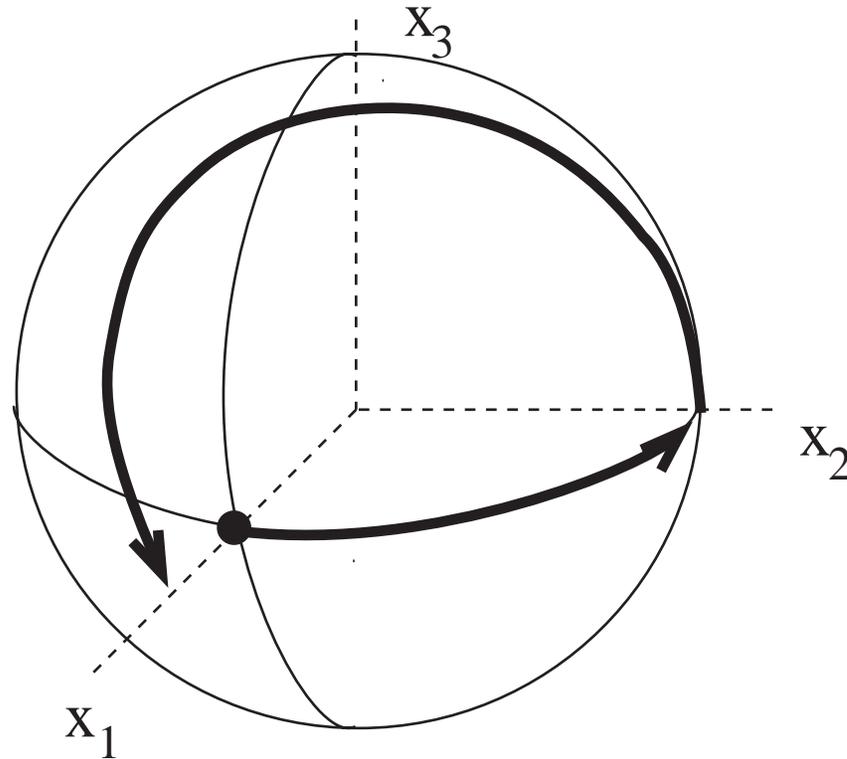
$$y = x_2 + n$$

Let w and n be white noise. The problem is to choose u to reduce the uncertainty in f , given the observation y .

Observe that there is a constant bias term. Intuitively speaking, one wants to transfer the bias present in x_1 to generate a bias for the signal x_2 which then shows up in y .

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.



The Continuous Wave Approach

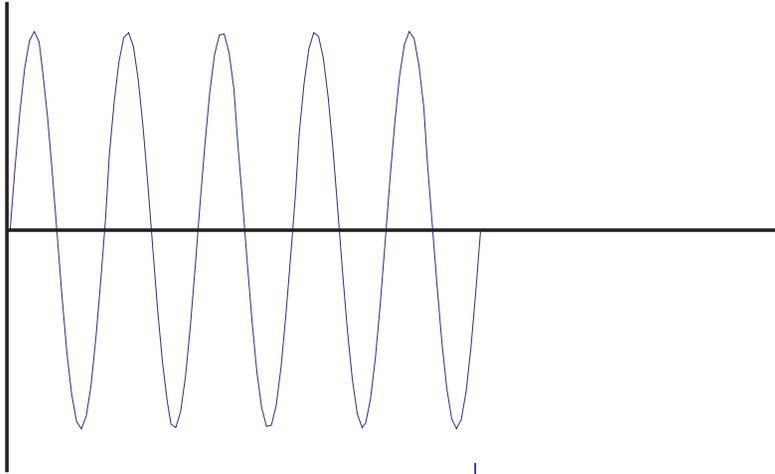
$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 & u & 0 \\ -u & -1 & f \\ 0 & -f & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$y = x_2 + n$$

