Optimization
Algorithms

RMSprop

deeplearning.ai

RMSprop

w_1, w_2, w_3

$w_3, w_4, \ldots$

slow

fast
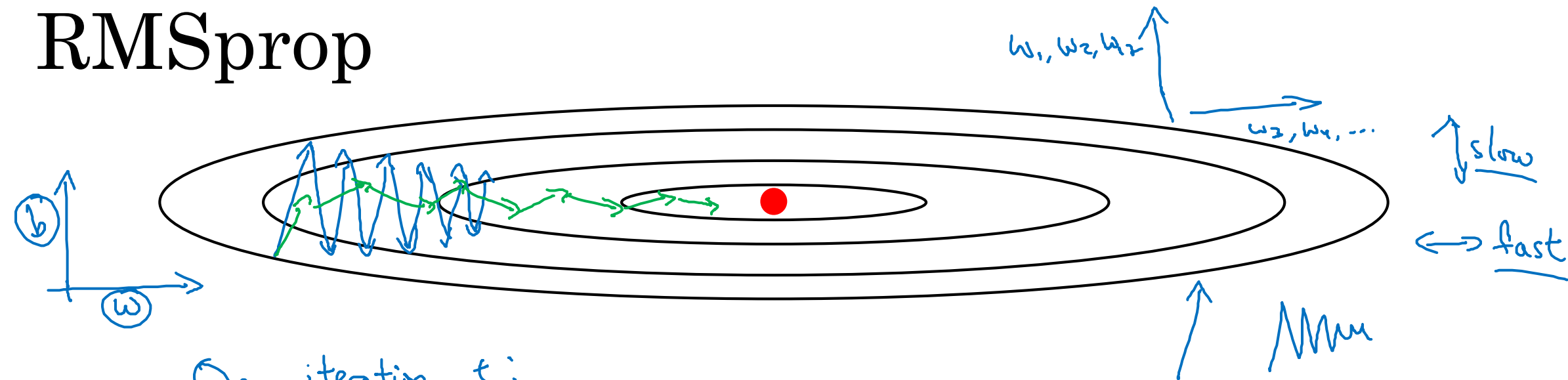
On iteration t:

Compute dW, db on current mini-batch

$S_{dW} = \beta_2 S_{dW} + (1-\beta_2) dW^2$ ← element-wise ← small

$\rightarrow S_{db} = \beta_2 S_{db} + (1-\beta_2) db^2$ ← large

$W := W - \alpha \dfrac{dW}{\sqrt{S_{dW}} + \varepsilon}$

$b := b - \alpha \dfrac{db}{\sqrt{S_{db}} + \varepsilon}$

$\varepsilon = 10^{-8}$

Andrew Ng