

Semi-Autonomous Formalization of the Vlasov–Maxwell–Landau Equilibrium

Clawristotle: Adventures in semi-autonomous formalization

Vasily Ilin — University of Washington

1 OVERVIEW

This project presents a complete Lean 4 formalization of the equilibrium characterization for the Vlasov–Maxwell–Landau (VML) system. The formalization closes the full AI-assisted research loop: Gemini DeepThink generated the proof idea, Claude Code translated it into Lean, Aristotle proved many lemmas, and the Lean kernel verified the final theorem. A single mathematician supervised the process over 10 days and *wrote zero lines of Lean code by hand*.

2 MAIN THEOREM

For smooth positive steady states of the VML system on $T^3 \times \mathbb{R}^3$ with Coulomb collisions:

- $f(x, v)$ is a spatially uniform Maxwellian
- $E(x) = 0$
- $B(x) = B_0$ is constant

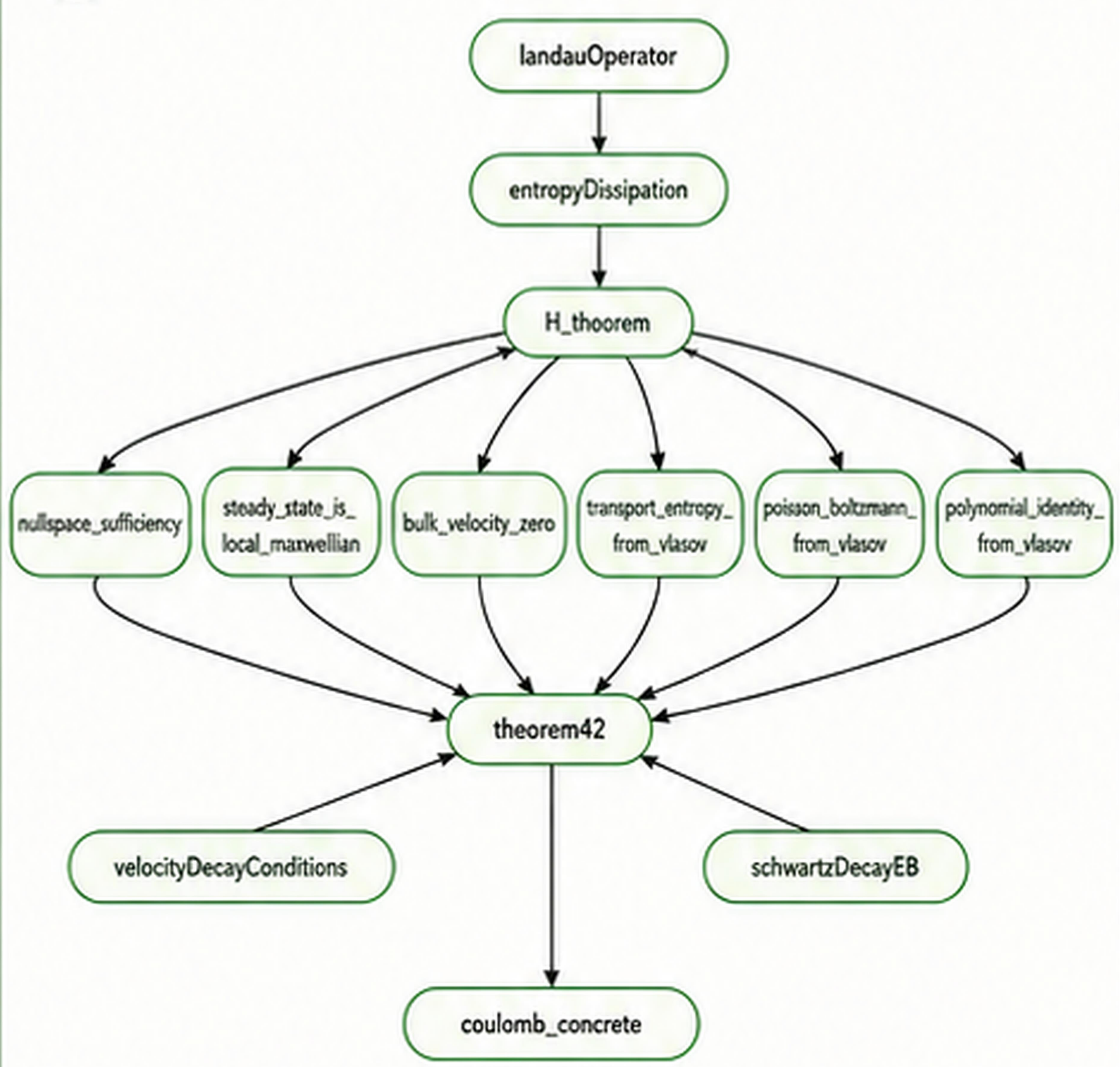
$$f(x, v) = \frac{\rho_{ion}}{(2\pi T)^{3/2}} \exp\left(-\frac{|v|^2}{2T}\right), \quad T > 0$$

where $\rho_{ion} > 0$ is the fixed ion density.

