Carina Curto – White paper for Physical and Mathematical Principles of Brain Structure and Function Workshop (Arlington, VA, May 2013).

A major obstacle impeding progress in brain science is the lack of beautiful models.

Let me explain. Many will agree that the existing (and impending) deluge of data in neuroscience needs to be accompanied by advances in computational and theoretical approaches – for how else are we to "make sense" of these data? What such advances should look like, however, is very much up to debate. How much detail should we be including in our models? How closely shall we try to match experimental observations? How well can we defend the biological realism of our theories? These considerations often put theorists on the defensive, and – ironically – they end up giving experimentalists relatively few ideas to guide their own research. They have also steered the field towards ever more complicated models intended more for exploration with a supercomputer than for mathematical analysis. These models are ugly, and ugliness in science has consequences. Ugly models are cumbersome to describe, the essential features are difficult to understand, and any results obtained are nearly impossible to build upon. When is the last time you took an ugly model as the starting point for a theoretical project or as inspiration for an interesting experiment? More often than not, it is abstract theoretical ideas – not concrete mathematical models – that are guiding neuroscience research, because the models are simply too ugly to be attractive or useful.

What are beautiful models? We have some examples in neuroscience, but far fewer than in physics. The Ising model is a beautiful model. It is simple to describe yet behaviorally rich, and it captures the essence of an interesting physical phenomenon (phase transitions in magnets). It is also mathematically tractable – i.e., one can prove theorems about the Ising model and do analytic computations. Interestingly, the model is known to be "wrong," but it nevertheless enjoys a secure reputation as one of the leading instigators of progress in statistical mechanics, and has recently gained some newfound influence in neuroscience. The Hopfield model is also a beautiful model, and the ideas of attractor neural networks – however flawed – continue to influence the way neuroscientists think about persistent activity and information storage in neural networks. It is worth noting that this model also had a nontrivial (and positive) sociological effect on theoretical neuroscience, inspiring a generation of talented physicists to enter the field.

Beautiful models are simple to describe and can be mathematically analyzed. They contain important kernels of "truth" in them, even if they are not messy enough to be completely accurate. The general principles behind these models are often easy to understand, making them attractive candidates to be tested experimentally. Equally important is the fact that beautiful models can synchronize a theoretical community around a coherent line of research, leading to work that can be understood and built upon by future generations. This is tremendously important to ensure that theory in neuroscience develops into more than a collection of one-off "spin" papers, whose lasting impact is rapidly swamped by fluctuations in fashionable trends. Beautiful models can also inspire the most talented young people with training in physics, mathematics, and computer science to enter a new field. We need these people in neuroscience, and not only as theorists. They will be difficult to recruit, however, if there is nothing beautiful to capture their imaginations and draw them in.