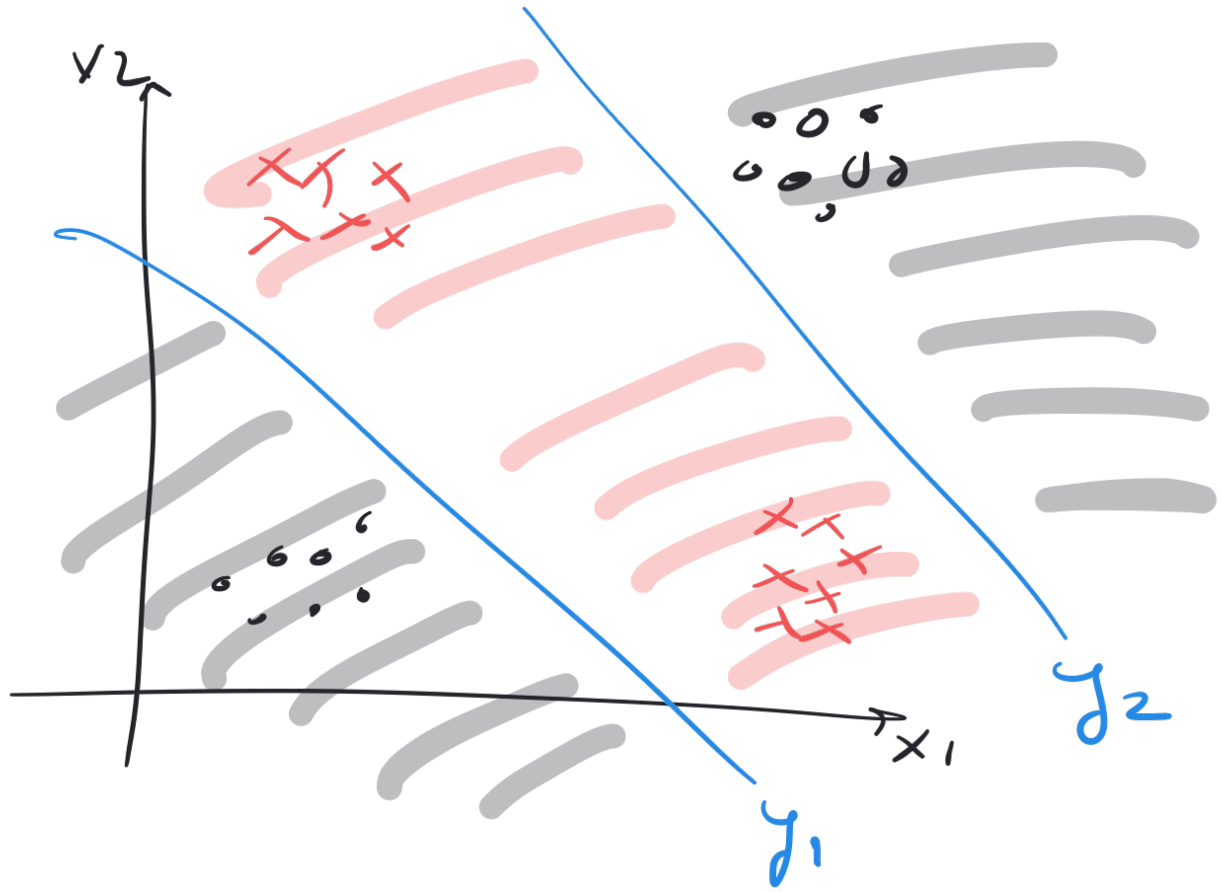
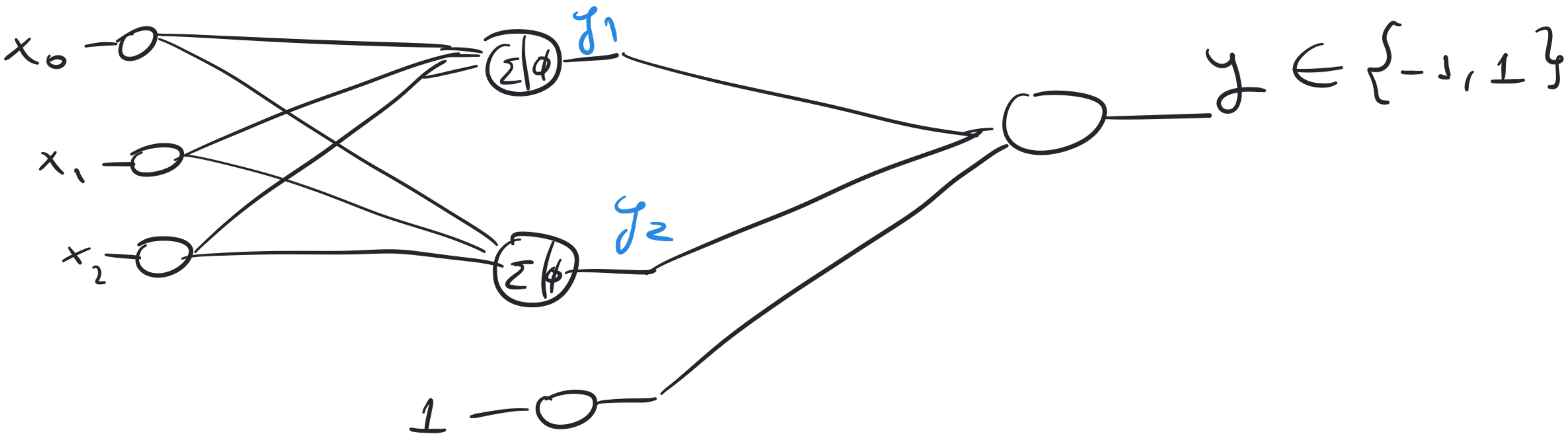