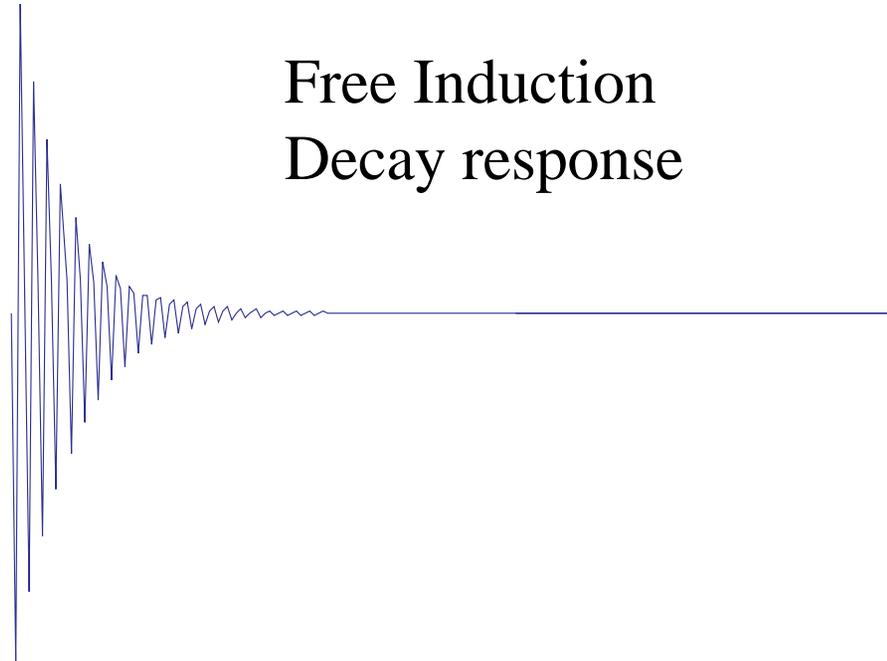
Let u be “slowly varying sine wave” $u = a \sin(b(t) t)$ with $b(t) = rt$. The benefit of the pulse goes away after the decay--the sine wave provides continuous excitation.

Possible Input-Output Response

Radio Frequency Pulse input

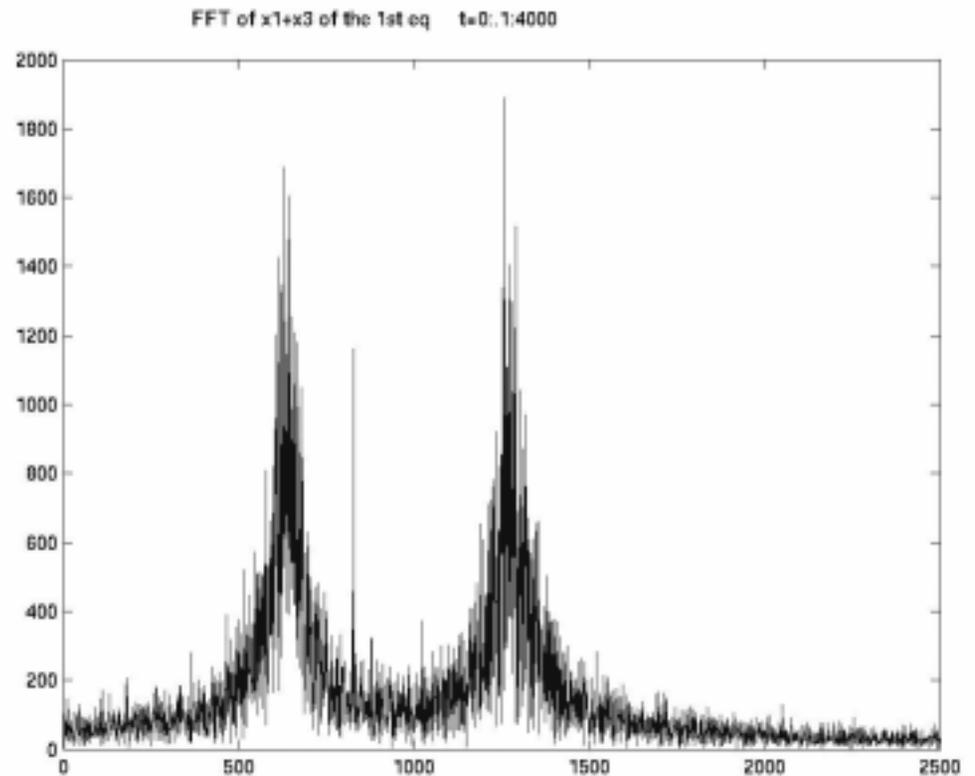
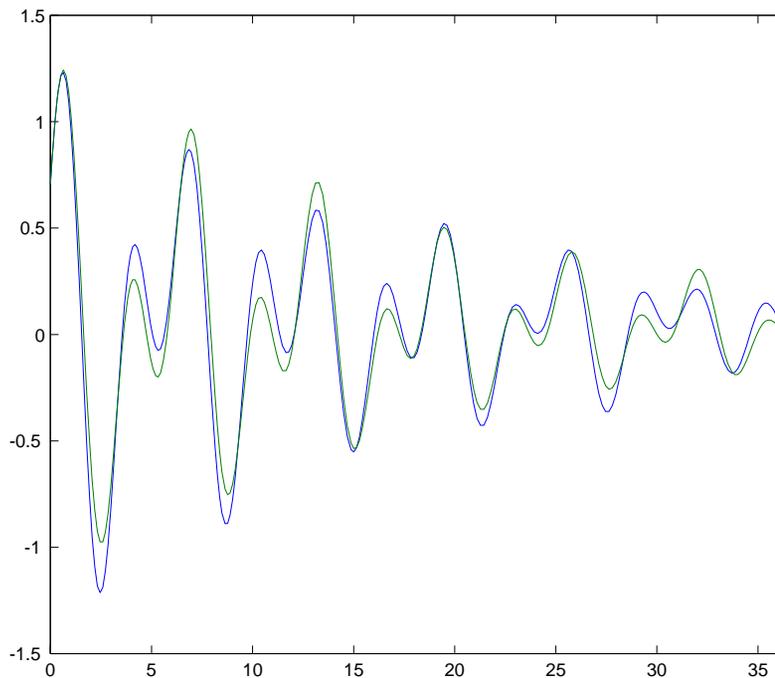


