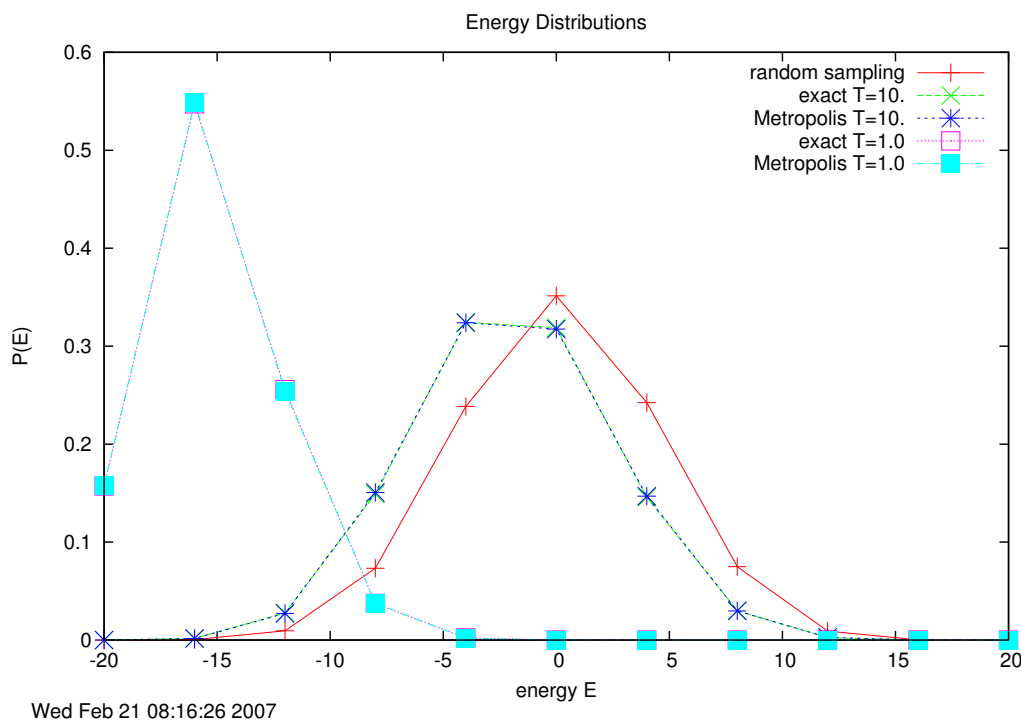


## 13. 780.20 Session 13

### a. Follow-ups to Session 12

- **Comment on the codes from Session 12.** The codes used in Session 12 use various “tricks” or shortcuts to make them more compact. In general, I don’t recommend coding this way (so these codes should be rewritten!). For example, the choice of energy units is exploited so that energies are integers and therefore can be used as array indices. It would be much better to define a function that returned an appropriate integer index when called with the energy. Another example is the `next_configuration` function, which uses recursion (i.e., the function calls itself) to count upward in binary. This is clever, but really hard to decipher!
- **Recap of energy distributions figure.**

Here is a typical figure you might have obtained from the `sampling_test` program:



Some observations:

- The “random sampling” selects configurations (recall that a configuration is one particular specification of spins in the chain) at random, so each one is equally likely to be chosen. Then the energy is calculated and histogrammed. Since there are only two out of about one million configurations that have the lowest energy of  $-20$ , the curve will be almost zero there. The peak in the middle means that there are more ways to have zero energy than any other energy. The curve is symmetric because there are equal numbers of configurations with a given positive energy as with the corresponding negative energy.

- The exact  $P(E)$  at temperature  $T$  combines two factors: the Boltzmann factor  $e^{-E/kT}/Z$  and the number of states with energy  $E$ . For  $T = 1$  (we use units where  $k = 1$ ), the first factor is largest for  $E = -20$  but the second is smallest there. The product is a moderate probability. For  $E = -16$ , the first factor is smaller but the second is sufficiently larger that the net result is the maximum probability. Even though the second factor is largest at  $E = 0$ , the Boltzmann factor is very small (remember that  $Z$  is large), so there is only a very small probability  $P(0)$ .
  - For larger temperatures, the Boltzmann factor is more uniform, and so the factor of the number of states with energy  $E$  dominates. Thus, the random sampling distribution corresponds to very high temperature.
  - The Metropolis distribution uses only a representative subset of the full set of configurations to estimate  $P(E)$ . But it does a very good job of reproducing the exact distribution at a given temperature.
- **Thermal averages.** Recall the basic idea for evaluating thermal averages. We generate a set of “configurations,” which we’ll label  $\mathbf{x}_0, \mathbf{x}_1, \dots$  or collectively as  $\{\mathbf{x}_i\}$ , that are distributed according to an appropriate probability distribution function (PDF). One of the most physical PDF’s is the Boltzmann distribution. In this case, the thermal average of  $A(\mathbf{x})$  is approximated by