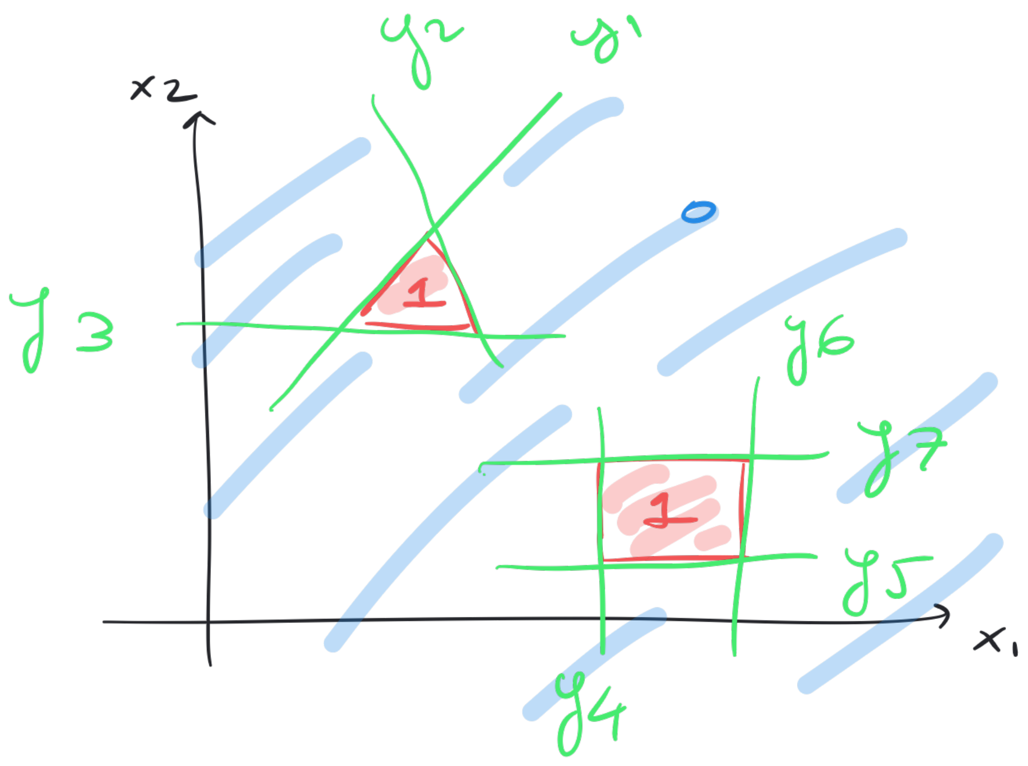


