9/28/09

880.05 Lecture 2

Handouts:
- a. SVM sample results
- b. MATLAB Ill-conditioning demo

Logistics:
- I sent email yesterday with a brief announcement about readings linked on the homepage to test my list. Anyone not get it?
- The reading will generally be listed a couple of days before the lecture and without an email. So keep checking the webpage. It is not essential that you go over them before the lecture, but you may find it helpful to at least skim them.

Recap:
- Last time we discussed features of nonrelativistic many-body systems at low energy and ways to make calculations tractable/easier or more systematic
  - Renormalization Group — RG. Many ways to implement. Generically modifies Hamiltonian so that most relevant degrees of freedom (dof's) are emphasized.
    - example: decoupling by SRG. of high- and low-energy parts of the potential. Matrix needed for calculations is smaller ⇒ better convergence.

$\langle k|V|k'\rangle_{initial} \longrightarrow \langle k|V|k'\rangle_{evolved}$ — RG terminology

  - Effective Field Theory — EFT. Systematically build a Hamiltonian
    - use freedom to replace short distance part of interactions because not resolved (low energy ⇒ long wavelength)
    - usually (regulated) delta function for short distance part (more later)
- Questions?? • 3-body forces are inevitable (but may be small)

9/28/09

Gameplan for this week:
- Follow pedagogical philosophy of putting on the table the important ideas of many-body calculations at the beginning using examples and analogies (heuristic overview)
    - rather than building slowly and thoroughly
    - keep returning to add details or generalize ⟹ greater **resolution**
- This week: SVM as example of "wavefunction methods" for many-body problems and model partition function as analogy for path integral partition function techniques.
- At each stage, focus on actual computing of quantities.

- Before going on, a brief aside about potentials like those from Lecture 1: (see also page (14)).

$V(r)$
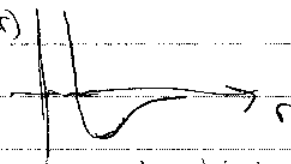
- We're used to seeing such a potential in a Schrödinger equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{x}) + V(\vec{x})\psi(\vec{x}) = E\psi(\vec{x})$$

with $V(\vec{x}) = V(|\vec{x}|) = V(r)$.

- This is a <u>local</u> potential. More generally,

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{x}) + \int d\vec{x}'\, V(\vec{x},\vec{x}')\psi(\vec{x}') = E\psi(\vec{x}')$$

so that V is <u>nonlocal</u>. If we make space discrete and label the points $i=1,2,\ldots$ then

$$\int d\vec{x}'\, V(\vec{x},\vec{x}')\psi(\vec{x}') \Rightarrow \sum_j V_{ij}\psi_j \qquad \left(V_{ij}\right)\left(\begin{smallmatrix}\psi_j\end{smallmatrix}\right)$$

and so a local potential is a diagonal matrix in coordinate space:
$V_{ij} \to V_i \delta_{ij}$ or $V(\vec{x},\vec{x}') \to V(\vec{x})\delta^3(\vec{x}-\vec{x}')$. [We'll often take $V(\vec{x}) \propto \delta(\vec{x})$ as well.]

9/28/09

So suppose we have a many-particle Hamiltonian written in "first quantization" for N particles as

$$H = \sum_{i=1}^{N}(T_i + U_i) + \frac{1}{2!}\sum_{i \neq j=1}^{N} V_{ij} + \frac{1}{3!}\sum_{i \neq j \neq k}^{N} V_{ijk} + \ldots \quad \Big| \text{ to avoid double-counting}$$

$$= \sum_{i=1}^{N} T_i + \sum_{i<j=1}^{N} V_{ij} + \sum_{i<j<k} V_{ijk} + \ldots$$

where T_i is the kinetic energy for the ith particle (e.g. $-\frac{\hbar^2}{2m}\nabla_i^2$ in coordinate rep.)
U_i is a one-body potential, V_{ij} is a two-body potential,
V_{ijk} is a three-body potential.

