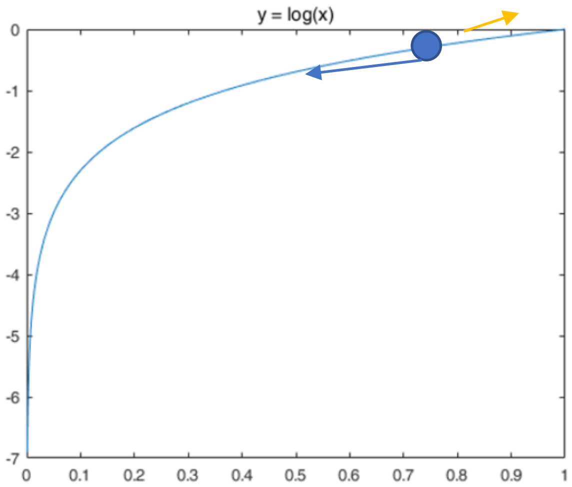
Unrolled gan

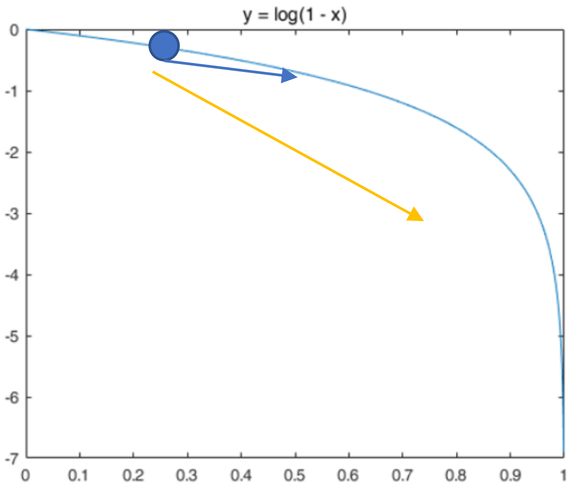
using the optimal discriminator(in the generator's objectives) impossible

- k steps of optimizing D and one step of optimizing G
- D: near optimal solution
- G: change slowly enough

x: data variable , z: noise variable, G(z): fake data

$$\min_G \max_D V(D, G) = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

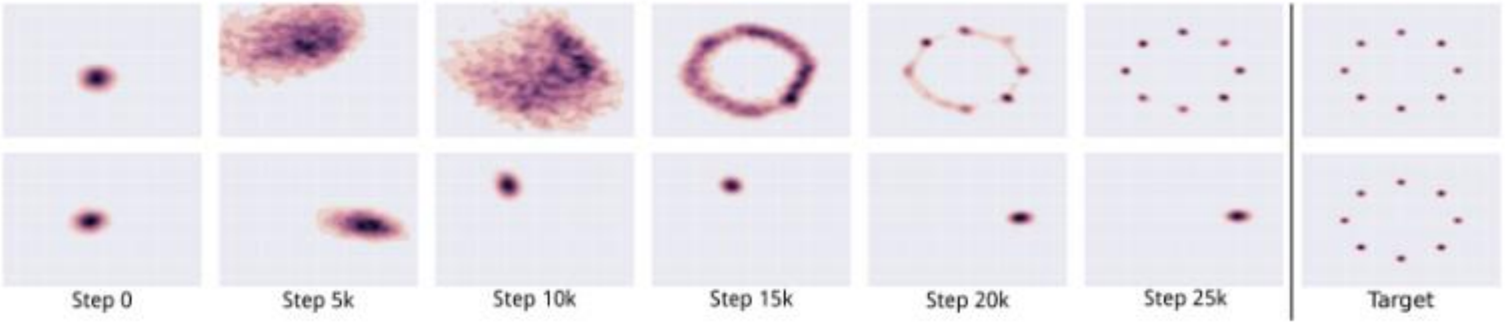




Unstability of gan

- Mode collapsing or dropping (generator)
- Generator and discriminator oscillating during training
- No learning when the power between generator and discriminator is unbalanced.