Free Induction
Decay response



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 signal processing necessary.



The Linear System Identification Problem

Given a fixed but unknown linear system

$$dx/dt = Ax + Bw \quad ; \quad y = cx + n$$

Suppose the A belongs to a finite set, compute the conditional probability of the pair (x, A) given the observations y . The solution is well known, in principle. Run a bank of Kalman-Bucy filters, one for each of the models. Each then has its own “mean” and “error variance”. There is a key weighting equation associated with each model

$$d(\ln \alpha)/dt = x^T C^T (y - Cx) - (1/2) \text{tr}(C^T C - \Sigma^{-1} B B^T \Sigma^{-1} (x x^T - \Sigma))$$

(weighting equation)

$$dx/dt = Ax - \Sigma C^T (Cx - y)$$

(conditional mean equation)

$$d\Sigma/dt = A\Sigma + \Sigma A^T + B^T B - \Sigma C^T C \Sigma$$

(conditional error variance)

The Mult-Model Identification Problem

Consider the conditional density equation for the joint state-parameter problem

$$\rho_t(t, \mathbf{x}, \mathbf{A}) = \mathbf{L}^* \rho(t, \mathbf{x}, \mathbf{A}) - (\mathbf{C}\mathbf{x})^2 / 2 \rho(t, \mathbf{x}, \mathbf{A}) + y \mathbf{C}\mathbf{x} \rho(t, \mathbf{x}, \mathbf{A})$$

This equation is unnormalized and can be considered to be vector equation with the vector having as many components as there are possible models. Assume a solution for a typical component of the form

$$\rho_i(t, \mathbf{x}) = \alpha_i(t) (2\pi^n \det \Sigma)^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{x} - \mathbf{x}_m)^T \Sigma^{-1} (\mathbf{x} - \mathbf{x}_m) \right)$$

$$d\alpha_i(t)/dt = \dots$$

$$d\mathbf{x}_i(t)/dt = \dots$$

$$d\Sigma_i(t)/dt = \dots$$

The Linear System Identification Problem Again

When the parameters depend on a control it may be possible to influence the evolution of the weights in such a way as to reduce the entropy of the conditional distribution for the system identification.

Notice that for the example we could apply a $\pi/2$ pulse to move the the bias to the lower block or we could let u be a sine wave with a slowly varying frequency and look for a resonance. It can be cast as the optimal control (say with a minimum entropy criterion) of

$$d(\ln \alpha)/dt = x^T C^T (y - Cx) - (1/2) \text{tr}(C^T C - \Sigma^{-1} B B^T \Sigma^{-1})(xx^T - \Sigma)$$

$$dx/dt = A(u)x - \Sigma C^T (Cx - y)$$

$$d\Sigma/dt = A(u)\Sigma + \Sigma A(u)^T + B^T B - \Sigma C^T C \Sigma$$

$$p_i = \alpha_i / (\sum \alpha_i)$$

Interpreting the Probability Weighting Equation

The first term changes α according to the degree of alignment between the “conditional innovations” $y-Cx$, and the conditional mean of x . It increases α if $x^T C^T (y-Cx)$ is positive. What about

$$(1/2)\text{tr}(C^T C - \Sigma^{-1} B B^T \Sigma^{-1})(x x^T - \Sigma)$$

It compares the sample mean with the error covariance. Notice that

$$C^T C - \Sigma^{-1} B B^T \Sigma^{-1} = -d\Sigma^{-1}/dt - \Sigma^{-1} A - A^T \Sigma^{-1}$$

Thus it measures a difference between the evolution of the inverse error variance with and without driving noise and observation.

