| **§** | **Heading** | **What goes in** | **Status** |
| --- | --- | --- | --- |
| 3.1 | Bridge from review | “From delta-fragility to minimax RL” – 1 crisp paragraph | ✍️ next message |
| 3.2 | Market & hedge dynamics | Discrete-time price, option, P&L recursion | queued |
| 3.3 | Latent regime model | Two-state HMM, reference matrix \bar P | queued |
| 3.4 | Uncertainty set 𝒫 | ε-rectangle, convexity lemma | queued |
| 3.5 | Risk measure | CVaR definition + why coherent | queued |
| 3.6 | Minimax objective | \min\_\pi\max\_{P∈𝒫}\text{CVaR}\_α\bigl(L(π,P)\bigr) | queued |
| 3.7 | Existence & properties | Sketch proof: stationary minimax π exists (Iyengar 2005) | queued |
| 3.8 | Algorithm choice | Why PPO (+ entropy) fits the theory | queued |

Section 3 will be the lay of the land. Above are the pieces of our puzzle.

**From delta-fragility to minimax RL**

**Classical delta-hedging collapses** the continuous-time Black–Scholes ideal onto a single-regime volatility.  Section 2 showed empirical evidence — from the 2020-22 SPX vol-spikes — that mis-specifying σ inflates tail P&L.  Deep-hedging papers inject neural nets, yet still **train on one fitted model**.  Recent “robust RL” studies (e.g. Rajeswaran 2017; Merton-Heston PPO 2025) instead cast hedging as a **two-player game**: *the hedger* chooses positions, while *Nature* perturbs dynamics inside a statistically plausible set. We now formalise that game for latent **regime-switch volatility** and show that the resulting minimax problem admits a stationary optimal policy and an efficient PPO-style solver.

**Market & Hedge dynamics**

We first look to discretise the trading horizon