· These can contain gradients, spin-orbit $\vec{L}\cdot\vec{S}$ operators, etc.
· We need to convert this into a concrete form: e.g. coordinate or momentum

· The time-dependent S-eqn is

$$i\hbar\frac{\partial}{\partial t}\Psi(1,2,\ldots,N;t) = H\Psi(1,2,\ldots,N;t)$$

· Where $1, 2, \ldots$ specify a complete set of coordinates.
   E.g., $\{\vec{x}_i, s_{z_i}, t_{z_i}, \ldots\} \rightarrow$ the position, spin, isospin projections

· Consider $N=1$, numerical methods to solve for the eigenvalues: $H\Psi = E\Psi$ and eigenvectors and think about generalizing to $N = 2, 4, 20, 100, \ldots$

generalization?
a. Solve differential equations in coordinate space (integrodifferential if nonlocal V)

few-body
Faddeev
b. Solve integral equation in momentum space

mBPT
c. Write $H = H_0 + H_1$ with H_0 solvable and do perturbation theory in H_1

yes to N≈50
d. Expand Ψ in a basis and diagonalize $H_{ij} = \langle \psi_i | H | \psi_j \rangle$

yes, using d.
e. Use the variational principle and minimize: $\dfrac{\langle \Psi_{trial} | H | \Psi_{trial}\rangle}{\langle \Psi_{trial} | \Psi_{trial}\rangle}$ with respect to parameters in trial wave function $|\Psi_{trial}\rangle$.

GFMC
f. Project out the ground-state starting from a wave function $|\psi_0\rangle$ with $\langle \psi_0 | \psi_0 \rangle \neq 0$ by applying $e^{-H\tau}|\psi_0\rangle$ as $\tau \rightarrow \infty$. ($\hbar = 1$ here)

~100 cold atoms in a box
extrapolates
to thermodynamic
limit

in 1d with one particle and $H\psi_n(x) = E_n\psi_n(x)$ the exact eigenstates,

with $e^{-H\tau} = 1 - H\tau + \frac{1}{2}(H\tau)^2 + \ldots$

$$e^{-H\tau}\psi_0(x) = a_0 e^{-E_0\tau}\psi_0(x) + a_1 e^{-E_1\tau}\psi_1(x) + \ldots \xrightarrow[\tau\to\infty]{} a_0 e^{-E_0\tau}\psi_0(x)$$

9/28/09

All of these works for $N=1$ and generally up to $N=3$ or 4. Then
we start running into problems.

· The generic problem happens when we try to represent
the wave function with a discrete sample of points, for example
equally spaced in each cartesian direction (x, y, z).

· We need the spacing Δx to be small compared to the variation
scale of the wave function.

example from Negele + Orland chapter 8.

   · If we look at the nuclear case, this is about $\Delta x \approx 0.2$ fm.
   · To describe a nucleus like ^{208}Pb, which has a
   radius of about 7 fm, we would need about 100
   mesh points in each cartesian direction
   $\Rightarrow$ to represent the wave function, need $(100)^{3 \, dimensions \times 208 \, particles}$
   $= 10^{1248}$ points!

   -Suppose we just want to evaluate I (e.g. to calculate matrix elements),
   then if our computer can do gigaflops:
   10^9 operators/sec $= 10^{16}$ evaluations per year $\Rightarrow 10^{1232}$ years!
   · Petaflops won't help! 10^{1226} years.

-Same result with liquid helium with couple hundred particles
(nuclei are actually worse because of spin and isospin dependence
to keep track of.)

-You can check that even for N much smaller we have a problem.

·So what can we do?
   · One way is to sample points stochastically (random numbers!),
   which scales much more favorably with N. (e.g. for e, and f.)
   · Another is to use a more efficient basis than coordinate representation
   and do it in a truncated basis.

·The first method will be applied to path integrals but first we consider
the second approach, but also introduce a stochastic element in
selecting the basis elements. $\rightarrow$ SVM!
   ·This will only get us so far, then back to the first method next time!

9/28/09

For many details on the Stochastic Variational Method (SVM), see the book "Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems" by Suzuki and Varga (listed on 880.05 page). Other references are:

Y. Suzuki, K. Varga, J. Usukura, Nucl. Phys. A 631 (1998) 91c

and K. Varga and Y. Suzuki, Comp. Phys. Comm. 106 (1997) 157.