Controlling an Ensemble with a Single Control

The actual problem involves many copies with the same dynamics

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

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

.....

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

$$y = (cx_1 + cx_2 + \dots + x_n) + n$$

The system is not controllable or observable. There are 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.

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, (Zeeman splitting). 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 some of the measurable aspects of spin. Protons were discovered to have spin in the late 1920's 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.

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 the total momentum is such that the spin can be only plus or minus $1/2$. 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

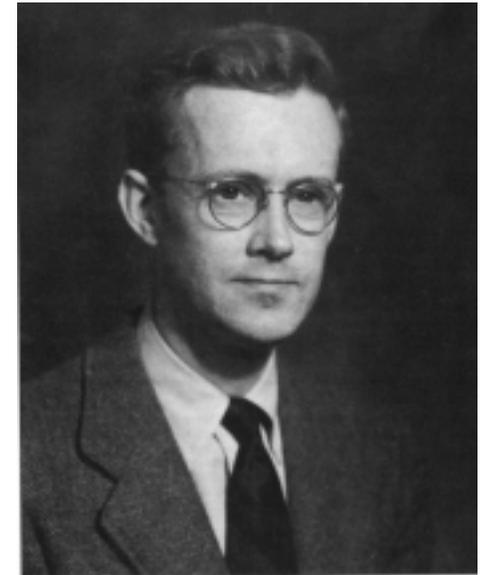


The Pioneers of NMR, Felix Bloch and Ed Purcell



$$dM/dt = B \times M + R(M - M_0)$$

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



Bloch
Nuclear Induction

$$dx_r/dt = Ax_r + b$$

$$A = \begin{bmatrix} -\frac{1}{T_2} & \omega - \omega_0 & 0 \\ -\omega + \omega_0 & -\frac{1}{T_2} & \omega_1 \\ 0 & -\omega_1 & -\frac{1}{T_1} \end{bmatrix}$$

Purcell
Absorption

ω is rf frequency,
 ω_0 is precession
frequency

In a Stationary (Laboratory) Coordinate System

$$dx/dt = Ax + b$$

$$\bar{A} = \begin{bmatrix} -\frac{1}{T_2} & -\omega_0 & \sin \omega t \\ \omega_0 & -\frac{1}{T_2} & \cos \omega t \\ -\sin \omega t & \cos \omega t & -\frac{1}{T_1} \end{bmatrix}$$

$$\bar{A} = \begin{bmatrix} -\frac{1}{T_2} & -\omega_0 & u(t) \sin \omega t \\ \omega_0 & -\frac{1}{T_2} & u(t) \cos \omega t \\ -u(t) \sin \omega t & -u(t) \cos \omega t & -\frac{1}{T_1} \end{bmatrix}$$

