Only reconstruction loss        Only KL divergence        Combination

The circular shape in the combined-loss case comes from the KL-Divergence aligning the latent space with a spherical Gaussian prior, since it has no directional bias, treating all directions equally, much like points on a sphere. The KL-Divergence penalizes deviations from this symmetry, pulling latent points toward the center (zero mean) and keeping the variance even in every direction that evenly distributes points around the origin.

The reconstruction loss makes sure that similar data points (like images of the same class) are mapped close together in the latent space. A regular autoencoder may form separate clusters and gaps, but the KL-Divergence acts as a constraint, forcing these clusters to stay within the boundaries set by the Gaussian prior.

Visually this produces a densely packed circle: the KL-Divergence creates the circular envelope by enforcing symmetry, while the reconstruction loss arranges meaningful groups within that envelope.