Mode collapsing



Unrolled GAN(위) vs. standard GAN(아래)

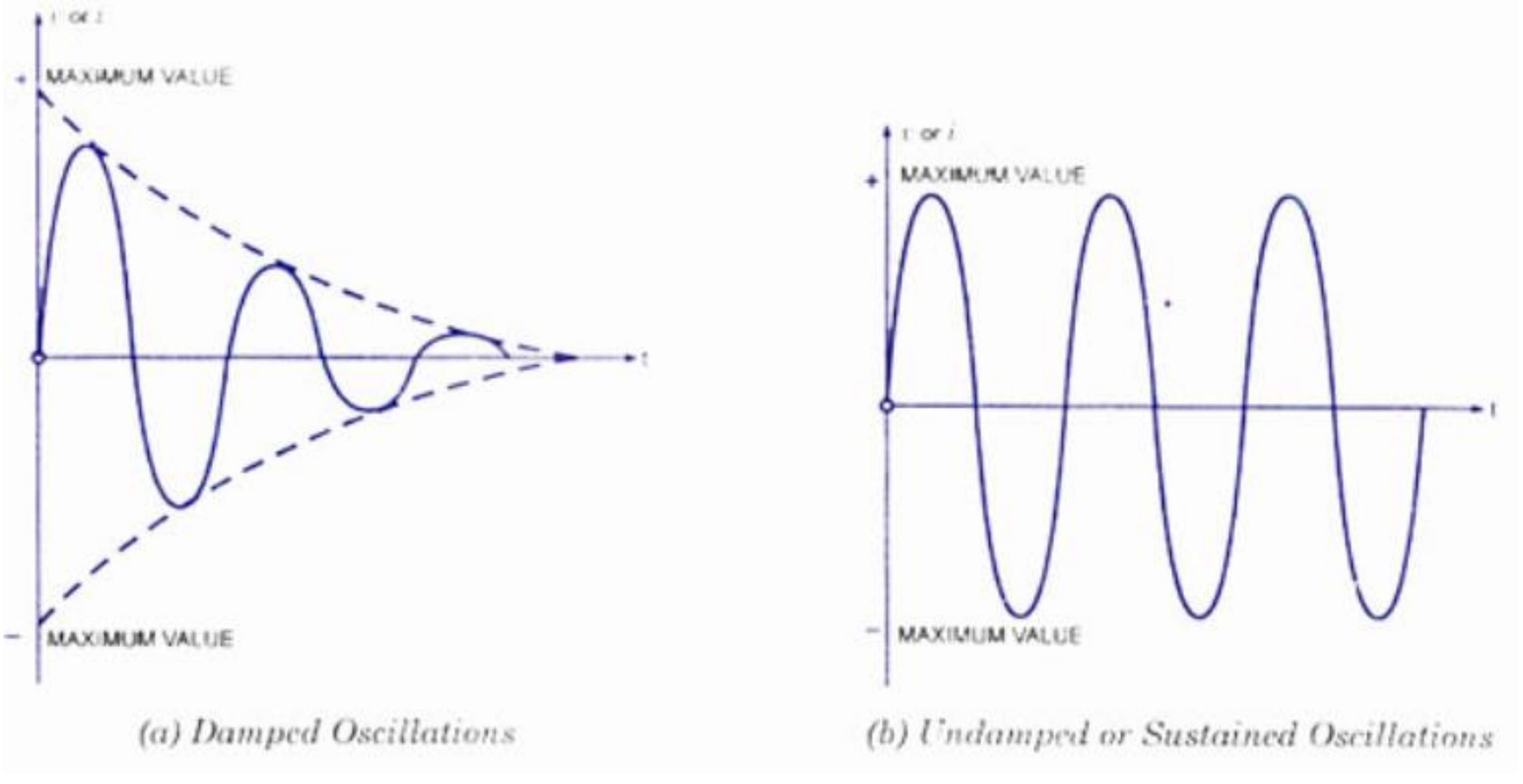
Neural network의 입장에서는 이러한 **minimax problem**과 아래와 같은 **maximin problem이 구별이 되지 않습니다:**

$$G^* = \min_G \max_D V(G, D).$$

$$G^* = \max_D \min_G V(G, D).$$

G에 대한 minimization 문제가 먼저 있기 때문에 generator의 입장에서는 **현재 고정되어있는 discriminator (non-optimal)**가 가장 헷갈려 할 수 있는 sample 하나만 즉, **value V를 가장 최소화할 수 있는 mode 하나만**을 내보내면 그만입니다.

