14. 780.20 Session 14

a. Simulated Annealing

Standard optimization methods are very good at finding local minima near where the minimization was started, but not usually good at finding the *global* minimum. Finding the global minimum of a function (such as the energy) is often (but not always) the goal. One strategy using conventional minimizers is to run multiple trials with the minimization started at different places in the parameter space (perhaps chosen at random) and then to keep the best minimum found of all the trials.

An alternative approach is to adapt the Metropolis Monte Carlo algorithm for generating a canonical Boltzmann distribution of configurations at a temperature $T$ to mimic how physical systems find their ground states (i.e., the energy minimum at $T = 0$). At high temperature (which means $kT$ large compared to characteristic energy spacings), the equilibrium distribution will include many states. If the system is cooled slowly, then it will have the opportunity to explore many states and then settle into the lowest energy state as $T$ goes to zero. This is called annealing. If the system is cooled quickly, it can get stuck in a state that is not the minimum ("quenching"); this is analogous to the routines we looked at in Session 10, which rapidly go "downhill" but only to local minima.

The strategy of *simulated annealing* is to mimic the annealing process by treating the function to be minimized as an energy (it might actually be an energy!), introducing an artificial temperature $T$, and generating a sequence of states in a canonical distribution via the Metropolis algorithm. Then we lower the temperature according to a "schedule" (this just means according to a definite pattern) and let the system settle into (hopefully!) a configuration that minimizes the energy.

In practice this is not so easy:

- The problem needs to be cast into a form appropriate for this technique. This means we need to have a description of possible configurations of the system and then a way to change the configuration randomly (i.e., the analogs to specifying all the spins on a lattice and generating a new configuration by randomly flipping a spin). At the same time we need to identify an appropriate energy function to be minimized; this may be immediate, if the problem is to minimize a function, or less obvious, if the problem is to find a solution to a problem such as the Traveling Salesman [1].

- We need to devise an appropriate annealing schedule for the control parameter $T$. For example, do we change $T$ after 10 or 100 or 1000 or ? random changes in the configuration? And how much do we change it each time? These are critical questions to the success of the procedure. If there is a physical connection to the problem, we may be able to use physics insight to determine the appropriate scales. Often it is more a trial-and-error procedure.

- For continuous control parameters, as opposed to combinatoric problems, the (common) possibility of long, narrow valleys in parameter space is a serious problem. If one takes steps at random, the most likely step will be uphill rather than along the valley. So one needs to modify the basic strategy, as discussed in Ref. [1].
We'll apply simulated annealing to one artificial and one real (i.e., physical) problem. The artificial problem is just the global minimization of a one-dimensional function:

\[ f(x) = e^{-(x-1)^2 \sin(8x)} , \tag{14.1} \]

which has multiple local minima. We'll compare a standard minimization routine (from GSL) to simulated annealing. You'll need to adjust the simulated annealing control parameters to make it work effectively. The second problem is the shape of molecules built from sodium (Na) and chlorine (Cl) atoms. This problem is described in a Session 10 handout from the book *An Introduction to Computational Physics* by T. Pang, in a section entitled “Geometric structures of multicharge clusters.” The idea is that one can write a potential energy function that depends on the relative positions of the elements of the clusters (here Na and Cl atoms). The parameters of the function are taken from experiment or theoretical calculations. The kinetic energy can be ignored, so the arrangement of the cluster is determined by minimizing the energy. There are many local minima corresponding to configurations that might be metastable but do not have the very lowest energy.

b. References