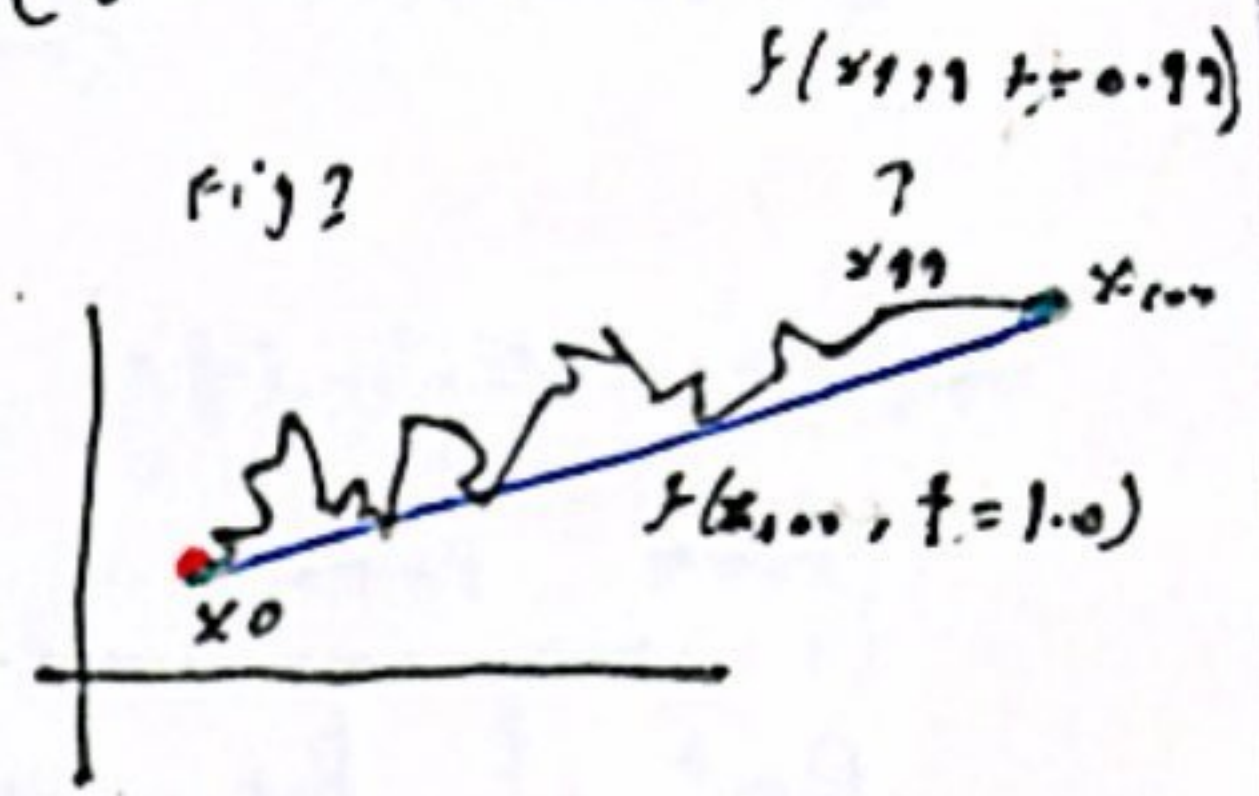


P-2 Vector Fields (Gradient Descent)

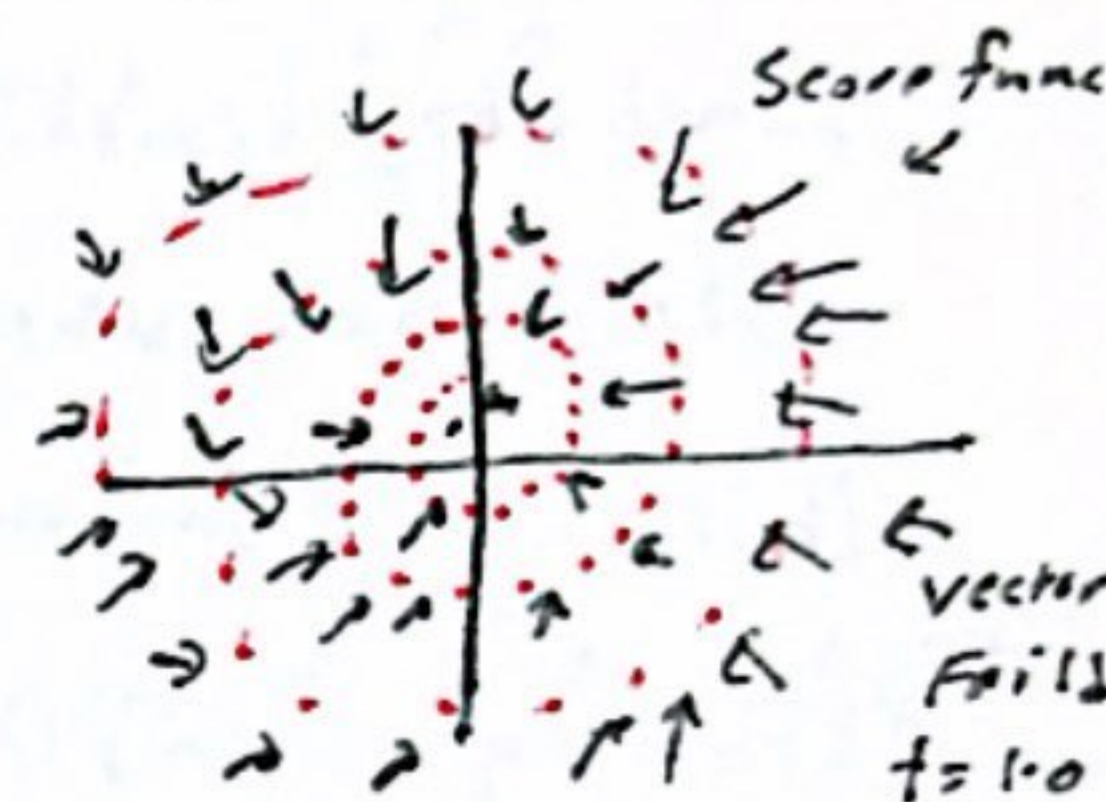
- if we consider a 100 step random walk the point at the end is the final end of point. here we want to go back to original point. here that's the vector from x_{100} to x_0 (original point).



$$E[x_{11} - x_{100} | x_{100}] = \frac{E[x_0 - x_{100} | x_{100}]}{100}$$

- learning to predict the noise added in the final step of our walk is equivalent to learning to predict the total noise added, divided by the number of steps taken. by training model to predict directly x_0 and not x_{99} we reduce the variance in training examples allowing model to learn efficiently without changing objective

- So for each point in our space our model learns the direction pointing back towards the original data distribution AKA, Score Function.



- in practice these learned direction depend on how much noise we add to original data, at 100 steps most points are far from original point so the model learns to move these points back in the general direction of our spiral. the more training we do the better the vector field is at pointing back our noisy points to original data to help this process look at Fig 2, instead of just passing coordinates of points to our model we also pass in the time var = # of steps taken in walk this helps our model learn a better vector field. Our equation will now be better at predicting vectors for all t values.

- Our img gen process now looks like points starting at random locations are working its way back to form a spiral.

- to visualize the random added noise with no added noise all points converge at the center of the spiral on the inside edge instead of nicely fitting the spiral hence why we add random noise so it fits the distribution nicely. in the space of imgs args look blurry, think of different parts of the spiral as parts of img like trees 82 sand, desert, etc when we remove noise we are being the avg of all possible imgs and looking bad

