b) POLYAK gradient descent

$$\boxed{x_{k+1} = x_k - \gamma \cdot \nabla f(x_k) + \mu(x_k - x_{k-1})}$$

I) 2-periodic

Polyak GD may result in non-constant 2-periodic seque[nce]

We pick a quadratic function: $f(x) = ax^2$ $\quad \nabla f(x) = 2ax$

$x_1 = 1$

$x_2 = -1$

$x_3 = 1$ $\quad$ ① $x_3 = x_2 - \gamma \cdot \nabla f(x_2) + \mu(x_2 - x_1) = 1 \Rightarrow x_3 = -1 - \gamma \cdot \nabla f(-1) + \mu(-1-1) = 1$

$x_4 = -1$ $\quad$ ② $x_4 = x_3 - \gamma \cdot \nabla(x_3) + \mu(x_3 - x_2) = -1 \Rightarrow x_4 = 1 - \gamma \cdot \nabla f(1) + \mu(1+1) = -1$

$\vdots$

① $\gamma = \frac{-2\mu - 2}{\nabla f(-1)} = \frac{-2\mu - 2}{-2a} = \underline{\underline{\frac{\mu+1}{a}}}$ $\quad$ ② $\gamma = \frac{2\mu+2}{\nabla f(1)} = \underline{\underline{\frac{\mu+1}{a}}}$ $\quad \checkmark$

We see that for a fixed a and μ the learning rate is constant which means we prove[d] we created a 2-periodic sequence. EXAMPLE: $\underline{\underline{f(x) = x^2}}$, $\underline{\underline{\nabla f(x) = 2x}}$, $\underline{\underline{\mu = 1}}$, $\gamma = \mu + 1 = \underline{\underline{2}}$

II) 3-periodic

Polyak GD may result in nonconstant 3-periodic sequence.

We show example from lecture notes:

$$f'(x) = \begin{cases} 25x & x < 1 \\ x + 24 & 1 \leq x \leq 2 \\ 25x - 24 & 2 < x \end{cases}$$

$x_1 \approx 0.65$

$x_2 \approx -1.8$

$x_3 \approx 2.12$

$x_4 \approx 0.65$

$\vdots$