The CPC article documents a freely available computer code (perhaps the worst written Fortran program I've seen!).

Let's first describe the method in the simplest case: We'll look for the $L=0$ (s-wave) bound states for one particle in a central potential (eg. the Coulomb potential — solving the hydrogen atom).

What is the usual way to solve this problem with a basis?

• Pick an orthonormal basis $\psi_i(r)$, $i = 1, 2, \dots N_{max}$ with a finite # of states.

• for example, harmonic oscillator wave functions specified by a parameter ω (or sometimes Ω)

$\langle \psi_i | \psi_j \rangle = \int d^3r\, \psi_i^* \psi_j = \delta_{ij}$

- As $N_{max} \to \infty$, the basis is complete $\Rightarrow$ can represent any wave function to arbitrary accuracy [discuss bound vs. continuum]

• Expand the exact ground state in the $\psi_i(r)$'s:

$$\Psi_{GS}(r) \doteq \sum_{i=1}^{N_{max}} c_i \psi_i(r) \equiv \Psi_{trial}(r) \qquad \text{[take } c_i\text{'s real]}$$

and minimize $E_{estimate} = \dfrac{\langle \Psi_{trial} | H | \Psi_{trial} \rangle}{\langle \Psi_{trial} | \Psi_{trial} \rangle} \geq E_0$ (exact eigenvalue)

• Since Ψ_{trial} is a linear combination of the basis functions, requiring $E_{estimate}$ to be stationary under arbitrary variations of the c_i's yields the eigenvalue problem [exercise for the student] $\left(\text{ie. do } \frac{\delta}{\delta c_i} E_{estimate} = 0\right)$

$$\underline{H}\,\underline{c} = E\,\underline{c} \quad \text{or} \quad \sum_{j=1}^{N_{max}} (H_{ij} - E) c_j = 0 \quad \text{were} \quad H_{ij} = \langle \psi_i | H | \psi_j \rangle = \int d^3r\, \psi_i^*(r) H \psi_j(r)$$
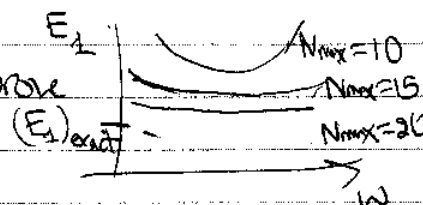$$i=1,\dots,N_{max}$$

$\Rightarrow$ get N_{max} eigenvalues (and eigenvectors) from H_{ij} [eg. eig(H) in MATLAB]
(Why do we call this "diagonalizing H"?)

9/26/09

- For a given N_{max}, the best E-estimate is the lowest eigenvalue as w is varied.
- As N_{max} is increased, all estimates improve

E_1

$(E_1)_{exact}$

$N_{max} = 10$
$N_{max} = 15$
$N_{max} = 20$

w

-Now lets do the SVM. Here the basis elements are always just Gaussians, i.e. of the form

$$\psi_i(r) = N_i e^{-\alpha_i r^2/2}$$

with α_i specifying the width of the gaussian, and N_i is a normalization constant.
- This is like the ground-state harmonic oscillator wave function. (If we want an $L \neq 0$ state, we just multiply by $Y_{\ell m}(\Omega)$ with $m=0$.)
- But instead of using excited states, we'll have only Gaussians, with the α's chosen randomly, keeping the best ones.
- If we choose a basis with K elements (so $K \sim N_{max}$), then

$$\Psi_{trial}(r) = \sum_{i=1}^{K} c_i \psi_i(r)$$

and the best estimate for the ground state is for the c_i determined by the generalized eigenvalue problem

$$\underline{\underline{H}}\,\underline{C} = E\,\underline{\underline{B}}\,\underline{C} \quad \text{or} \quad \sum_{j=1}^{K} (H_{ij} - E B_{ij}) c_j = 0 \quad (i=1,\dots,K)$$

where $\underline{\underline{H}}$ and $\underline{\underline{B}}$ are the $K \times K$ matrices $H_{ij} = \langle \psi_i | H | \psi_j \rangle$, $B_{ij} = \langle \psi_i | \psi_j \rangle$

