$$D_{KL} = \Big(\underbrace{Q_\phi\left(\tfrac{z}{x}\right)}_{\text{known}} \,\Big\|\, \underbrace{P_\theta\left(\tfrac{z}{x}\right)}_{\text{not known}} \Big) =$$

$$= \int_z Q\,\phi\left(\tfrac{z}{x}\right) \log\left(\frac{Q\,\phi\left(z/x\right)}{P_\theta\left(\tfrac{z}{x}\right)}\right) dx$$

$$= \underbrace{\sum Q_\phi\left(\tfrac{z}{x}\right) \log\left(\frac{Q\,\phi\left(\tfrac{z}{x}\right)}{P_\theta\left(\tfrac{z}{x}\right)}\right)}_{\text{KL div. Formula}}$$

$$E_{z \sim Q\left(\frac{z}{x}\right)} \log\left(\frac{Q\,\phi\left(\tfrac{z}{x}\right)}{P_\theta\left(\tfrac{z}{x}\right)}\right)$$

$$E_{z \sim Q\left(\frac{z}{x}\right)} \left(\log Q_\phi\left(\tfrac{z}{x}\right) - \log P_\theta\left(\tfrac{z}{x}\right)\right)$$

$$E_{z\sim Q\left(\frac{z}{x}\right)}\left(\log Q_\phi\left(\frac{z}{x}\right) - \log\left(\frac{P_\theta\left(\frac{x}{z}\right)\cdot P(z)}{P(x)}\right)\right)$$

$$E_{z\sim Q\left(\frac{z}{x}\right)}\left(\log Q_\phi\left(\frac{z}{x}\right) - \left(\log P_\theta\left(\frac{x}{z}\right) + \log P_\theta(z)\right) - \log P_\theta(x)\right)$$

$$E_{z\sim Q\left(\frac{z}{x}\right)}\left(\log Q_\phi\left(\frac{z}{x}\right) - \log P_\theta\left(\frac{x}{z}\right) - \log P_\theta(z) + \log P_\theta(x)\right)$$

move out to left side

$$D_{KL}\left(Q_\phi\left(\frac{z}{x}\right) \,\|\, P_\theta\left(\frac{z}{x}\right)\right) - \log P_\theta(x)$$

$$= E_z\left(\log Q_\phi\left(\frac{z}{x}\right) - \log P_\theta\left(\frac{x}{z}\right)\right.$$

$$- \log P_\theta(z))$$

~~Add -ve sign both side~~

$$\log P_\theta(x) - D_{KL}\left(Q_\phi\left(\frac{z}{x}\right) \| P_\theta\left(\frac{z}{x}\right)\right)$$

$$= E_z\left(\log P_\theta\left(\frac{x}{z}\right)\right) - E_z\left(\log Q_\phi\left(\frac{z}{x}\right) - \log P_\theta(z)\right)$$

$$\log P_\theta(x) - D_{KL}\left(Q_\phi\left(\frac{z}{x}\right) \| P_\theta\left(\frac{z}{x}\right)\right)$$

$$-E_{Q_\phi\left(\frac{z}{x}\right)}\left(\log p_\theta\left(\frac{x}{z}\right)\right) - D_{KL}\left(\frac{Q_\phi\left(\frac{z}{x}\right)}{p_\theta(z)}\right)$$

$L(\theta, \phi)$

Reconstruction term

L

VAE objective

KL div. loss

regularization