

Global Normalization and Fermat's Last Theorem

(An Explanation for a Mathematician without Programming)

Abstract

This text provides a *non-programmer's* explanation of the *global normalization* approach to Fermat's Last Theorem (FLT): we introduce a single parameter $o > 1$ and state that *any* hypothetical counterexample to the equation $x^n + y^n = z^n$ for $n > 2$ *implies* the equality

$$o^n = 2n,$$

after which, from the *principle of maximum coverage*, it follows that the *only possible* choice is $o = 2$, and the equality $2^n = 2n$ is possible only for $n \in \{1, 2\}$. Thus, for $n > 2$, a contradiction arises, which is consistent with FLT. Importantly: the **normalization statement** (that every counterexample indeed implies $o^n = 2n$ for the same o) is accepted as a *hypothesis*; everything else follows elementarily. Formal analogues of definitions and lemmas are provided and machine-checked in Coq (with no need to know programming languages, we will provide an interpretation of all steps in plain mathematical language).¹

¹The formal definition `covers_with`, the “bridge” lemma, and the theorem on maximum coverage are in the module *GlobalNormalization*; see the file with the Coq code. See also the explanatory preprint with historical commentary and a step-by-step reconstruction.

Contents

1	Short Roadmap	3
2	Intuition Comprehensible without Programming	3
3	Formulation in Mathematical Language	4
3.1	Definition and Principle	4
3.2	Structure of the set $S(o)$	4
3.3	Elementary Estimate of Exponential Growth	4
4	Main Conditional Theorem (Logic of FLT Derivation)	5
5	Where the Normalization Premise Comes From	5
6	Bridge from Real Equalities to Natural Powers	5
7	Elements of the Classics: Parametrization, Parity, Binomial Symmetry	6
7.1	Parametrization $(z, x) = (m^n + p^n, m^n - p^n)$	6
7.2	Odd/Even Binomial Sums	6
8	What Exactly Is Formally Checked and What Is Not	6
9	Why the Choice $o = 2$ Is Unique and “Natural”	6
10	Mini-FAQ for the Reader	7
11	Detailed Appendices (Elementary Proofs)	7
11.1	Why $2^n > 2n$ for $n \geq 3$	7
11.2	Strict decrease of $(2n)^{1/n}$	7
11.3	Odd binomial indices in the difference	7
11.4	Parametrization and parity	7
12	How to Read Formal Names (Mini-Glossary without Programming)	8

1 Short Roadmap

1. **The Idea of Normalization.** We fix a number $o > 1$. We say that “ o covers the exponent n ” if $o^n = 2n$ holds. This idea is conveniently packaged as follows: *every* counterexample for $n > 2$ *implies* the coverage of n by the same o .
2. **Maximum Coverage.** Consider $S(o) = \{n \in \mathbb{N} : o^n = 2n\}$. We will show that for $o = 2$, the set $S(o)$ consists of exactly $\{1, 2\}$, and for $o \neq 2$, it is smaller in cardinality. Hence, the *only* choice compatible with the principle of “maximum scope” is $o = 2$.
3. **The Elementary Reason for the Contradiction.** For $o = 2$, we have $2^n = 2n$ only for $n \in \{1, 2\}$, because for $n \geq 3$, $2^n > 2n$. Consequently, a hypothetical counterexample for $n > 2$ cannot exist.
4. **What Remains a Hypothesis.** We do *not* prove that for any counterexample, $o^n = 2n$ is inevitable with the *same* $o > 1$. This is the *normalization premise*. Everything else is rigorous (and even formally verified) mathematics.

2 Intuition Comprehensible without Programming

The goal is to *conditionally* (assuming the normalizing premise) reduce FLT to a simple growth estimate. The idea of normalization arose naturally from the classical transformation

$$(z, x) = (m^n + p^n, m^n - p^n),$$

which is convenient for parity analysis and binomial expansions: for an odd n , the difference $(m^n + p^n)^n - (m^n - p^n)^n$ contains only odd binomial indices and is a multiple of n .²

Two things are important here:

- **Combinatorics of the Binomial Theorem.** When subtracting $(A+B)^n - (A-B)^n$, only terms with odd powers of B remain, and there is a common factor of n —this allows *normalizing* the difference to the form $2 \cdot n \cdot (\text{something})$.
- **Growth of Exponential vs. Linear Function.** The equality $o^n = 2n$ cannot hold for large n if $o \geq 2$; and if $1 < o < 2$, it can be satisfied *at most once* due to the monotonicity of $(2n)^{1/n}$ with respect to n . This leads to the idea of choosing the o where the coverage is maximal—and it turns out to be exactly $o = 2$, covering only $n = 1$ and $n = 2$.

²In Coq, this is reflected through elementary facts about the parity of sums/differences, see lemmas on parity and divisibility of binomial differences.