- These are not orthogonal functions, so $B_{ij} \neq \delta_{ij}$ (as with the orthonormal harmonic oscillator basis).
- This sounds like a big disadvantage — why would we want to use such functions?
   - Gaussians are versatile basis functions — choosing different α's allows the trial function to imitate different scales in the exact wave function — like the width of the main part vs. the fall-off at large r (the "tail"). HO wave functions have great trouble with this.
   (What about continuum wf's?)

see figures and comments below

- With enough Gaussians, the basis becomes functionally complete (that is, we can reproduce [almost] any wave function)
- Gaussians are great for calculating integrals ⇒ we can arrange to do the integrals *analytically*! (more below)

✗ - Most importantly for our purposes, then are generalizations to N>1 that work very well.

⇒ In particular, we generalize $e^{-\alpha r^2/2} \rightarrow e^{-\frac{1}{2}\sum_{j>i=1}^{N} \alpha_{ij} (\vec{r}_i - \vec{r}_j)^2} = e^{-\frac{1}{2}x^T A x}$

- need to include spin, isospin, L≠0 angular part, $\frac{N(N-1)}{2}$ parameters [see below] and then (anti) symmetrize the whole thing. [come back to this] [for these]

Let's first describe the basic procedure and then make comments on its special features. Note: everything will go through the same for N>1 with $\alpha \rightarrow \{\alpha_{ij}\}$

- We start by picking N values of α_1, the first basis element. Label them α_1^n, (n=1,...N). In practice, N might be something like 5 or 10. Let's suppose it is 10.
- We select them randomly in a "physical" interval: $\alpha_{min} \leq \alpha_1^n \leq \alpha_{max}$ where α_{min} and α_{max} characterize the natural scales (e.g., the size of the nucleus or atom will set one end). We might want a non uniform (e.g. logarithmic) distribution.
- Calculate the energy estimate for each of the N values of α_1 (so N estimates; see below for how to do this) and pick the best (lowest) result to fix α_1.
- Now select another N values for α_2, that is α_2^n (n=1,...N). With α_1 fixed, compare the new energy estimates obtained from the N 2×2 generalized eigenvalue problems.
    - Theorem 3.5 of the Chp. 3 excerpt from Suzuki and Varga says that all of the estimates will be better (lower) or at least not worse than the estimate from the first round.
  - Once again, keep the best one. ⇒ fix α_2.
- Keep repeating this process until we have K basis states with α_i, (i=1,...K), and our best estimate for the energy.

9/28/09

· Before going any further, we should see if it works!
You'll have a chance to play with a code that implements
this SVM, but for now let's look at graphs and tables
for a test case from the Suzuki/Varga book which a simple
"self-bound" three-body system: the positronium negative ion Ps^-
(two electrons and a positron) with the usual Coulomb potential
binding the system. [A "few-" rather than "many"-body system, but good for exploring.]
   · We'll come back below to discuss the "correlated gaussian"
   wave function — all we need to know now is that instead
   of one α chosen randomly to specify the wf, there are 3 $[\alpha_{12}, \alpha_{23}, \alpha_{13}]$
   Otherwise, the procedure goes through as before.

· Consider Fig 4.2 and Tables 4.4 and 4.5.
   · The graph in Fig 4.2 shows the convergence of the energy
   estimates for 5 different runs of the procedure toward
   the "exact" result of -0.262005 ("exact" here means the
   result to that many digits if only the pure Coulomb interaction
   is considered and not other physical contributions).
   · Note the monotonic decrease with the basis size K (the x-axis)
   [note also the expanded scale!]
   · The rate of decrease in the energy estimate [in "atomic units" or a.u.]
   gives an idea of how close to convergence we are.
   · The sets of parameters are different — which is quite a change
   from the usual orthonormal basis — clearly we can substitute some
   parameter sets for others and get as good a result for 80 or 100 basis states.
   · Table 4.4 shows results for basis size K=10 for N=1
   and N=10 [3rd and 4th columns]. The "Powell" column is a   } very difficult
   deterministic (rather than stochastic) method to select the α_i's.   to do non-
   Table 4.5 shows similar results for K=100.   linear
                                                                                optimization
                                                                                ⇒ global minimum
                                                                                hard to find
   · The last column shows the improvement with "refining" cycles. This procedure corrects
   for the fact that the early choices i=1,2,3 for the α_i's may no longer be
   optimal once K α_i's are chosen. So, one at a time, simply go back with all other
   α's frozen and check N more of each state to see if there is improvement. It works!