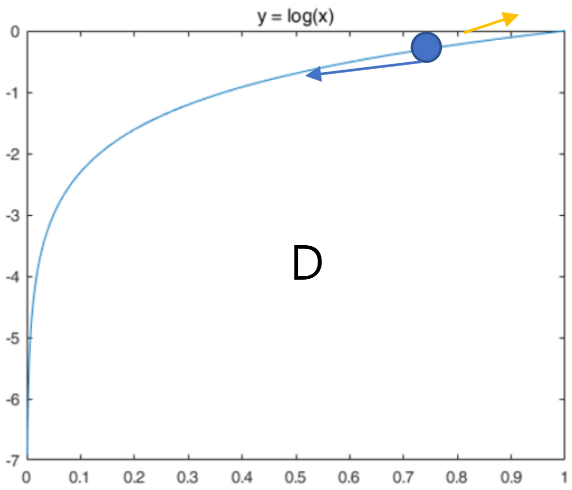
Undamped oscillation

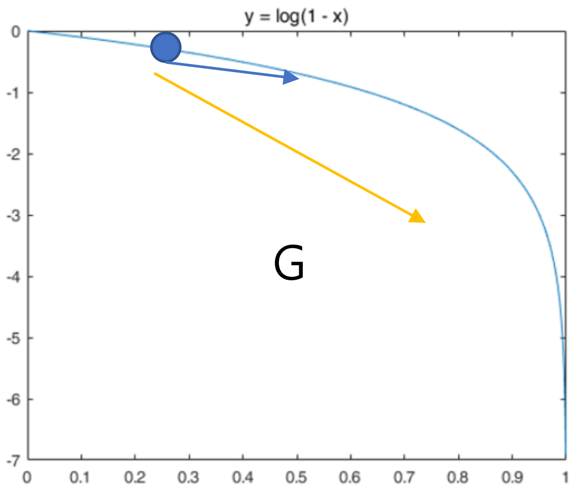


Damped and Umdamped Oscillations

f (θG,θD) is typically very far from convex in θG and concavein θD

$$f(\theta_G, \theta_D) = \min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$





$$D_G^*(\boldsymbol{x}) = \frac{p_{data}(\boldsymbol{x})}{p_{data}(\boldsymbol{x}) + p_g(\boldsymbol{x})}$$

Not concavein

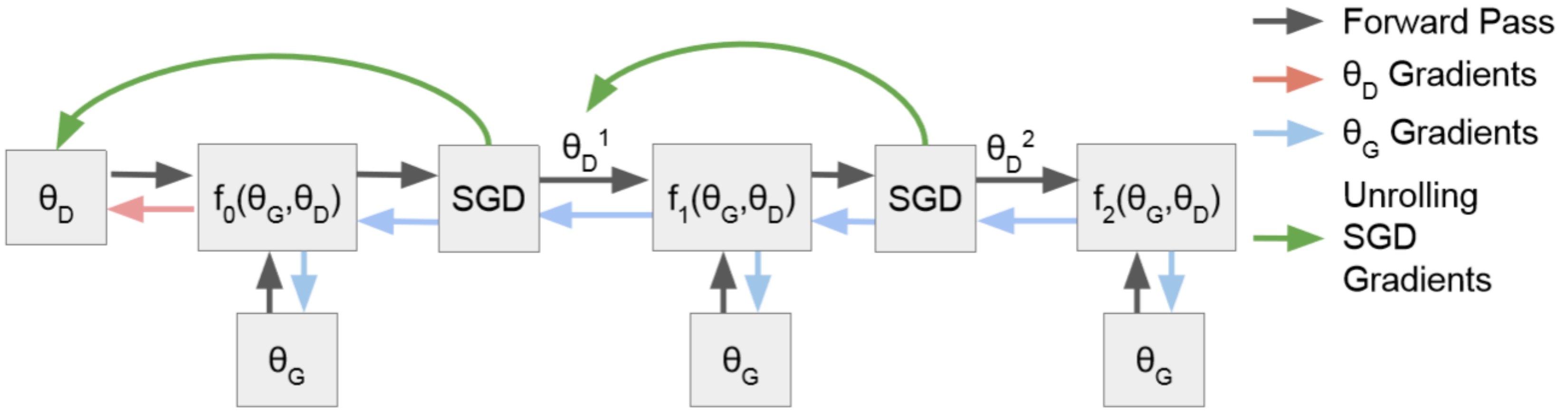
Not convex

→ impossible

we will introduce a surrogate objective function fK (θG,θD)

