War Room: An AI-assisted mathematical review workflow

Illustrated by a Navier–Stokes case study

Version v1.0.0 – October 14, 2026 (EDT)

Matthew J. Goss Jr.

Abstract

We present War Room, an AI-assisted workflow for organizing mathematical review and inspectable evidence. An 80-round campaign examining OpenAI's *Finite Time Blowup for Navier–Stokes* was followed by source-grounded reconstruction, isolated AI challenges and executable scalar QA. The retained evidence supports two conditional arguments: pressure comparison and exact reserved-base preservation. The report explains the apparatus, acceptance criteria and bounded conclusions.

1. Audience, purpose and questions answered

Intended audience. This report addresses mathematicians evaluating AI-assisted proof review, analysts examining the Navier–Stokes arguments, and researchers considering MathLab/War Room.

What we offer. We provide a checkable verification record: source locations, assumptions, reconstructed reasoning, challenge reports and scalar checks. Readers can examine the apparatus and judge the conclusions from the evidence.

What we found. Our calculations support both selected arguments in the source. The follow-up answers three concrete questions:

1. **Does pressure comparison require an additional pressure-growth assumption?** No additional assumption is needed in the inspected argument: a smooth competitor with the same force and zero datum, uniformly bounded in spatial L^2 on fixed $[0, T]$, $T < 1$, agrees with the smooth compact reference solution.
2. **Why does the reserved mean base survive exactly?** Support ordering and zero axial moment force the primitive and both cutoff terms to vanish on the reserved interval, preserving the fixed power-law base under the source's preparation conditions.
3. **What supports these judgments beyond AI agreement?** Derivations, source locators, six QA questions, hypothesis-removal checks, challenge reasoning and ten exact arithmetic checks support the judgments [E1–E3].

Relationship to the original proof. The source already contains these arguments. We have not established that these questions were previously unanswered or that our derivations are mathematically novel. Our contribution is an organized, inspectable examination of two delicate steps and of the evidence supporting their verification.

2. The apparatus and follow-up procedure

2.1 Campaign and focused source-grounded follow-up

The project used MathLab/War Room to organize an 80-round automated review against a frozen internal TeX master, with analyst, critic, verifier and synthesizer role labels. The present report concerns the focused mathematical follow-up of two inspectable dependencies. Pressure comparison was selected for its pressure ambiguity and nonlocal flux estimates; reserved-base preservation was selected for its primitive, support and cutoff dependencies. This is purposive selection.

The follow-up used ChatGPT/Codex Work contexts for retrieval, analytical reconstruction and challenge review. The original 166-page PDF governed interpretation. An internal TeX reconstruction assisted retrieval; the source pair was fixed by SHA-256. Original-PDF page and equation locators link the comparison proof to pp.121–123 and the mean-base chain to the preparation, coefficient and inverse passages listed in [E1].

The evidence procedure had five steps:

1. **Identify the source and target.** Fix source identities and locate the claim, proof and supporting passages across sections and appendices.
2. **State the premises.** Retain the solution class, fixed-time bounds, support ordering, moment equations and fixed-parameter conditions in a claim ledger.
3. **Reconstruct the reasoning.** Derive pressure recovery, flux bounds and closure; trace the primitive and both cutoff terms to exact mean-patch preservation.
4. **Challenge the implications.** Use separate AI contexts to attack analytical steps and remove selected hypotheses to test their necessity.
5. **Reconcile and retain evidence.** Judge objections against source passages and calculations, record the resulting bounded conclusions, and retain executable arithmetic checks.

2.2 Source-first panel and coordination

Following separate challenges of the reconstructions, three fresh reviewer contexts received the frozen source and neutral, role-specific questions. They did not receive our draft, supporting derivations, preferred verdicts or one another's reports. Pressure and mean reviewers could inspect the original PDF and frozen ingestion TeX. The adversarial reviewer began with 29 identified original-PDF pages and consulted the full PDF to resolve definitions and update formulas. The coordinator read the reports after their first reviews.

The three tasks were pressure/source reconstruction, mean/source reconstruction and an adversarial six-question review. The mean reviewer also followed the preparation chain through Appendix C. Reconciliation retained support ordering, the prepared comparison state and the distinction between the fixed base and the later updated velocity. The supplied reports show the reasoning behind those decisions; role agreement alone was not the acceptance criterion.

3. Mathematical QA: what was tested

The follow-up organized six questions, P1–P3 and M1–M3. Each has a source locator, analytical test, retained premise and resulting conclusion in [E2].