9/28/09

• Even for this very small system, the gain in time or Neval over the deterministic method is impressive.
• With refining, the final result is very good. (Note: refining K=100 states is much better than increasing K.)

We haven't mentioned a key feature that makes the entire method feasible. If we had to actually solve the generalized diagonalization problem every time we consider another K, the cost would be prohibitive (matrix manipulations for matrices of size K have times that increase like K^3 or worse in general). The point is that we can find how the lowest eigenvalue changes when adding a single new basis function <u>without</u> doing a full diagonalization.

    • This is worked out in the handout on variational methods, where it is shown that given the K eigenvalues ϵ_i, $i=1,...K$, we can calculate matrix element h_i, $i=1,...,K+1$ using the new ψ_{K+1} basis element, then the new eigenvalues are the roots of

$$E - h_{K+1} = \sum_{i=1}^{K} \frac{|h_i|^2}{E - \epsilon_i}$$

as shown in Fig. 3.1. The new eigenvalues are the intersections of the straight line and the other curves. It demonstrates the theorem about improving eigenvalues and shows that <u>all</u> of the eigenvalue estimates improve.

Excited states. The theorems in the Chapter 3 excerpt
as applied to excited states are not widely known.
They basically say that not only is the lowest eigenvalue an
upper bound to the lowest bound state energy, the 2nd lowest
eigenvalue is an upper bound to the 2nd lowest bound state energy,
and so on.

- So we automatically get variational estimates of the
excited states.

- However, choosing the α_i's to lower the ground state energy doesn't
mean in general that it is the best choice for an excited state.

- But we can apply procedures as shown in Fig 4.4 and Table 4.7
for the first five ^{1}S states of the Helium atom.

- Procedure B optimizes each state separely (so 5 different
bases).

- Procedure A yields one common basis by choosing α_i to
lower the ground state for the first k=100 states, then choosing
α_i to lower the 2nd state for the next 100 (up to k=200),
and so on. Fig. 4.4 shows it works very well.

Improving the Hamiltonian with RG.

- Figure 6.1 to 6.6 show how well Gaussian expansions work to
describe even very non-gaussian looking functions.

- But the Gaussians do have trouble with the function $f(r) = r^{12}e^{-r^2}$,
which arises with Lenard-Jones or other repulsive-core
potentials.

- The solution is not to try to fix the expansion but to change the
Hamiltonian! This is what the Similarity Renormalization Group (SRG)
does (see Handout from Lecture 1). The potential is modified such that
the hard core is removed and then variational calculations work much
better.

- This is a consequence of decoupling: the short wavelength parts are no longer
needed to describe the low-energy states.

9/28/09

What can go wrong?

· There is an important numerical issue associated with the generalized eigenvalue problem that is a generic matrix problem we will encounter again when calculating path integrals numerically. This is the problem of "ill-conditioning"

· Ill-conditioned matrices occur here when our non-orthogonal basis becomes close to being linearly dependent. This means that one of the basis elements is almost a linear combination of the others.

· As shown in the Chapter 3, exercise, linear independence is equivalent to having only the solution $\underline{C} = 0$ to $B\underline{C} = 0$, where $B_{ij} = \langle \phi_i | \phi_j \rangle$ is the overlap matrix. This in turn implies that $\det B \neq 0$, which means the eigenvalues are all greater than zero.

· Near linear dependence then is signalled by one or more very small eigenvalues. This can lead to a catastrophic sensitivity to matrix elements (of H or B).   ⎤ see the MATLAB handout for an example.

· A practical fix for the SVM is to reject states unless their overlap of basis states (normalized) is smaller than a specified value (which probably is determined empirically).

