Documenting Possible Gaps, Counterexamples, and Formalization Deviations

2026-04-29

1 Purpose

When a formalization relies on a cited mathematical source, the cited source should remain intact. It is a source, not an errata file. If a possible gap, missing hypothesis, incorrect numerical constant, or other mathematical discrepancy is found, the matter should be recorded in a separate LaTeX note. In this repository such notes live under `docs/paper-gaps/`. The same principle applies when the formal development has already diverged from the cited source, for instance by changing a statement so that it compiles, or by using a different mathematical route because an available formal-library theorem has a slightly different form. Such a note is a mathematical note first, and a record of the formalization second: it should be readable by a mathematician or mathematical physicist who has not followed the issue discussion.

The note should isolate the mathematical mechanism. It should say what the cited source asserts, what the cited results actually imply, where a term or hypothesis is lost, and what statement the formal development proves instead. If the formal proof takes a different route, the note should explain the replacement argument and the hypotheses on which it rests. This is especially important when the proof uses a theorem, construction, or convention from Mathlib which is not part of the cited argument. In that case the note should state the Mathlib result in ordinary mathematical language, explain how it is specialized to the notation of the paper, and compare its hypotheses and conclusion with the step it replaces. It should not merely say that Lean did not accept a proof, that the modified proof compiles, or that a named library lemma was applied.

The documents in this directory should themselves follow this standard. A policy or template which reads like a checklist will tend to produce check-

list notes; the examples and templates should therefore be written as mathematical prose.

2 Concision and Readability

A paper-gap note should be short enough that a reader can understand the issue in one sitting. The first page should answer four questions: what the source uses, how hard the remaining formal work is, what the formalization or audit checked, and what the current verdict is. Begin with a short *At a glance* section when the note concerns an open or recently resolved formalization task. That section should estimate the difficulty, the rough implementation weight, the split between Mathlib reuse and project-local work, and the key Mathlib APIs or modules. Use honest ranges, not false precision.

The *Key theorem forms* section should be Section 2, immediately after *At a glance*. Put the theorem forms before the source details so the reader knows what the formalization is aiming to prove. After that, keep the source assertion, the mechanism being checked, the blueprint/formal comparison, and the conclusion. Use more sections only when a counterexample or long derivation genuinely needs room.

Keep background to the minimum needed for the local point. Prefer one clear special case over a general restatement of the surrounding theorem. Use displayed equations for the source assertion, the key calculation, and the theorem forms; avoid displays for traceability data, declaration names, or long lists. The blueprint/formal comparison should be compact and should not repeat the theorem forms. The section "Form of the Note" below specifies where traceability data belongs; the body should read as mathematical prose, not as a log.

If the verdict is that there is no mathematical discrepancy, say so plainly and stop. Keep the conclusion highly compact, normally under twenty source lines. If the only remaining work is an implementation task, name it briefly and link the issue or pull request in a footnote.

3 When the Investigation Starts

The investigation starts when the obstruction appears to be mathematical rather than merely formal. No final judgement is needed at this stage. It is enough that, after reading the relevant source and attempting the corresponding formal statement, the proof seems to require a different scalar, an

additional hypothesis, a different tensor placement, a distinction between two objects which the prose identifies, or an auxiliary assertion not supplied by the cited argument.

The same investigation should also start when such a deviation is discovered after the fact. If the existing formal development or informal blueprint has already changed the statement, strengthened a hypothesis, weakened a conclusion, inserted an unadvertised lemma, or replaced the cited proof by a different formal-library argument, the mathematical change should be documented even if all relevant files already compile.

At this early stage the purpose is to preserve the question while it is still local. The note may later conclude that the matter is only notational, that an intermediate estimate in the cited source needs a harmless correction, that a missing hypothesis is genuinely needed, or that no discrepancy remains. Thus the trigger is not the severity of the eventual verdict, but the appearance of a real mismatch between the cited reasoning and the statement that can be formalized.

A failed tactic search, a missing formal lemma, or an inconvenient coercion is not enough. The obstruction must survive translation back into ordinary mathematics.

Sometimes an investigation uncovers an improvement that is not a discrepancy: the formalization works correctly, but it could be simplified by using an existing library, replacing a custom type, or adopting a cleaner proof route. These are not, by themselves, paper-gap notes. They are action items. When such an improvement is identified, open a GitHub issue or pull request that describes the proposed change. If the investigation also produces a paper-gap note, link the issue or pull request from that note: the note records the mathematical justification, and the issue or pull request drives the implementation.

4 When to Write a Note

A note should be written once the source argument and the formal proof no longer match as mathematical arguments. This includes, for instance, a numerical estimate whose constant does not follow from the cited lemma, a displayed conclusion which uses an unstated hypothesis, or a step in which two objects are identified although they must remain distinct. In analytic, algebraic, probabilistic, or quantum arguments, such distinctions may concern the domain of a map, the placement of a tensor factor, the type of an outcome, the distinction between a structure and a substructure, or the

scalar carried by an approximation relation.

A note is required when an audit finds that a Lean declaration associated with a source-labelled statement has acquired an extra non-source hypothesis, such as bridge, residual, repair, producer, or obligation-structure data. Such a declaration is temporary proof debt, not a replacement for the source theorem, and should not be introduced merely to avoid a `sorry`. The note must say plainly that the Lean result is conditional, identify the source-faithful theorem form that remains to be proved, and give the intended obligation discharger or removal plan. The blueprint should not mark the source-labelled statement with `\leanok` by pointing to this conditional helper.

Changing a Lean theorem away from the statement in `references/ldt-paper/` is strongly discouraged. A paper-gap note should not make such a change look routine: it should explain the mathematical necessity, or else classify the change as proof debt to be removed.