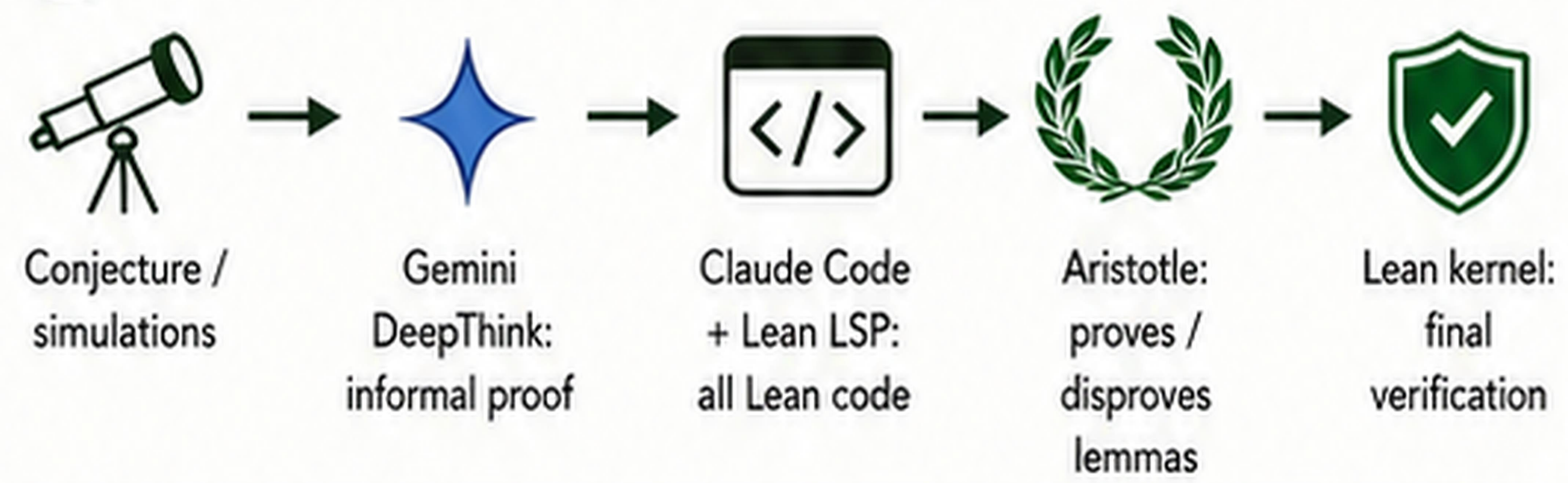
Lean theorem excerpt

```
theorem CoulombConcreteTheorem42 ... :
  ∃ (T_eq : ℝ) (B_0 : Fin 3 → ℝ),
  0 < T_eq ∧
  (∀ x v, f x v =
    equilibriumMaxwellian ρ_ion T_eq v) ∧
  (∀ x, E x = 0) ∧
  (∀ x, B x = B_0)
```

PROOF ARCHITECTURE



AI-ASSISTED WORKFLOW



Late Feb 2026: Gemini one-shots the proof • Mar 1–9: formalization sprint

PROJECT STATISTICS

	34 Lean 4 files		229 human prompts	~50 hours active supervision
	10,445 lines of code		220 Aristotle submissions	
	39 theorems		111 proved (50%)	
	186 lemmas		28 disproved (13%)	
	28 definitions		66 returned with sorry (30%)	
	0 sorry's		15 failed (7%)	0 lines of Lean written by hand
	10 days		2.8 billion input tokens	
	213 git commits		11 million output tokens	
			\$200 actual cost	

LESSONS LEARNED

- ✓ **WHAT WORKED**
- ✓ Lean LSP tools + Aristotle made Claude much stronger
- ✓ Start manually, automate only as needed
- ✓ Gemini was useful as a mathematical consultant
- ✓ Manual inspection of the main theorem was essential
- ✓ Research-level formalization is becoming accessible



- ⚠ **FAILURE MODES**
- ✗ Hypothesis creep
- ✗ Definition-alignment bugs
- ✗ Context management is hard
- ✗ Code quality can be poor
- ✗ LLMs can be overconfident or misleading
- ✗ Human supervision is still critical

Takeaway: AI can now help close the loop from conjecture to machine-verified proof — but theorem integrity still depends on expert human review.