P1 – Pressure-gradient recovery. Derive time-averaged H^{-3} control from the conservative momentum equation. Check that the harmonic difference has Fourier support at the origin and weighted L^2 regularity, hence vanishes. The calculation supports recovery without an extra pressure-growth premise.

P2 – Cutoff flux and legitimate testing. Reconstruct the L^1-to-$L^{4/3}$ commutator estimate and the fixed-radius spatial approximation. Check local integrability and compact space-time testing. Constant pressure offsets cancel because the integral of the compact test field's divergence is zero; the cutoff test itself is not divergence-free.

P3 – Energy closure. Track weighted interpolation, Young absorption and radius-independent constants. The relevant radius remainders have powers 4, 4/3 and 14/5. The reference solution supplies the bounded gradient coefficient for Gronwall; no global competitor derivative bound is introduced. The result is $E_R \le C_T/R$ for each fixed $T < 1$.

M1 – Primitive cancellation. Convert the solved axial moment using $dX = R\,dR$, giving $\int U_n\,dX = 0$. Check that all relevant supports precede the reserved mean interval uniformly in the parameter. Thus $F_n = 0$, its parameter derivatives vanish and the radial coefficient V_n vanishes there.

M2 – Cutoff and realized sum. Differentiate the cutoff streamfunction before taking the velocity. Retain both χV_n and $\chi' F_n$ in the radial bracket. Their exact zeros, together with $U_n = E_n = 0$, annihilate every positive-order velocity term on the interval. Local finiteness for $q > 0$ passes this equality to the summed base.

M3 – Prepared state and normalization. Follow heat preparation, axis joining and cone moment restoration to the already prepared input. Distinguish local velocity from cumulative pressure/stress data. Verify the $X = x^2/2$ coefficient conversion and its positive bound for fixed $\lambda > 0$. Later updates are checked against their own full-defect formulas; the power law belongs to the fixed base.

Hypothesis-removal checks

Two explicit examples test why the premises matter. A spatially constant velocity $v = c(t)$ with pressure $P = -c'(t) \cdot x$, zero force and $c(0) = 0$ permits a nonzero harmonic pressure gradient; any nonzero $c(t)$ violates whole-space L^2. For the mean base, equal-integral bumps on opposite sides of the interval give zero total moment but a nonzero primitive within it. This violates the source's support ordering. Both checks distinguish the actual claims from weaker statements that would be false.

Ten exact rational checks [E3] cover the weighted L^3/L^4 interpolation, radial kernel powers, commutator radius scaling and the three Young remainders. They check scalar arithmetic; the analytic implications are supported by the displayed derivations and review reasoning.

4. The two mathematical calculations

4.1 Pressure comparison under bounded energy

Lemma 10.5 (original pp.121–123) compares the constructed compactly supported solution u with a smooth competing solution v of the same equation and force, starting from zero, with a uniform spatial L^2 bound on $[0,T]$, $T<1$. No spatial pressure-growth or global derivative bound is imposed on v.

Set $w=v-u$, $\pi=P-p$ and $g=v\otimes v-u\otimes u$, where P and p are the competing and reference pressures. The energy bounds give $g\in L^\infty_t L^1_x$. Define the Riesz pressure $\pi_*=\sum_{i,j}R_iR_jg_{ij}$. A bounded Fourier multiplier on L^1 data places π_* in H^{-s} for $s>3/2$, without asserting Riesz boundedness on L^1.

Time averaging the conservative difference equation against a compact time test places $\int a(\nabla\pi-\nabla\pi_*)\,dt$ in H^{-3}. This distribution is harmonic. Its Fourier transform is a weighted L^2 function supported at the origin, hence zero. This recovers the pressure gradient without adding an unstated growth condition.

The cutoff commutator has kernel bounded by $|z|^{-3}\min\{|z|/R,1\}$ and $L^{4/3}$ norm $O(R^{-3/4})$. It therefore accepts L^1 stress input. Weighted Sobolev and interpolation reduce boundary flux to dissipation powers $3/2$, $1/2$ and $3/4$. Young absorption leaves radius decay at least R^{-1}. With ϕ_R the spatial cutoff, set $E_R=\int\phi_R^8|w|^2$. Gronwall then gives $E_R\le C_T/R$, implying $w=0$ on every fixed ball as $R\to\infty$.

The pressure-gradient control follows from the time-averaged equation; the competing solution's uniform spatial L^2 bound is essential. Detailed estimates and ten rational exponent checks appear in the analytical supplement. A cutoff energy test $\chi_R w$ is not divergence-free: spatially constant pressure offsets contribute zero because the integral of its divergence is zero for a compactly supported test field.