· You'll do a simple example illustrating ill-conditioning in the homework.

Come back to the generalization

$$e^{-ar^2} \longrightarrow e^{-\frac{1}{2}\sum_{i,j=1}^{N}\alpha_{ij}(\vec{r}_i-\vec{r}_j)^2} = e^{-\frac{1}{2}\underline{x}^T \underline{\underline{A}}\,\underline{x}} \equiv G_{\underline{\underline{A}}}(\underline{x})$$

· If we go to the $N=2$ case with $\vec{r}_1$ and $\vec{r}_2$ the particle coordinates and $V \to V(\vec{r}_1-\vec{r}_2)$, then we usually switch variables to

$$\vec{x} = \vec{r}_1 - \vec{r}_2 \quad \text{and} \quad \vec{X} = \tfrac{1}{2}(\vec{r}_1+\vec{r}_2)$$

and then our wave function is taken to be a function of only the relative coordinate $\vec{x}$. So this is just the case already considered.
· What about the dependence on $\vec{X}$? What is it and where does it go?

answer: It is the center-of mass coordinate (com) and V does not depend on it by Galilean invariance. For self-bound systems, the com only shows up as a non-dynamical contribution to the kinetic energy, which we can subtract off or just set to zero by working in the com frame.
· What about antisymmetry is these are fermions? Need a spin antisymmetric
· define $\vec{x}_1 \equiv \vec{x}$ and $\vec{x}_2 \equiv \vec{X}$ for $N=2$.
$\quad$ wf $\Rightarrow \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\uparrow\downarrow\rangle)$
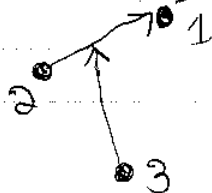$\quad$ if only one type. (Other choices if we have protons and neutrons.)

· What about $N=3$? $\vec{r}_1, \vec{r}_2, \vec{r}_3$
Again, we want to eliminate the com
$$\Rightarrow \vec{x}_3 = \tfrac{1}{3}(\vec{r}_1+\vec{r}_2+\vec{r}_3)$$
The others are "Jacobi coordinates", which (for equal mass particles we define as

$$\vec{x}_1 = \vec{r}_1 - \vec{r}_2 \qquad \vec{x}_2 = (\vec{r}_1+\vec{r}_2)/2 - \vec{r}_3$$

9/28/09

For general N, $\boxed{\underline{x} = \underline{\underline{U}}\, \underline{r}}$ ($\underline{x}, \underline{r}$ are vectors of vectors!)

and $\underline{r} = \underline{\underline{U}}^{-1} \cdot \underline{x}$ (these are not unitary matrices)

For $N=2$,

$$\underline{\underline{U}} = \begin{pmatrix} 1 & -1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \Rightarrow \underline{\underline{U}}\begin{pmatrix} \vec{r}_1 \\ \vec{r}_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} \vec{r}_1 \\ \vec{r}_2 \end{pmatrix} = \begin{pmatrix} \vec{r}_1 - \vec{r}_2 \\ \frac{1}{2}\vec{r}_1 + \vec{r}_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

For $N=3$,

$$\underline{\underline{U}} = \begin{pmatrix} 1 & -1 & 0 \\ \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \Rightarrow \underline{\underline{U}}\begin{pmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 \end{pmatrix} = \begin{pmatrix} \vec{r}_1 - \vec{r}_2 \\ \frac{1}{2}(\vec{r}_1+\vec{r}_2)-\vec{r}_3 \\ \frac{1}{3}(\vec{r}_1+\vec{r}_2+\vec{r}_3) \end{pmatrix}$$

For general N,

$$\underline{\underline{U}} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & -1 & & 0 \\ & & \ddots & & \\ \frac{1}{N-1} & \frac{1}{N-1} & & & -1 \\ \frac{1}{2} & \frac{1}{2} & \cdots & & \frac{1}{N} \end{pmatrix} \qquad \underline{\underline{U}}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{N} & 1 \\ -\frac{1}{2} & \frac{1}{3} & \cdots & -\frac{1}{N} & 1 \\ 0 & \frac{2}{3} & & & \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & -\frac{N-1}{N} & 1 \end{pmatrix}$$

So we can switch from $\vec{r}_i - \vec{r}_j$ to x_K by

$$\boxed{\vec{r}_i - \vec{r}_j = \sum_{k=1}^{N}\left((U^{-1})_{ik} - (U^{-1})_{jk}\right)x_k}$$

which means that $\{x_{ij}\}$ and $\{A_{ij}\}$ can be related by linear transformation.

