John Hertz QC2024: From Spin Glasses to Deep Networks

One can learn a lot about how neural networks for AI work by using the tools of statistical physics, particularly the theory developed to describe spin glasses. In these lectures I will begin with a brief survey of spin glass physics and then show how theoretical methods developed there can help us understand learning in neural networks, from the simplest models to multilayer nets with extensive nonlinear interactions between units.

1. The prototypical spin glass: the Sherrington-Kirkpatrick model

It was proposed as a model for which mean field theory would be exact, but that mean field theory turned out to have a form no one had imagined. I'll sketch this bit of history and then focus on how to study its dynamics.

2. Another kind of model: the p-spin glass

Instead of spins just interacting in pairs, how about letting them interact in larger groups? This is the "p-spin" glass, and it is quite different (and in many ways simpler!) than the Sherrington-Kirkpatrick model. We will study its asymptotic dynamics and the phenomenon of "weak ergodicity breaking".

3. Neural networks and how to train them

I'll start with the simplest kind of network, called the perceptron. It has a single layer of connections, and I'll show how to use spin glass theoretical methods to calculate how many arbitrary input-output pairs it can learn. I'll then go on to describe how we train networks with multiple layers and ones with interactions within the layers.

4. The learning transition in deep recurrent networks

I'll treat the transition between imperfect and perfect learning as one increases the size of a deep (many-layer) net with recurrence (interactions between units within layers). We'll see that this transition appears to be in the same class as the those we have seen in the perceptron and several kinds of spin glasses. Finally, I'll try to make a quantum connection and describe how one might begin to make a quantum version of such a network.