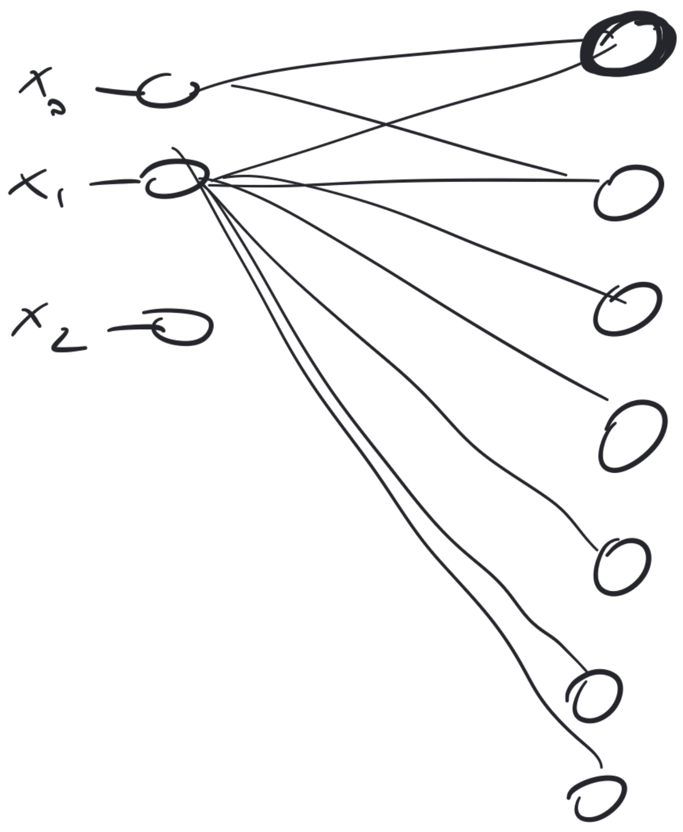
XOR

or

Exclusive OR







...

to be continued when we discuss multi-Layer Perceptron (MLPs).

$$y = \phi(w^T x + w_0)$$

where $\phi(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$

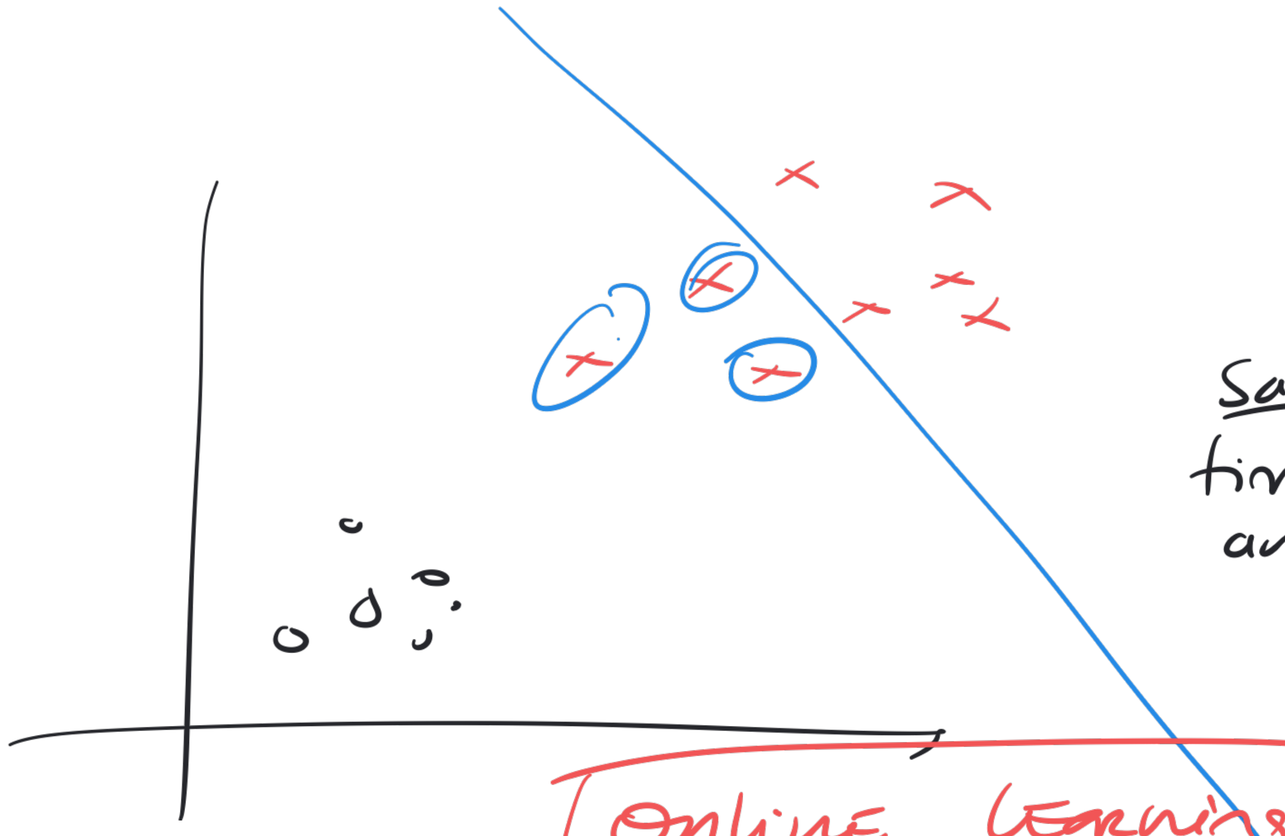
Objective fct: $E_p(w, w_0) = - \sum_{n \in M} (w^T x_n + w_0) t_n$