4.2 Exact preservation of the reserved mean base

The mean-correction inverse requires an exact power-law slow base on a reserved interval, not merely a small asymptotic error. We trace this requirement from the four reserved intervals after (A.9), through profile preparation, positive-order coefficient corrections, and summation cutoffs, to (8.24).

The third and fourth intervals retain $U=0$ and $E=c_{\mathrm{patch}}(1+\eta^2)^{-1}X^{-1/2-\lambda}$. Heat compensation edits the second interval and distant tail. It can change cumulative pressure and stress data on the later intervals even though their local velocity profiles stay exact. Cone modulation restores five moments to the already prepared input before the later reserved intervals; it need not restore every datum to the unprepared reference.

At positive coefficient order, U_n, E_n are supported to the left of the mean interval. Their supports lie in fixed radial regions, uniformly in η. The zero total axial moment makes the primitive $F_n=\int_0^X U_n\,dx$ vanish beyond those supports, and then $V_n=0$ there as well. This is essential: local vanishing of U_n alone would leave a possible nonzero primitive and a cutoff-generated radial velocity proportional to $\chi' F_n$. Equations (5.18) and (5.45) remove that term exactly.

With $X=x^2/2$, the resulting base coefficient is $a(\eta)=2^{1/2+\lambda}c_{\mathrm{patch}}(1+\eta^2)^{-1}$, bounded below by $2^{\lambda-1/2}c_{\mathrm{patch}}>0$. This supplies the distinct power-moment blocks of Lemma 8.7 for fixed $\lambda>0$. This preservation implication uses the upstream constructions and moment solves specified above. The claim concerns the fixed base, not the full state after later perturbations. The preparation chain, primitive calculation, cutoff derivative and normalization are provided in the analytical supplement.

5. Evidence, reproduction and interpretation

The public package provides four connected kinds of evidence:

- **Analytical details [E1]:** exact source locations, pressure estimates, the preparation ledger and the mean-base derivation.
- **Mathematical QA ledger [E2]:** P1–P3/M1–M3 mapped to premises, tests, conclusions and the retained review excerpts.
- **Review reasoning [E2]:** separate reconstruction challenges and relevant excerpts from the source-first pressure, mean and adversarial reports. An extraction map records the retained originals' hashes and the excerpt boundaries.
- **Executable checks [E3]:** ten rational scalar checks, their results and a verifier for the package's file identities. Publication QA also checks rendering, notation and correspondence between the report and supplement.

A reader can begin at either mathematical finding, locate its source passages in the original PDF, follow the reconstructed calculation, inspect the challenges and run the scalar script. The public excerpts permit scrutiny of the selected follow-up; they do not reproduce the complete historical campaign or proprietary application.

Value to the reader and scope of the claim

Beyond reading the original proof, a reader receives a linked verification record: premises mapped to source passages, reconstructed estimates, attempted counterexamples and challenge responses. This makes the selected reasoning easier to inspect and contest. The case documents use of MathLab/War Room and the follow-up apparatus. It has not demonstrated an unmet mathematical need or measured an accuracy advantage over other review approaches; those would require separate evidence.

Scope and review status. The two implications use the source's asserted smooth compact reference construction, upstream profile/coefficient and moment solves, and stated support and parameter conditions. Constants are for fixed $T < 1$ and fixed positive λ. The complete blowup construction and accompanying formalization were not independently certified. Completed challenge contexts used one configured model family; context isolation is documented, and external human peer review is not claimed.

Authorship, AI use and interests

Matthew J. Goss Jr. commissioned and monitored the project and is associated with MathLab/War Room. ChatGPT/Codex assisted retrieval, reconstruction, challenge review, scripts and drafting. This product interest and AI assistance are disclosed. OpenAI authored the examined source; it does not author or endorse this report.

References and accompanying evidence

[S1] OpenAI. *Finite Time Blowup for Navier–Stokes*. Original PDF, 166 pages. Source identities are recorded in [E1].

[S2] OpenAI. *NavierStokesAndEuler*. Accompanying repository.

[E1] `ANALYTICAL_DETAILS.md`. [E2] `MATHEMATICAL_QA_LEDGER.md`, `REVIEW_EVIDENCE_MAP.json` and the five review excerpts. [E3] `check_comparison_exponents.py`, `COMPARISON_EXPONENT_RESULTS.json`, `verify_release.py` and `MANIFEST.json`.