Why are Radio Frequency Pulses Effective

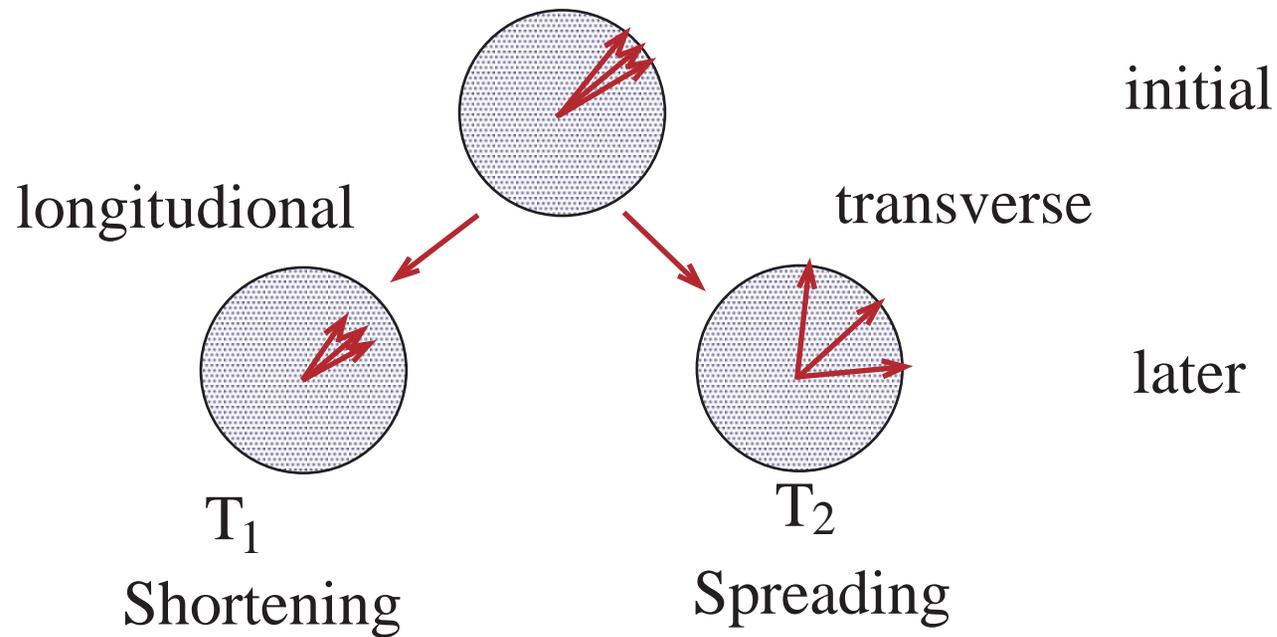
$$dx/dt = (A+u(t)B)x$$

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

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

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 $\exp(At)\exp(Ft)x(0)$.

Distinguishing Two Modes of Relaxation

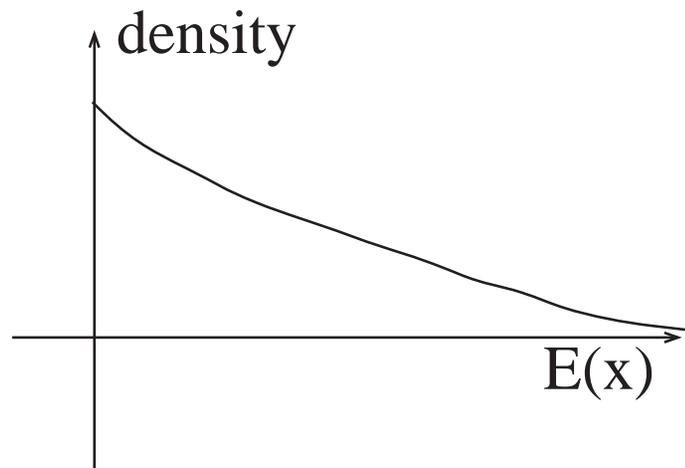


A view looking down on the transverse plane.

Boltzmann Distribution for a Physical System in Equilibrium at Temperature T

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

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.



Quantum Evolution Equations after Schrodinger

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi$$

Schrodinger Equation for a particle

$$\psi_i = \sum c_{ij} \phi_j$$

Expansion in terms of an orthonormal basis.

$$\rho = \frac{1}{N} \sum \psi_i \psi_i^T$$

The average behavior of many non-interacting particles

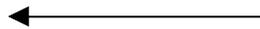
The last equation defines the so called density matrix of statistical mechanics and can be expressed in terms of the coefficients c_{ij} . These coefficients are complex and it happens that the coherence of the various quantum transitions is revealed by the off diagonal terms ρ_{ij}

The Hilbert Space for Spin

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 pure spin systems, unlike, say, the quantum description of a harmonic oscillator, the Hilbert space is finite dimensional.



John von
Neumann



Paul Dirac



The Meaning of the Density Matrix, Decoherence

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 noninteracting particles the relative phase angle matters. The size of the off-diagonals in ρ measures the consistency 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

The density matrix satisfies a linear equation derived from the wave equation. In studying NMR it is almost always simplified by eliminating many of the degrees of freedom. The resulting equation looks more complicated but it is more easily related to measurements.

