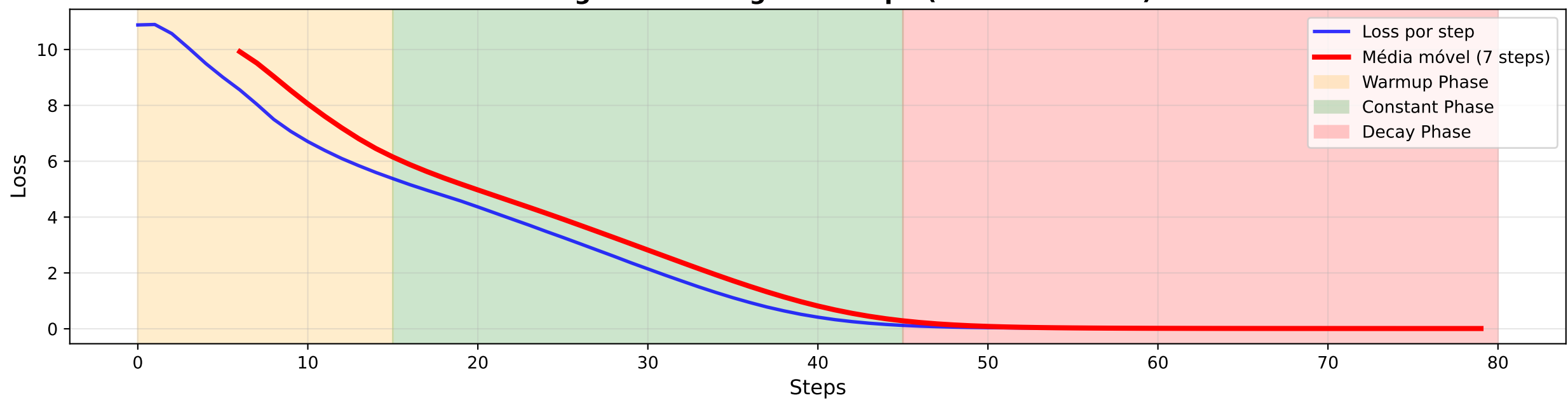
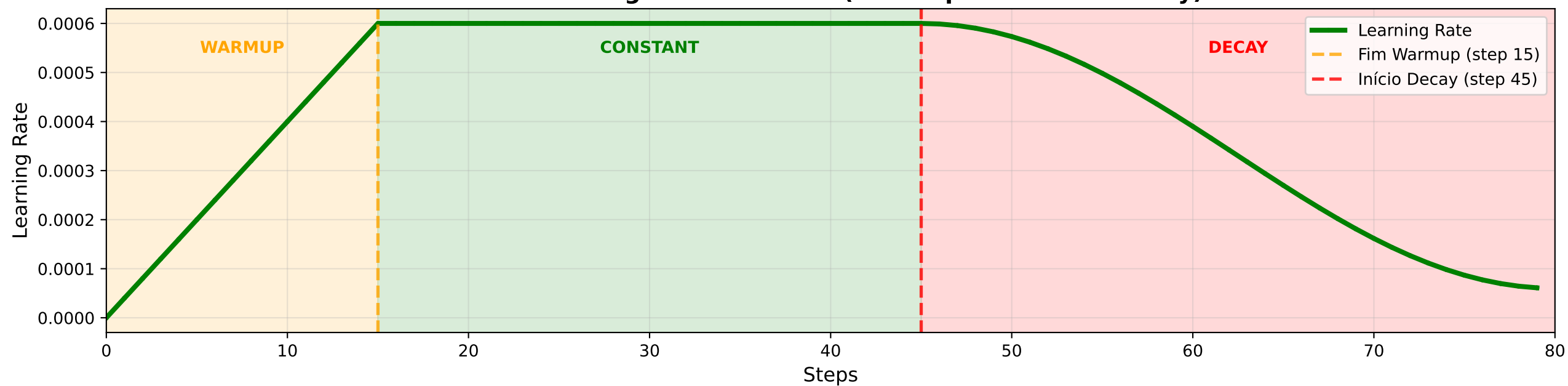


Training Loss ao longo do tempo (com fases WSD)



WSD Learning Rate Schedule (Warmup → Stable → Decay)



Loss vs Learning Rate (colorido por step)