What about antisymmetrization?

- Here again we have simple linear transformations that permute (switch) indices:

$$\text{antisymmetrizer } \hat{a} = \sum_{i=1}^{N!} (\text{sgn } P_i) \times P_i$$

where the P_i exchange the indices in all possible ways (i is just a label) and the sign is according to whether it is an even or odd permutation.

- For example, if $N=3$ the permutation $(3\ 1\ 2)$ is carried out by the matrix

$$C = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{acting on} \begin{pmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 \end{pmatrix}, \Rightarrow \begin{pmatrix} \vec{r}_3 \\ \vec{r}_1 \\ \vec{r}_2 \end{pmatrix}$$

- The action of P_i on $G_A(\vec{x})$ is simply to transform $\underline{\underline{A}} \rightarrow P_i^T \underline{\underline{A}} P_i$ where the P_i matrices can be found in the SVM book.
  $\Rightarrow$ all of this is easy to program.

- We've neglected the spin and isospin wavefunctions for the most part $\Rightarrow$ a complete set will need to be included.

[See the Blume/Daily arXiv: 0909.2701 for an example of specifying the spin and antisymmetrization.]

9/28/09

The great thing about Gaussians is that we can do all sorts of integrals. We'll be doing this quite a bit, so lets just recall some simple ones.

First, $\boxed{I \equiv \int_{-\infty}^{\infty} dz \; e^{-\alpha z^2/2} = \left(\frac{2\pi}{\alpha}\right)^{1/2}}$

[Usual proof in polar coordinates:

$$I^2 = \iint_{-\infty}^{\infty} dx\, dy \; e^{-\alpha(x^2+y^2)/2} = \int_0^{2\pi} d\theta \int_0^{\infty} r\, dr \; e^{-\alpha r^2/2} = \pi \int_0^{\infty} du \; e^{-\alpha u/2} = \frac{2\pi}{\alpha}$$
$$\underset{\substack{u=r^2 \\ du = 2r\,dr}}{}$$
]

• What about a matrix? From Zinn-Justin handout (and many others),

$$I(\underline{\underline{A}}, \underline{b}) = \int_{-\infty}^{\infty} \prod_{i=1}^{n} dx_i \; \exp\left(-\sum_{i,j=1}^{n} \tfrac{1}{2} x_i A_{ij} x_j + \sum_{i=1}^{n} b_i x_i\right)$$

$\underset{\substack{\underline{\underline{A}}^T = \underline{\underline{A}} \\ \text{or} \\ A_{ij} = A_{ji} \\ (\text{and real})}}{}$

$$= \int_{-\infty}^{\infty} dx_1\, dx_2 \cdots dx_n \; e^{-\frac{1}{2} \underline{x}^T \underline{\underline{A}} \underline{x} + \underline{b}\cdot\underline{x}}$$

$$= (2\pi)^{n/2} (\det A)^{-1/2} \; e^{\frac{1}{2} \underline{b}^T (\underline{\underline{A}}^{-1}) \underline{b}}$$