M = set of misclassified points

Learning algorithm: Gradient Descent

$$w^{(t+1)} \leftarrow w^{(t)} - \eta \cdot \frac{\partial J(w, w_0)}{\partial w^{(t)}}$$

$$w_0^{(t+1)} \leftarrow w_0^{(t)} - \eta \cdot \frac{\partial J(w, w_0)}{\partial w_0^{(t)}}$$



it uses **one** sample at a time to make an update to w and w_0

Online Learning

$$\frac{\partial J(w, w_0)}{\partial w} = -x_n \cdot t_n$$

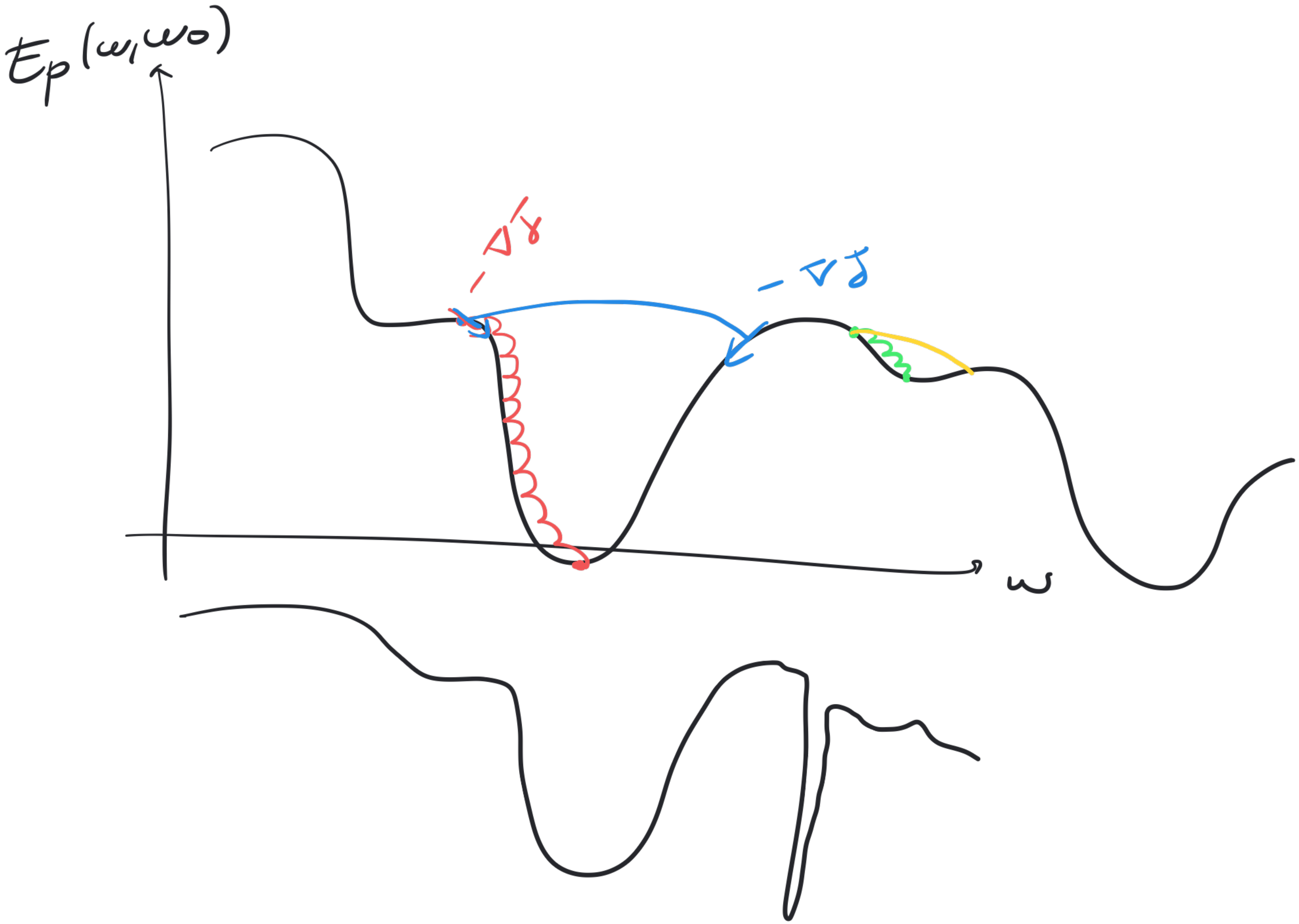
$$\frac{\partial J(w, w_0)}{\partial w_0} = -t_n$$

$t \equiv$ iteration

Update Eq. :

$$w^{(t+1)} = w^{(t)} + \eta \cdot x_n \cdot t_n$$

$$w_0^{(t+1)} = w_0^{(t)} + \eta \cdot t_n$$



$E_p(w, w_0)$
$-\nabla\gamma$
$-\nabla\gamma$
w