

$$\nabla_{\theta} \mathcal{L}(\theta, \{x, y\})$$

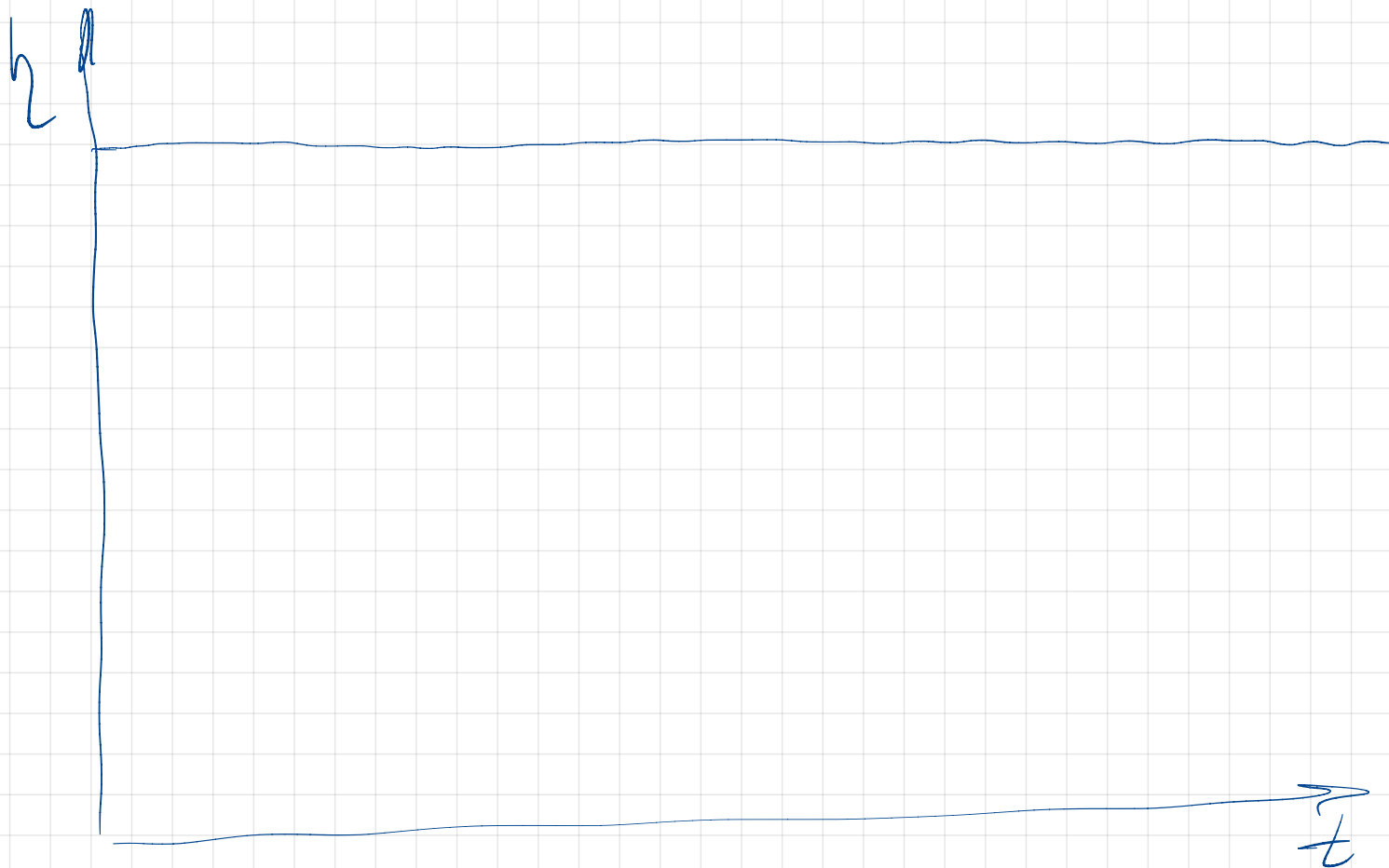
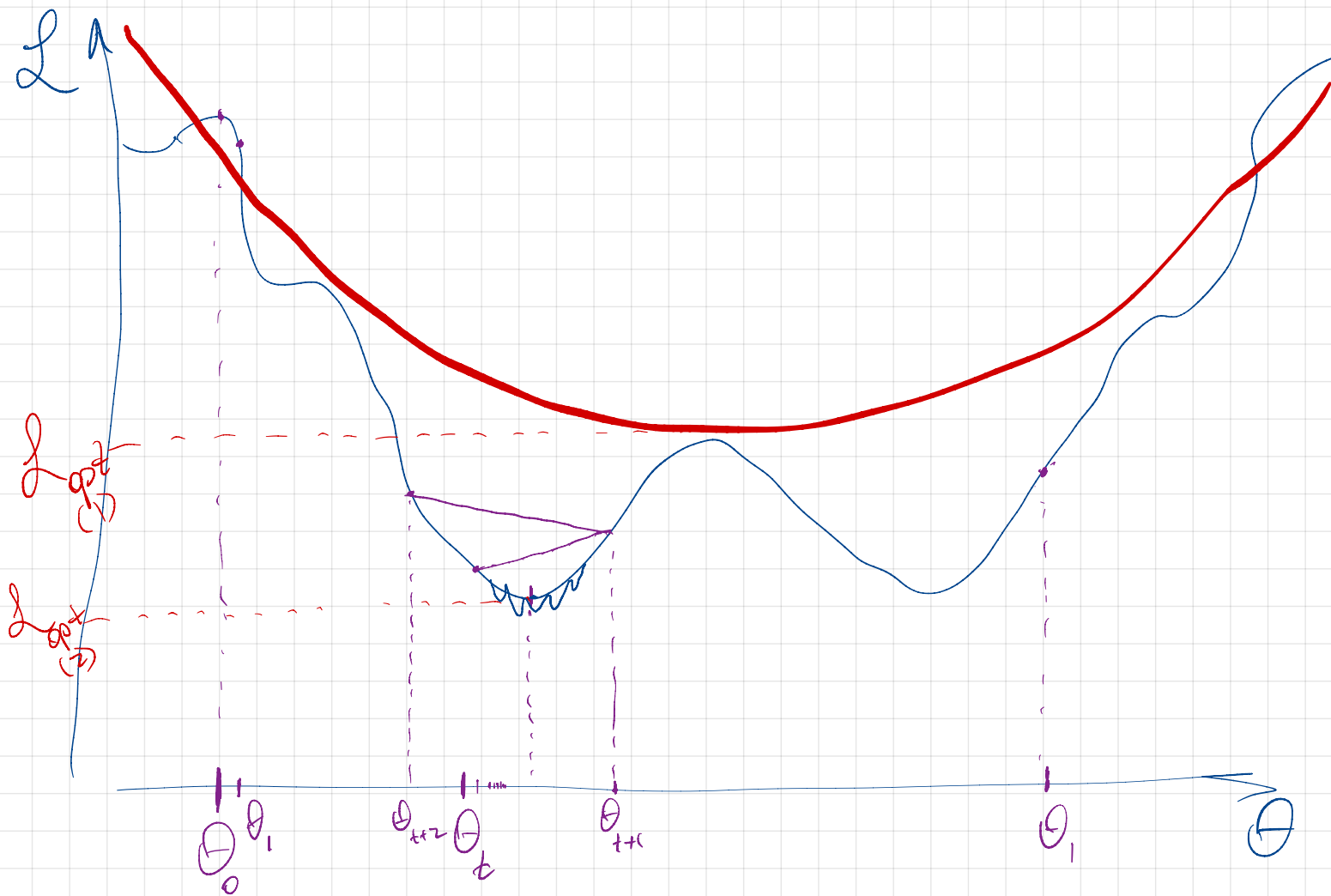

---

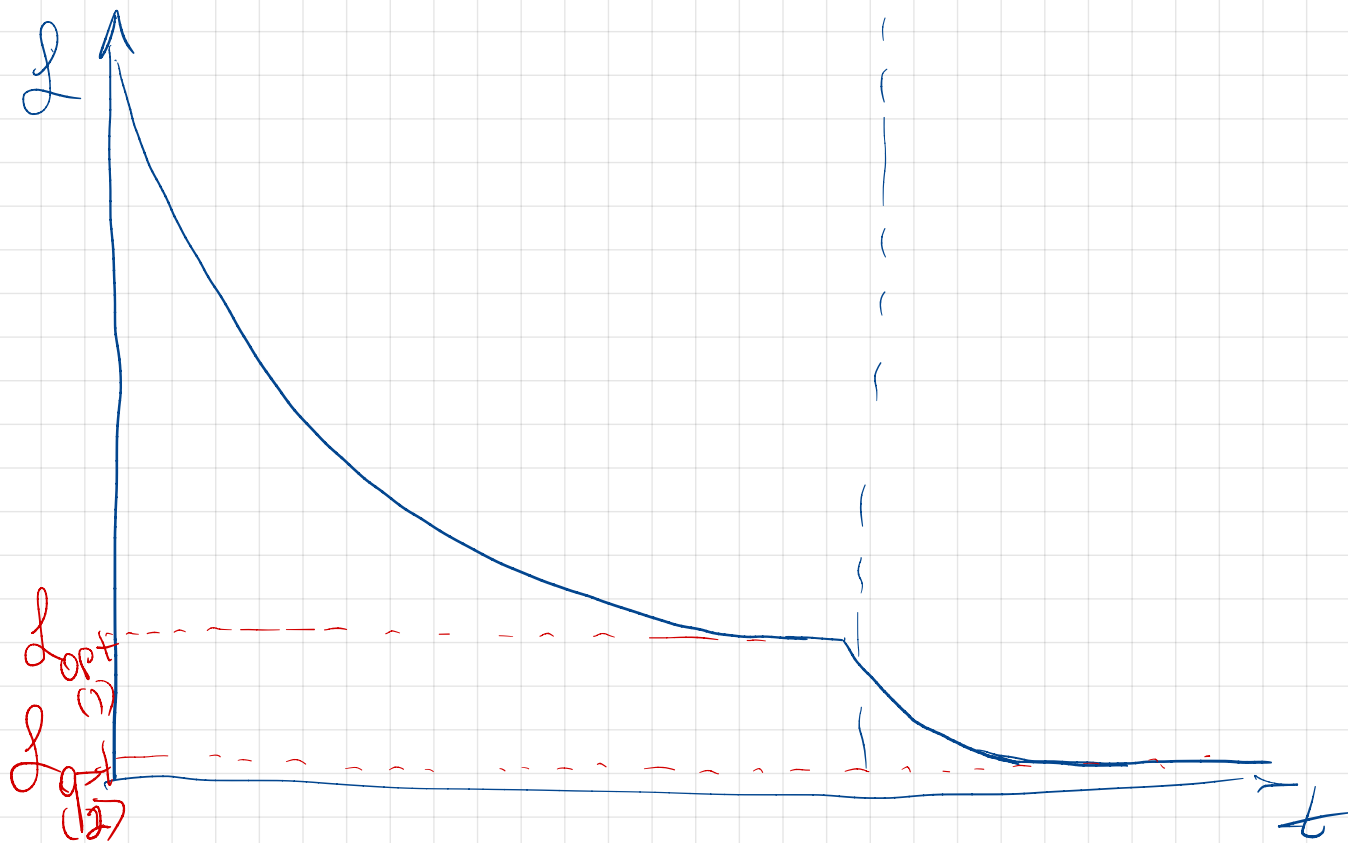