How do we prove something like this?
 • For many matrix manipulations, the path to a proof or evaluation is i) find the answer for a diagonal matrix and ii) note that we can diagonalize (many) matrices.
 • Here, consider the $\underline{b}=0$ case.
   • Diagonalize $\underline{\underline{A}}$ with $\underline{\underline{A}}_D = \underline{\underline{M}} \underline{\underline{A}} \underline{\underline{M}}^{-1}$ with $\underline{\underline{M}}$

$$\Rightarrow I(\underline{\underline{A}}, 0) = \int_{-\infty}^{\infty} dx_1 \cdots dx_n \; e^{-\frac{1}{2} \underline{x}^T \underline{\underline{M}}^{-1}(\underline{\underline{M}}\underline{\underline{A}}\underline{\underline{M}}^{-1})\underline{\underline{M}} \underline{x}}$$

$$\underset{\substack{\underline{x}' = \underline{\underline{M}}\underline{x}}}{=} \int_{-\infty}^{\infty} dx_1' \cdots dx_n' \; e^{-\frac{1}{2} \underline{x}'^T \underline{\underline{A}}_D \underline{x}'} = \prod_{i=1}^{n} \int_{-\infty}^{\infty} dx_i' \; e^{-\frac{1}{2} x_i' \epsilon_i x_i'} = \prod_{i=1}^{n} \left(\frac{2\pi}{\epsilon_i}\right)^{1/2} = \frac{2\pi^{n/2}}{(\det A)^{1/2}}$$
$$\underset{\text{Jacobian?}}{} \qquad\qquad \underset{\text{eigenvalue}}{\uparrow\epsilon_i}$$

9/28/09

In doing that proof, we used that:
- a symmetric matrix can be diagonalized by a rotation of the coordinates (orthogonal transformation)
- the Jacobian for that transformation is one.
- The diagonal elements of a diagonal matrix are the eigenvalues (what are the eigenvectors?)
- $\det(\underline{M}\,\underline{A}\,\underline{M}^{-1}) = (\det \underline{M})(\det \underline{A})(\det \underline{M}^{-1}) = \det \underline{A}$

  since $\det \underline{M}^{-1} = (\det \underline{M})^{-1}$
- The determinant of a diagonal matrix is the product of its eigenvalues.
- What is the trace of a matrix $\underline{A}$? (Both by definition and in terms of eigenvalues.)

We'll see more such manipulations as we go!

For the correlated Gaussians, the claim is that in D spatial dimensions

$$\langle G_{A'} | G_A \rangle = \int dx_1 \cdots dx_{N-1} \, e^{-\frac{1}{2} \underline{x}^T \underline{A}' \underline{x}} \, e^{-\frac{1}{2} \underline{x}^T \underline{A} \underline{x}}$$

$$= \left(\frac{(2\pi)^{N-1}}{\det(\underline{A} + \underline{A}')} \right)^{D/2}$$

Is this correct? Is there an issue about $\underline{A}$ and $\underline{A}'$ commuting?

Similarly (but with longer proofs), there are formulas for evaluating the matrix element of the kinetic energy and the potential.
  → can do them analytically, which is a big help for the SVM.

Lead in to lecture 3 and beyond...

- The SVM and other diagonalization approaches (FCI) eventually run into a wall that limits the maximum number of particles possible.
    - This is basically because of the factorial growth in the number of possible many-particle basis states, which means the matrix to be diagonalized grows factorially.

- So it is useful but with limits.

- From now on, we take a different path, which is based on a thermodynamics/statistical mechanics point of view.
    - We had a hint of this direction earlier when we considered hitting a starting (guess) wave function $|\psi_0\rangle$ with $e^{-\hat{H}\tau}$ $(\tau=1)$

recall from

$$\boxed{\text{M}} \Rightarrow \quad e^{-\hat{H}\tau}\psi_0(x) \xrightarrow[\tau\to\infty]{} a_0 e^{-E_0\tau}\psi_0(x) + (\text{small})$$

(ground-state energy) (ground-state wf)

- This might remind you of:
    i) Boltzmann factors $e^{-\beta\hat{H}} \Rightarrow \tau \Leftrightarrow \frac{1}{\beta}$
    ii) the formal solution to the time evolution of a wave function

$$\boxed{i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \hat{H}\Psi(x,t) \Rightarrow \Psi(x,t) = e^{-i\hat{H}t/\hbar}\Psi(x,0)}$$

so $it \to \tau \Rightarrow$ evolution in imaginary time.

- Make the connection through the partition function $Z = e^{-\beta\hat{H}}$
    $\Rightarrow$ review some stat. mech. and thermo next time and introduce a model for our path integral expressions for Z with external fields added.