When an existing conditional helper is audited, the first task is to extract any argument it already contains into source-faithful lemmas. If the remaining bridge hypothesis cannot be derived from the paper hypotheses, the formal development should restore the source theorem statement and record the missing proof obligation explicitly, rather than retaining the strengthened statement as though it were the paper theorem.

This requirement is not meant to classify every explicit Lean side condition as a discrepancy. Positivity, nonemptiness, decidability, field-model instances, and denominator nonvanishing hypotheses may be faithful encodings of the domain in which the paper's theorem is stated. The note should distinguish such boundary hypotheses from hypotheses that supply an unproved step of the paper argument.

A note is also appropriate when the proof becomes formal only after changing an intermediate scalar, weakening a conclusion, strengthening a hypothesis, or inserting an additional lemma which is not simply an expansion of the cited argument. These are mathematical changes, even when they are harmless for the final theorem. It is likewise appropriate when an available formal-library theorem proves the needed result under a somewhat different formulation from the cited source. In that case the note should be deliberately pedagogical. It should explain, in ordinary mathematical language, what the replacement result says, how it is specialized, which hypotheses are new or differently packaged, and why the resulting statement is a legitimate substitute for the step in the cited argument. If the issue remains open after reasonable attempts, the note should say what has been checked and what remains unknown.

5 Counterexamples

If the cited assertion appears to be false, the note should not stop at saying that the proof cannot be formalized. It should give the counterexample, or the smallest relevant part of it, in ordinary mathematical notation. The example should be accompanied by enough verification that the reader can see both sides: the hypotheses of the cited assertion hold, and its conclusion fails.

When feasible, the counterexample should also be formalized in Lean. This is especially valuable when the example is finite, algebraic, computational, or otherwise small enough that the failure can be checked without reproducing the whole surrounding theory. The note should then state the mathematical content of the checked counterexample and identify the corresponding Lean declaration in a footnote. The Lean file is not a substitute for the prose: the prose should still explain why the example refutes the cited assertion.

6 Revision of a Note

A paper-gap note may be provisional. It may begin as soon as the mathematical obstruction has been isolated well enough to state the question, and it may be revised as the calculation, the blueprint, or the corresponding Lean declarations become clearer. The first version need not know the final classification, but it should say plainly which mathematical point is known and which point remains unsettled.

Later revisions should not make the note into a chronological log. They should rewrite the document so that it remains a coherent account of the present mathematical understanding. If the verdict changes, the conclusion should state the current verdict; the linked issue or pull request can preserve the earlier discussion.

7 Form of the Note

Each note should stand on its own. It should give the date, cite the relevant source article, and provide enough traceability for the reader to locate the original passage and the project discussion. GitHub issues and pull requests may be linked with `\href`, and source files, line ranges, and labels should be recorded when they help the reader check the claim. These references

should serve the exposition rather than govern it: the main text should be organized around the mathematical assertion and its proof.

The note should introduce its notation before using it. Symbols for measurements, tensor factors, approximation relations, and error parameters should be named in prose. If an approximation relation is state-dependent, the state should be named, or its defining expression should be given when this can be done briefly. Only after the notation is in place should the note state the assertion in the cited source, using a bibliographic citation such as `\cite[Section~...]{...}` when the section or chapter is known.

Only the relevant hypotheses should be reproduced. If the question is a scalar substitution, a missing term, or a missing hypothesis, the note should not quote a long unrelated part of the proof. It should give the cited lemma or proposition in the special case actually used, then display the calculation or logical implication in which the discrepancy appears.

Displayed equations should be used with restraint. They are appropriate for the source estimate being checked, the cited estimate in the form actually used, and the calculation where a term or hypothesis appears or disappears. When a displayed equation is identified by a source label, the label should appear on the right of the display, while the surrounding prose refers to the mathematical assertion itself. The standard `amsmath` command `\tag` is the right mechanism for this; a separate local convention should be introduced only if the standard mechanism is genuinely insufficient. For a multi-line aligned display, tag the individual lines that carry the source information rather than forcing a long combined label below the equations. Short scalar definitions and elementary inequalities should remain inline when the sentence reads better that way. As in ordinary mathematical writing, a displayed formula is part of the surrounding sentence and should be punctuated accordingly.

Bibliographic entries used only by these notes should live in `docs/paper-gaps/references.bib`. A new note should ordinarily begin from `docs/paper-gaps/template.tex`, which records the common LaTeX preamble, the local bibliography, and a model for right-hand source labels in displayed equations.

Blueprint prose cites a note like any other source, with `\cite{gap:<slug>}`; the bibliography entry in `blueprint/src/references.bib` carries the published URL of the note, whether the note lives in this repository or in another one. Lean docstrings and comments use the repository path form `docs/paper-gaps/<slug>.tex`.

The note should then compare the cited source with the informal blueprint, if there is one, and with the formal statement. This comparison

should be written in ordinary mathematical prose. Formal declaration names and source paths should be supplied only to identify the exact formal statements, and normally belong in footnotes.

The conclusion should give a verdict. It may say that there is no discrepancy, that the matter is only notational, that an intermediate estimate needs a harmless correction, that a hypothesis is missing, that there is a theorem-level gap, or that the question remains open. The classification should reflect the mathematics, not the amount of formal work required.

8 Decision Principle

The formal development should not silently replace the cited assertion by a different one. If the formal statement differs from the cited source, the reason for the difference should be written down. Conversely, the note should not overstate the conclusion. A wrong intermediate scalar is not a fatal gap in the theorem if the corrected scalar is bounded by the error parameter appearing in the theorem statement.