Training Loss: Numerical Stability

Legend:
- Ours (Stable)
- Episode 920
- Stability threshold
- Stable region

Loss remains stable (7.08)