This is the graph of the perceptron being trained over 100 linearly-separable points. Most of the learning happens at the beginning, with the last half or so of the iterations used to fine-tune the line. Because the points are generated so as to be linearly separable, the perceptron does end up classifying all the data perfectly.