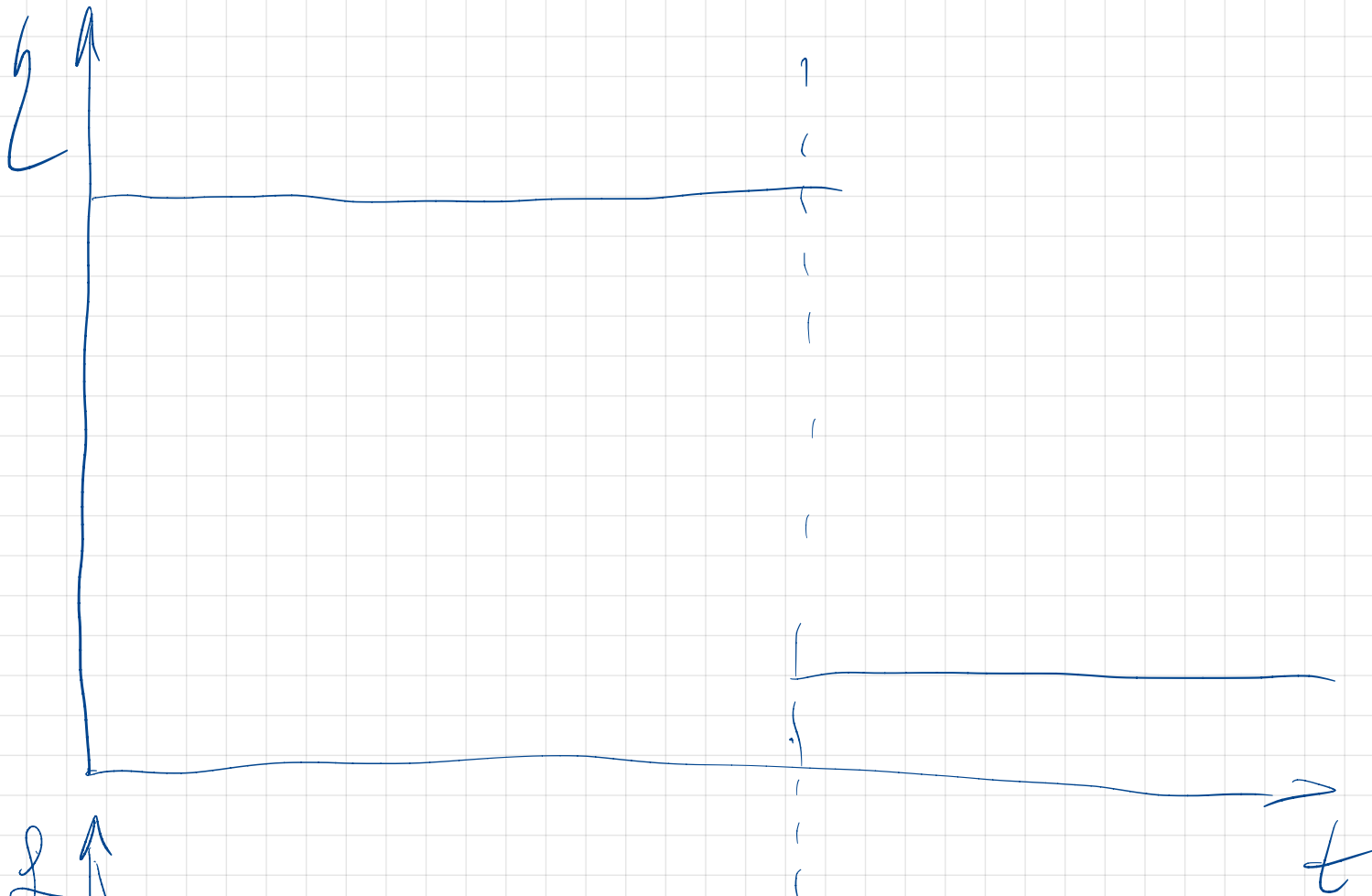
problem:  $\mathcal{L}(\theta, \{x, y\}) \Rightarrow \nabla_{\theta} \mathcal{L}(\theta, \{x, y\})$

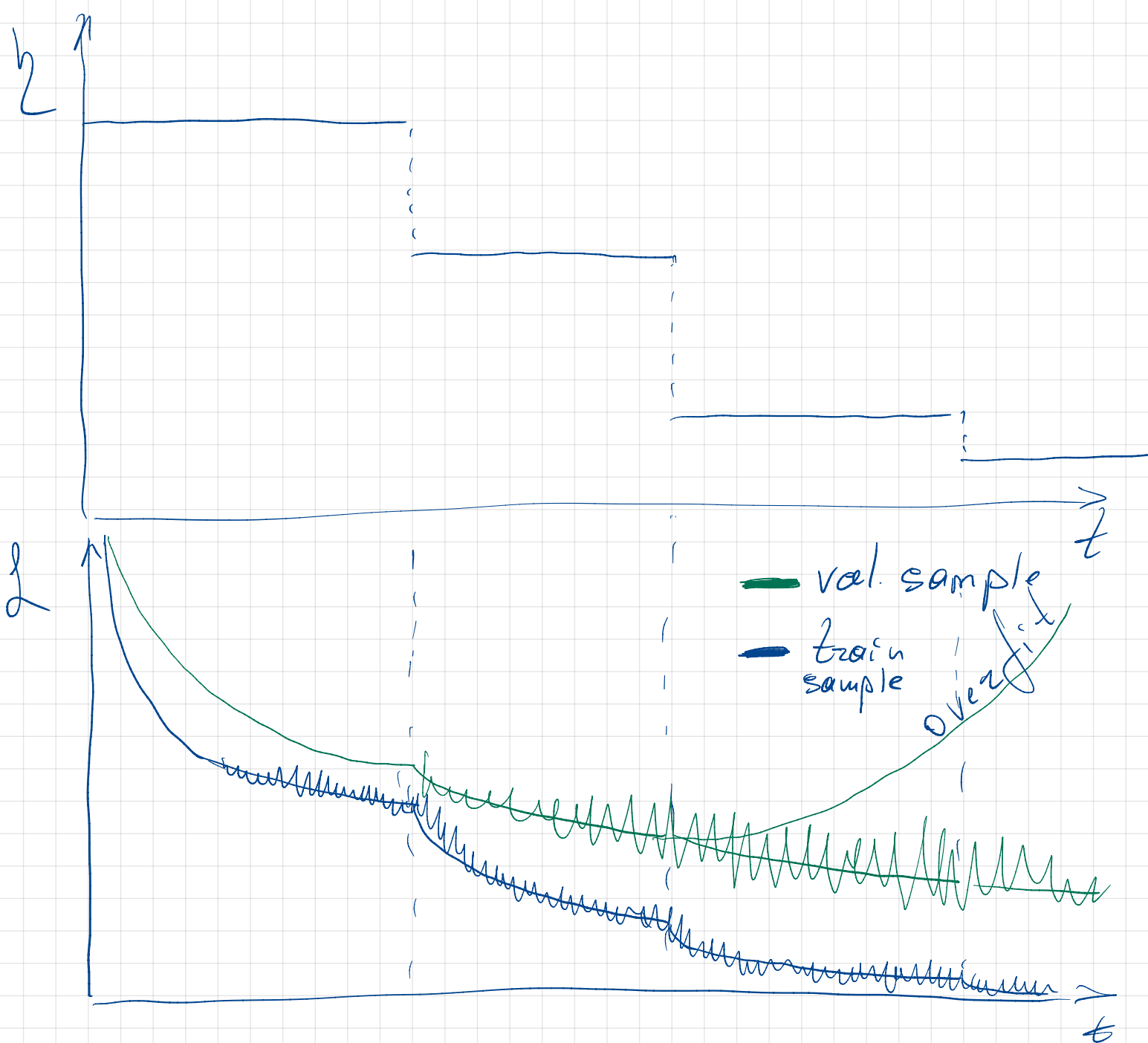
$$\theta_0 \sim \dots (?)$$

<ul style="list-style-type: none"> <li>• <math>m_t = (1 - \beta_1) \nabla_{\theta} \mathcal{L}_{t-1} + \beta_1 m_{t-1}</math></li> <li>• <math>s_t = (1 - \beta_2) (\nabla_{\theta} \mathcal{L}_{t-1})^2 + \beta_2 s_{t-1}</math></li> <li>• <math>\theta_t = \theta_{t-1} - \eta \frac{m_t}{\sqrt{s_t + \epsilon}}</math></li> </ul>	<div style="display: flex; align-items: center;"> <div style="font-size: 4em; margin-right: 10px;">}</div> <div> <p>AdaGrad</p> <p><math>s_t^* = \max(s_t, s_{t-1})</math></p> </div> </div>
---	--

↙







$$f = k_1 e^{-k_2 t} + \epsilon$$

