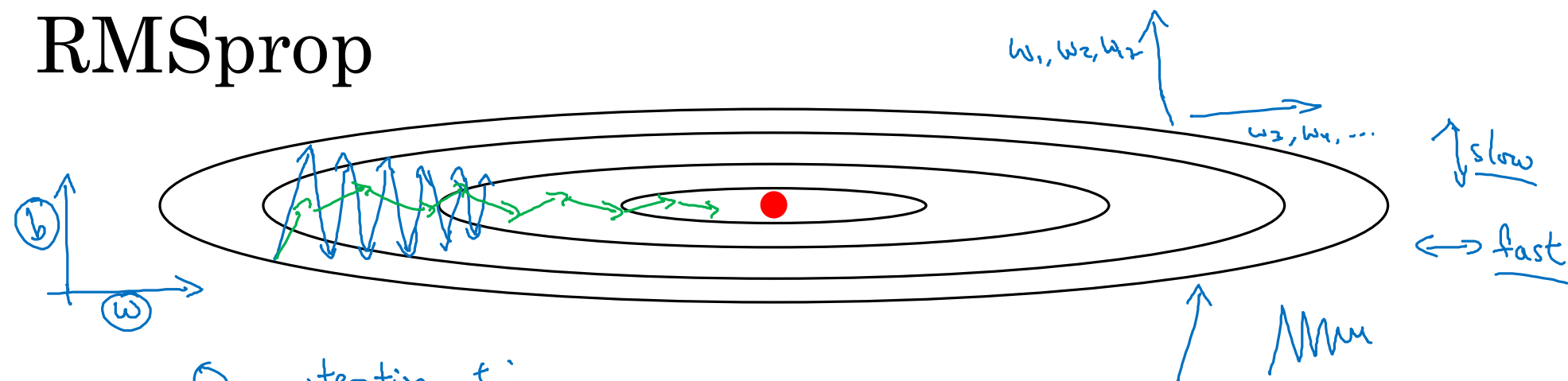deeplearning.ai

Optimization
Algorithms

RMSprop

RMSprop

On iteration t:

Compute dW, db on current mini-batch

$$S_{dW} = \beta_2 S_{dW} + (1-\beta_2) dW^2 \longleftarrow \text{small}$$

element-wise

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

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

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

$$\varepsilon = 10^{-8}$$

w_1, w_2, w_2

$w_2, w_4, \ldots$

$\updownarrow$ slow

$\leftrightarrow$ fast

Andrew Ng