$$\langle A \rangle_T \equiv \frac{\int d\mathbf{x} A(\mathbf{x}) e^{-\mathcal{H}(\mathbf{x})/kT}}{\int d\mathbf{x} e^{-\mathcal{H}(\mathbf{x})/kT}} \approx \sum_i A(\mathbf{x}_i). \quad (13.1)$$

Thus, we find the average simply by summing the quantity  $A$  (e.g., the magnetization) over our representative set of  $\mathbf{x}_i$ ’s (e.g., spin configurations). At high temperatures, there will be a large spread of configurations contributing while at low temperature, states close to the ground state will increasingly dominate (i.e., appear more times in the set of  $\mathbf{x}_i$ ’s).

- **Ensuring representative configurations.** When carrying out the sum over configurations, we want to use ones that are actually representative. Three things we need to do:
  1. Throw out configurations from the beginning until the system has “equilibrated.” How might we tell when this has happened?
  2. Don’t use each configuration resulting from a MCS (“Monte Carlo Step”) but skip every  $n$  configurations to avoid correlations. Basically,  $n$  should be enough MCS’s so that the next configuration used has lost its memory of the previous one used. You can experimentally determine an appropriate “ $n$ ” by looking at an autocorrelation function. We’ll see how this works in Session 13.
  3. Adjust the “step size” between successive  $\mathbf{x}$ ’s so that the Metropolis success rate is reasonable. A rate around 50% is the canonical trade-off between computational efficiency and exploring enough of the configuration space.
- **Efficiencies.** When looking at `ising_model.cpp` compared to `ising_opt.cpp`, please note what has been done to speed up the code. All of the upgrades described in the Session 12 guide are helpful. Why does having a look-up table speed things up? Answer: if there are only a fixed number of energies, we don’t need to calculate the Boltzmann ratio  $e^{-\Delta E/kT}$  over

and over again. Exponentials are expensive (i.e., they take a relatively long time to calculate) while looking up values in an array is cheap. The general idea of a look-up table is one that is frequently useful in speeding up codes.

- **Phase transition.** The optimized Ising model code with a ferromagnetic interaction has a phase transition in two dimensions. There is a “Curie temperature” below which the magnetization is nonzero. Can you see evidence of this phase transition from running `ising_opt`?
- **Further info.** Take a look at the chapters from Hjorth-Jensen, *Computational Physics*, which are posted on the 780.20 webpage, for a good summary of Monte Carlo methods and applications. The excerpt from the Binder/Heermann book (and the rest of the book) has good suggestions for the practical implementation of Monte Carlo approaches.

## b. Exchange Integral $J$

[This summary is based on the discussion in Shang-Keng Ma’s book on *Statistical Mechanics*.]

Let’s look at the physical origin of the exchange integral  $J > 0$  for a ferromagnet. The short-distance exchange interaction is the strongest among electron spins, and comes from the interplay of electrostatic repulsion and the Pauli exclusion principle. Consider the interaction between electrons at two sites. If it were just a magnetic dipole interaction, then  $\uparrow\uparrow$  would not be energetically favored and we would not have ferromagnets! To find the true origin of the interaction, we must turn to quantum mechanics.