3 Formulation in Mathematical Language

3.1 Definition and Principle

Definition 3.1 (Coverage). For $o > 1$ and $n \in \mathbb{N}$, we say that o covers n if

$$o^n = 2n.$$

We denote the set of covered exponents by $S(o) = \{n \in \mathbb{N} : o^n = 2n\}$.

Definition 3.2 (Global Normalization, Principle of Maximum Coverage). We say that *global normalization holds* if there exists an $o > 1$ such that **any** hypothetical counterexample $x^n + y^n = z^n$ for $n > 2$ implies the coverage of n by the same o , i.e., $o^n = 2n$ (the same o for all n). The principle of *maximum coverage* dictates choosing the $o > 1$ for which the cardinality $|S(o)|$ is maximal among all permissible o .

3.2 Structure of the set $S(o)$

Consider the function $f(n) = (2n)^{1/n}$ for $n \geq 1$. Then $o^n = 2n$ is equivalent to $o = f(n)$.

Lemma 3.3 (Monotonicity of f). For $n \geq 2$, the function $f(n)$ is strictly decreasing, $f(1) = 2$, $f(2) = 2$, while for $n > 2$ we have $f(n) < 2$ and $\lim_{n \rightarrow \infty} f(n) = 1$.

Proof (sketch). Consider $\ln f(n) = \frac{\ln(2n)}{n}$ and its derivative with respect to n in the continuous relaxation: $\frac{d}{dn}(\ln f(n)) = \frac{1 - \ln(2n)}{n^2} < 0$ for $n \geq 2$. The calculations $f(1) = f(2) = 2$ are trivial, and the limit to 1 is standard (use $\lim_{n \rightarrow \infty} \frac{\ln n}{n} = 0$). \square

Proposition 3.4 (Description of $S(o)$). For a fixed $o > 1$:

- 1) if $o \geq 3$, then $S(o) = \emptyset$ (since $o^n \geq 3^n > 2n$);
- 2) if $1 < o < 2$, then $|S(o)| \leq 1$ (by Lemma 3.3, the level o is crossed at most once for $n \geq 2$, and $f(1) = 2 > o$);
- 3) if $o = 2$, then $S(2) = \{1, 2\}$, since $f(1) = f(2) = 2$, and for $n > 2$, $f(n) < 2$.

Consequently, the maximum cardinality $|S(o)|$ is attained uniquely at $o = 2$ and is equal to 2.

Corollary 3.5 (Choice of Normalization). From the principle of maximum coverage, it follows that $o = 2$ and $S(o) = \{1, 2\}$.

3.3 Elementary Estimate of Exponential Growth

Lemma 3.6. For every $n \geq 3$, it is true that $2^n > 2n$.

Proof. By induction on n , or by comparison with binomial coefficients: $2^n = \sum_{k=0}^n \binom{n}{k} \geq \binom{n}{0} + \binom{n}{1} + \binom{n}{n-1} + \binom{n}{n} = 2 + 2n > 2n$ for $n \geq 3$. \square

Corollary 3.7. The equality $2^n = 2n$ is possible only for $n \in \{1, 2\}$.

4 Main Conditional Theorem (Logic of FLT Derivation)

Theorem 4.1 (FLT from Global Normalization). *Suppose there exists an $o > 1$ such that for **any** hypothetical natural triple x, y, z with $n > 2$ satisfying $x^n + y^n = z^n$, the coverage $o^n = 2n$ holds (with the same o for all n). Then the equation $x^n + y^n = z^n$ has no solutions in \mathbb{N} for $n > 2$.*

Proof. By the principle of maximum coverage and Proposition 3.4, we have $o = 2$ and $S(o) = \{1, 2\}$. Then any required coverage for $n > 2$ is impossible, since by Corollary 3.7 the equality $2^n = 2n$ does not hold. This is a contradiction. \square

5 Where the Normalization Premise Comes From

Substantively, the premise comes from the *binomial analysis* of the difference

$$(m^n + p^n)^n - (m^n - p^n)^n,$$

where for an odd n , only odd binomial indices survive, and the entire sum is a multiple of n ; moreover, in constructions of the form

$$\left(\frac{z+x}{2}\right)^{1/n}, \quad \left(\frac{z-x}{2}\right)^{1/n}$$

for different natures of n (even/odd), a common factor of n appears and, roughly speaking, a “universal” radical that can be collapsed into the same norm o for all n .³

It is important to understand: **this transition is left as a hypothesis**. The preprint states directly: normalization is *postulated* and used as an *explicit hypothesis*, rather than being derived from arithmetic by the author in a complete form.⁴

6 Bridge from Real Equalities to Natural Powers

The key “bridge” fact: if for the same o , two exponents n and m are covered,

$$o^n = 2n, \quad o^m = 2m,$$

then, by raising to powers and eliminating o , we obtain an equality of integer powers

$$(2n)^m = (2m)^n.$$

This is convenient when one needs to transfer the reasoning from \mathbb{R} back to \mathbb{N} and apply the arithmetic of powers.

