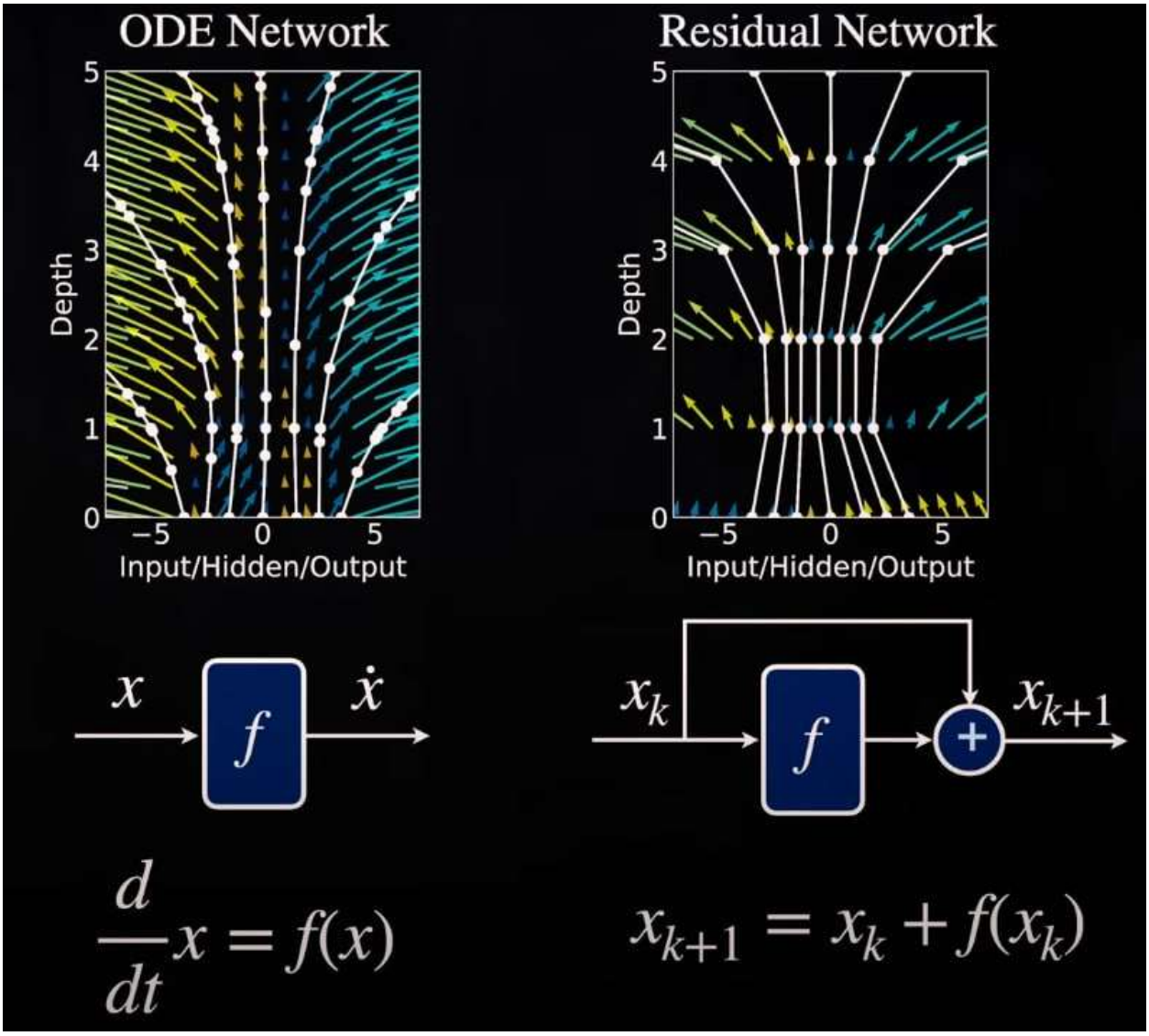
ODE Network
Residual Network
Depth
Input/Hidden/Output
0
1
2
3
4
5
−5
0
5
x
f
$\dot{x}$
x_k
x_{k+1}
+
$\frac{d}{dt}x = f(x)$
$x_{k+1} = x_k + f(x_k)$

$$\text{BN}(x) = \gamma \cdot \frac{x - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} + \beta$$

Where:

- μ_B: Mean of the batch for that channel
- σ_B^2: Variance of the batch
- γ, β: Learnable scale and shift (if `affine=True`)

During **inference**, it uses **running mean and variance** instead of batch stats.