Consider the wave function of electrons 1 and 2 as a product of a spatial part  $\varphi$  and a spin part  $\chi$  (this factorization is possible if there is negligible spin-orbit coupling, which we assume):

$$\psi(1, 2) = \varphi(\mathbf{r}_1, \mathbf{r}_2) \chi(s_1, s_2) , \quad (13.2)$$

where the  $\mathbf{r}_i$ ’s are the positions of the electrons and the  $s_i$ ’s are the spins. Since electrons are identical *fermions*, the overall wavefunction must be *antisymmetric* under particle exchange (i.e.,  $\psi(2, 1) = -\psi(1, 2)$ ). That leaves two possibilities:  $\varphi$  is symmetric while  $\chi$  is antisymmetric, or *vice versa*. For example,

$\chi$	implication for $\varphi$
$\uparrow\uparrow$	$\varphi(\mathbf{r}_1, \mathbf{r}_2) = -\varphi(\mathbf{r}_2, \mathbf{r}_1)$
$\frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow)$	$\varphi(\mathbf{r}_1, \mathbf{r}_2) = +\varphi(\mathbf{r}_2, \mathbf{r}_1)$

Thus the symmetry of the spins dictates the symmetry of the spatial part; this is the key!

Consider first the non-interacting case and imagine the spatial wavefunctions are like hydrogen ground state wavefunctions centered at the origin  $O$  and point  $R$ . So we make the identification

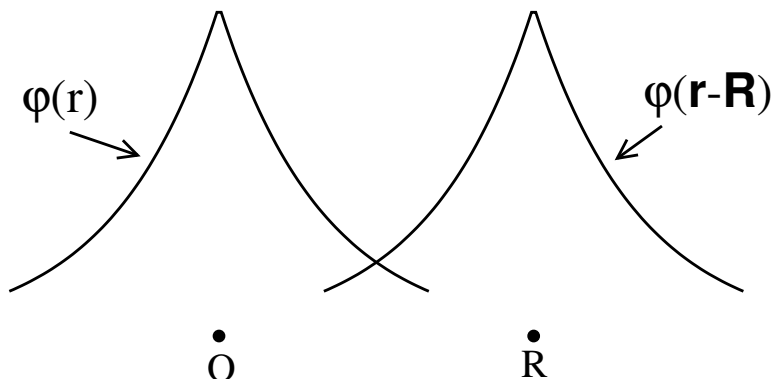
$$\varphi_1(\mathbf{r}) = \varphi(r) , \quad \varphi_2(\mathbf{r}) = \varphi(\mathbf{r} - \mathbf{R}) , \quad (13.3)$$

as pictured in the figure. Then the noninteracting wave function is either the symmetric ( $\varphi_+$ ) or

antisymmetric ( $\varphi_-$ ) combination:

$$\varphi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) \pm \varphi_1(\mathbf{r}_2)\varphi_2(\mathbf{r}_1)] \quad (13.4)$$

Notice that  $\varphi_-$  vanishes when  $\mathbf{r}_1 = \mathbf{r}_2$ .



Now consider the Coulomb repulsion between the electrons:

$$U(\mathbf{r}_1 - \mathbf{r}_2) = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} . \quad (13.5)$$

If we calculate the change in energy  $\delta E$  in first-order perturbation theory, then

$$\delta E = \langle \varphi(1, 2) | U | \varphi(1, 2) \rangle . \quad (13.6)$$

The spin part is diagonal and thus contributes only in the sense of telling us whether  $\varphi_+$  or  $\varphi_-$  is the spatial wave function. We can label the two possibilities  $\delta E_{\pm}$ , with

$$\delta E_{\pm} \equiv \langle \varphi_{\pm} | U | \varphi_{\pm} \rangle = \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 |\varphi_1(\mathbf{r}_1)|^2 |\varphi_2(\mathbf{r}_2)|^2 \frac{2e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \pm \frac{J}{2} , \quad (13.7)$$

which defines the *exchange integral*  $J$  as

$$J \equiv \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \varphi_1(\mathbf{r}_2)\varphi_2(\mathbf{r}_1)\varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) \frac{2e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} > 0 . \quad (13.8)$$

So what do we learn?

- $J$  is positive definite (for our example), so  $J > 0$  means *lower* energy for  $E_-$ .
- But  $E_-$  implies  $\varphi_-$ , which means an anti-symmetric spatial wave function and therefore symmetric (e.g., parallel) spins.
- Thus the parallel alignment is preferred! In short, if the spins are parallel, the electrons are less likely to be near each other, and therefore will have a lower average potential energy from electrostatic repulsion.
- We can evaluate Eq. (13.8) using hydrogenic wave functions to get a quantitative estimate for  $J$ .