Surrogate objective funtion

$$f_K(\theta_G, \theta_D) = f(\theta_G, \theta_D^K(\theta_G, \theta_D))$$



Local optimum θ_D^*

$$\theta_D^0 = \theta_D$$

$$\theta_D^{k+1} = \theta_D^k + \eta^k \frac{df(\theta_G, \theta_D^k)}{d\theta_D^k}$$

$$\theta_D^* = \lim_{k\to\infty} \theta_D^k$$

Update

$$\theta_G \leftarrow \theta_G - \eta \frac{df_K(\theta_G, \theta_D)}{d\theta_G}$$

$$\theta_D \leftarrow \theta_D + \eta \frac{df(\theta_G, \theta_D)}{d\theta_D}.$$

θ_D 업데이트에는 f_k를 적용하지 않는다

What is different from standard gan?

- Standard gan: several update steps of the discriminator parameters should be run before each single update step for the generator

$$\theta_G \leftarrow \theta_G - \boxed{\eta \frac{\partial f\left(\theta_G, \theta_D^K\left(\theta_G, \theta_D\right)\right)}{\partial \theta_G}}$$

- unrolled gan: several update steps of the discriminator parameters should be run before each single update step for the generator

$$\theta_G \leftarrow \theta_G - \boxed{\eta \frac{\mathrm{d} f(\theta_G, \theta_D^K(\theta_G, \theta_D))}{\mathrm{d}\theta_G}}$$

$$\boxed{\frac{\mathrm{d} f_K\left(\theta_G, \theta_D\right)}{\mathrm{d}\theta_G}} = \boxed{\frac{\partial f\left(\theta_G, \theta_D^K\left(\theta_G, \theta_D\right)\right)}{\partial \theta_G}} + \frac{\partial f\left(\theta_G, \theta_D^K\left(\theta_G, \theta_D\right)\right)}{\partial \theta_D^K\left(\theta_G, \theta_D\right)} \frac{\mathrm{d}\theta_D^K\left(\theta_G, \theta_D\right)}{\mathrm{d}\theta_G}.$$

the second term reflects that as the generator collapses towards a delta function

$$\frac{\mathrm{d}f_K(\theta_G,\theta_D)}{\mathrm{d}\theta_G} = \frac{\partial f\left(\theta_G,\theta_D^K(\theta_G,\theta_D)\right)}{\partial\theta_G} + \frac{\partial f\left(\theta_G,\theta_D^K(\theta_G,\theta_D)\right)}{\partial\theta_D^K(\theta_G,\theta_D)}\frac{\mathrm{d}\theta_D^K(\theta_G,\theta_D)}{\mathrm{d}\theta_G}.$$

Delta function: An optimal generator for any fixed discriminator

When K = 0 and when K → ∞
: standard Gan loss = unrolled gan loss

K = 0

$$f_K(\theta_G,\theta_D) = f(\theta_G,\theta_D^K(\theta_G,\theta_D))$$

$$\theta_D^0 = \theta_D$$

$$f_0(\theta_G,\theta_D) = f(\theta_G,\theta_D^0) = f(\theta_G,\theta_D)$$

K = ∞

$$f_K(\theta_G,\theta_D) = f(\theta_G,\theta_D^K(\theta_G,\theta_D))$$

$$\theta_D^* = \lim_{k\to\infty}\theta_D^k$$

$$f_\infty(\theta_G,\theta_D) = f(\theta_G,\theta_D^\infty) = f(\theta_G,\theta_D^*)$$

K=1,2,3, ……

$f_k(\theta_G,\theta_D)$ captures additional information about the response of the discriminator to changes in the generator

$$\frac{\mathrm{d}f_K(\theta_G,\theta_D)}{\mathrm{d}\theta_G} = \frac{\partial f\left(\theta_G,\theta_D^K(\theta_G,\theta_D)\right)}{\partial\theta_G} + \frac{\partial f\left(\theta_G,\theta_D^K(\theta_G,\theta_D)\right)}{\partial\theta_D^K(\theta_G,\theta_D)}\frac{\mathrm{d}\theta_D^K(\theta_G,\theta_D)}{\mathrm{d}\theta_G}.$$