The Bloch equation might be regarded as an extreme simplification of a reduced equation of this form

$$\frac{d\rho}{dt} = [iH, \rho]$$

$$\frac{d\sigma}{dt} = [iH, \sigma] + L(\sigma) + n$$

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$$

$$\sigma = \rho_{11}$$

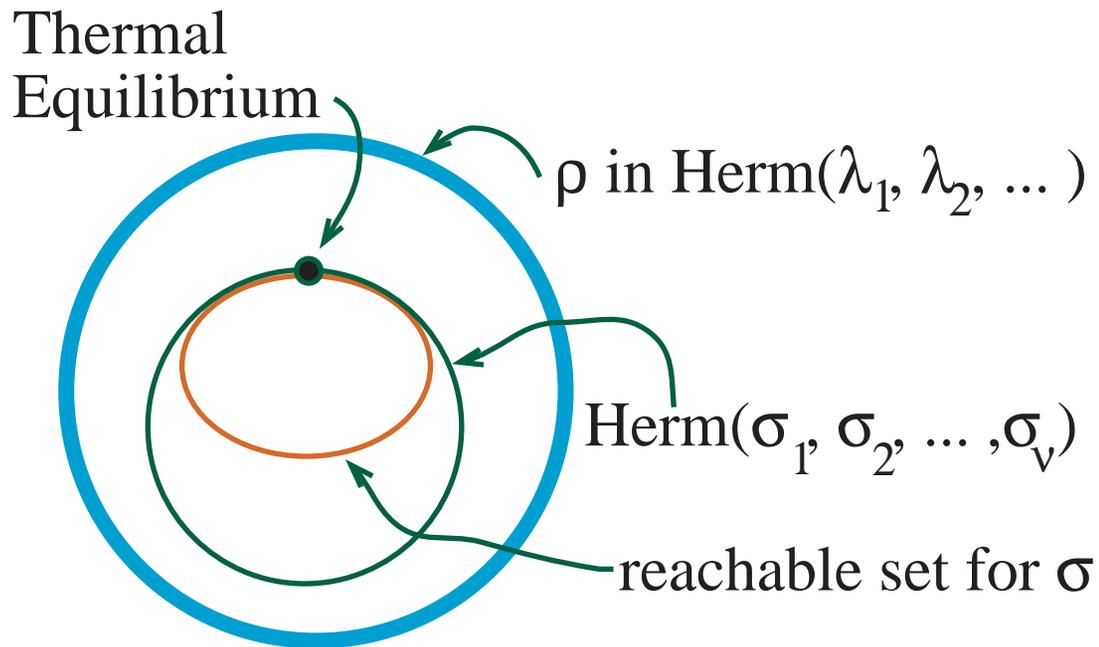
Isospectral Equation from Statistical Mechanics

The complete density equation is isospectral because it is of the form $d\rho/dt = [iH, \rho]$ form. iH simply infinitesimally conjugates the initial condition. This gives the initial condition considerable significance.

The reduced equation comes about by considering ρ to be a two by two block and focusing on the 11 term. It is then no longer isospectral. As a phenomenological equation the over-riding constraint applies to the steady state, which must be the Boltzmann distribution.

The Reduced Density Equation

For tractability, separate the “lattice dynamics” from the spin dynamics, replacing the former by an effective random term. The resulting equation 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” as above.

Back to Control Theory

Control theory can help by solving the problem of transferring the state of the reduced equation from its original value to an interesting “excited” value in minimum time. In this way the decoherence effects are minimized. For this purpose one may often ignore the dissipation and treat the reduced equation as if it were on a co-adjoint orbit. In this way the theory of controllability on Lie groups arises in the form

$$dx/dt = (A+uB)x$$

Controllability depends on the way in which A and B generate the Lie algebra. In some situations the Lie group is a rank one symmetric space and the time-optimal control can be solved for explicitly. (see recent paper by Navin Khaneja et al. In Physics Review B.)

Some Interesting Questions

1. We have framed the problem of optimal signal design in terms of minimizing the entropy of the distribution associated with conditional probabilities of the systems. Conventional practice in NMR makes extensive use of the Fourier Transform. Can we find a point of view from which the Fourier Transform defines an optimal or nearly optimal, i.e., conditional distribution generating, filter?
2. Can we find effective means for designing pulse sequences for point to point control on co-adjoint orbits of greater complexity?
3. Can we either improve on or prove the optimality of the various “two dimensional” signal processing schemes now in use in NMR?

What Kind of a Research Program Makes Sense?

1. Alternative views of computation involving an analysis of different data representations schemes and computational methods is essential if we are to get past the current status.
2. We need a better understanding of how to make use of memory in computation, and situation recognition. This includes an understanding of relational databases and their maintenance.
3. In some adaptive problems we might better think of A to Tree rather than A to D, so that we generate appropriate classification schemes.
4. Many of the issues that come up here were first articulated as computer vision problems. For example, the bottom/up -- top/down paradigm arises in that context. Computer vision is a continuing source of test cases.