### c. Variational Monte Carlo

Here's the basic idea of the variational method for estimating quantum mechanics energies. If we have a guess  $\psi_{\alpha}(\mathbf{R})$  for a ground-state wavefunction that depends on a set of parameters  $\alpha = (\alpha_1, \dots, \alpha_s)$  and the positions of  $N$  particles  $\mathbf{R} = \{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ , then

$$\langle E_{\alpha} \rangle = \frac{\langle \psi_{\alpha} | H | \psi_{\alpha} \rangle}{\langle \psi_{\alpha} | \psi_{\alpha} \rangle} = \frac{\int d\mathbf{R} \psi_{\alpha}^*(\mathbf{R}) H(\mathbf{R}) \psi_{\alpha}(\mathbf{R})}{\int d\mathbf{R} \psi_{\alpha}^*(\mathbf{R}) \psi_{\alpha}(\mathbf{R})} \quad (13.9)$$

is an estimate of the energy, which is always an *upper bound* to the exact ground-state energy. Thus our best estimate is obtained by *minimizing*  $\langle E_{\alpha} \rangle$  with respect to the parameters  $\alpha$ .

Start with the simplest case, that of a single particle (i.e.,  $N = 1$ ). We might take as our trial wavefunction

$$\psi_{\alpha}(\mathbf{r}) = C(1 + br^2)e^{-ar}, \quad (13.10)$$

in which case  $\alpha = \{a, b\}$ . Note that the normalization constant  $C$  is irrelevant (and is *not* an element of  $\alpha$ ), since it cancels between numerator and denominator in evaluating  $\langle E_{\alpha} \rangle$ . We calculate the three-dimensional integrals in Eq. (13.9) and minimize the result with respect to the two parameters  $a$  and  $b$ .

The expression for  $\langle E_{\alpha} \rangle$  is just a multi-dimensional integral, with  $3N$  integrations in general. In realistic systems, the wave function  $\langle \mathbf{R} | \psi_{\alpha} \rangle$  for the many-body system is very small for most configurations  $\mathbf{R}$  (in this context, a “configuration” just means a specification of the positions of all the particles). Solution? Use importance sampling with the Metropolis algorithm!

The strategy is to use a collection of “random walkers” to build a set of  $\mathbf{R}$ 's. Boltzmann weighting is not relevant in this problem; what do we use instead? We need a probability distribution, which naturally suggests using the wave function squared. Thus we'll distribute points according to

$$\rho(\mathbf{R}) \equiv \frac{\psi_{\alpha}^2(\mathbf{R})}{\int d\mathbf{R} \psi_{\alpha}^2(\mathbf{R})} \quad (13.11)$$

(we assume  $\psi_{\alpha}$  is real). Note that

$$\int \rho(\mathbf{R}) d\mathbf{R} = 1 \quad (13.12)$$

by construction. If we define the *local energy*

$$E_L(\mathbf{R}) = \frac{H\psi_{\alpha}(\mathbf{R})}{\psi_{\alpha}(\mathbf{R})}, \quad (13.13)$$

then

$$\langle E \rangle = \int d\mathbf{R} E_L(\mathbf{R}) \rho(\mathbf{R}) \quad (13.14)$$

is the integral we treat with Metropolis. This is carried out for a practice case in Session 13.

The basic procedure is to compare probabilities when deciding whether to move from  $\mathbf{R}$  to  $\mathbf{R}'$ . That is, we compare  $\rho(\mathbf{R}')$  to  $\rho(\mathbf{R})$  and use the Metropolis decision algorithm:

1. If  $\rho(\mathbf{R}')/\rho(\mathbf{R}) > 1$ , then we accept the move;
2. if  $\rho(\mathbf{R}')/\rho(\mathbf{R}) < 1$ , then we generate a uniform random number  $s \in [0, 1]$  and accept the step if

$$\frac{\rho(\mathbf{R}')}{\rho(\mathbf{R})} \geq s . \quad (13.15)$$

See the `variational_SH0` code for a sample implementation.

#### d. References

- [1] R.H. Landau and M.J. Paez, *Computational Physics: Problem Solving with Computers* (Wiley-Interscience, 1997).
- [2] M. Hjørth-Jensen, *Computational Physics*. These are notes from a course offered at the University of Oslo. See the 780.20 webpage for links.
- [3] W. Press *et al.*, *Numerical Recipes in C* (Cambridge, 1992). Individual chapters are available online from <http://lib-www.lanl.gov/numerical/bookcpdf.html>. There are also versions for Fortran and C++.