Remark 6.1 (Why this is needed). The bridge lemma helps to vary the points of view: normalization lives in \mathbb{R} , while $x^n + y^n = z^n$ itself lives in \mathbb{N} . The equality $(2n)^m = (2m)^n$ is a pure integer symmetry, compatible with the idea that the set of covered exponents is extremely limited.

³In the explanatory text, this is formulated as steps 2–6 with the extraction of the common factor n and the transition to an equality of the form (some expression)ⁿ = 2n; see the section *Possible Proof*.

⁴A paraphrase of the quote: “The normalization premise is kept as an explicit hypothesis over \mathbb{N} ; bridge lemmas connect the real equation $o^n = 2 \cdot n$ with the integer comparison.”

7 Elements of the Classics: Parametrization, Parity, Binomial Symmetry

7.1 Parametrization $(z, x) = (m^n + p^n, m^n - p^n)$

This notation by itself does not guarantee that m, p are integers; however, it provides strong *necessary* conditions on the parities of $z \pm x$ and on the form of the expressions $\frac{z \pm x}{2}$. In particular,

$$z + x = 2m^n, \quad z - x = 2p^n,$$

so that $z \pm x$ are even, and if we *additionally* require $m, p \in \mathbb{Z}$, then both halves must be perfect n -th powers. And vice versa: if these conditions are met, then m, p can be reconstructed (with caveats about signs for even n).

7.2 Odd/Even Binomial Sums

The sums of even and odd binomial coefficients are equal and give 2^{n-1} :

$$\sum_j \binom{n}{2j} = \sum_i \binom{n}{2i+1} = 2^{n-1},$$

which is convenient in checks for $n = 1, 2$ and in expansions of $(A+B)^n \pm (A-B)^n$.

8 What Exactly Is Formally Checked and What Is Not

- **Formally in Coq:** the definition of coverage $o^n = 2n$; the theorem on maximum coverage (from $o^1 = 2$ it follows that $o = 2$ and only $n \in \{1, 2\}$ are covered); the *bridge lemma*; elementary lemmas about the growth of 2^n relative to $2n$; *sanity-goals* like “from $o = 2$, coverage of $n = 3$ does not follow”.⁵
- **Left as a premise:** *global normalization* (“any counterexample \Rightarrow coverage by the same o ”). This is formulated and emphasized in the preprint as an explicit hypothesis; it is motivated by binomial calculations and its “universal” form, but is *not* fully derived within classical elementary number theory.

9 Why the Choice $o = 2$ Is Unique and “Natural”

Uniqueness is a consequence of the strict decrease of $f(n) = (2n)^{1/n}$ for $n \geq 2$: only the level $o = 2$ passes through two points $n = 1, 2$; any other o yields at most one intersection point (or none at all). This is the content of Proposition 3.4. In terms of *method*: the principle of maximum coverage suggests a “natural norm” that coincides with the base of the binary exponential.

⁵Sanity checks at the end of the file serve as regression tests: for $o=2$, coverage for $n \geq 3$ is false, etc.

10 Mini-FAQ for the Reader

- **Weak spot?** The only one is the *normalization premise*. All other steps are elementary and/or formally verified.
- **Doesn't this contradict Wiles?** No. This is *not* a new independent proof of FLT, but a *reduction* of FLT to a specific normalization statement. Wiles's proof remains the only unconditional one to date.
- **Why use Coq?** So that the minimal algebraic details (growth, the bridge lemma, “only $n = 1, 2$ for $o = 2$ ”) are machine-checked, eliminating human carelessness.
- **Is there a p -adic perspective?** The code contains a sketch of a “two-adic bracket” for a universal parameter and the fact that $v_p(o)$ vanishes for odd p ; this is illustrative and not used in the main growth antagonism.

11 Detailed Appendices (Elementary Proofs)

11.1 Why $2^n > 2n$ for $n \geq 3$

See Lemma 3.6. The proof can also be done by induction: in the step $n \rightarrow n+1$, from $2^n > 2n$ it follows that $2^{n+1} = 2 \cdot 2^n > 4n \geq 2(n+1)$ for $n \geq 2$.

11.2 Strict decrease of $(2n)^{1/n}$

See Lemma 3.3. Another option is to use the AM–GM inequality:

$$(2n)^{1/n} \leq \frac{2 + \overbrace{1 + \dots + 1}^{n-1 \text{ times}}}{n} \xrightarrow{n \rightarrow \infty} 1.$$

And for strict decrease for $n \geq 3$, one can compare $f(n)$ and $f(n+1)$ directly.

11.3 Odd binomial indices in the difference

The difference $(A+B)^n - (A-B)^n$ eliminates all even indices, leaving

$$2 \sum_j \binom{n}{2j+1} A^{n-(2j+1)} B^{2j+1},$$

which yields the common factor $2 \cdot B$ and, in appropriate substitutions, the factor n (via $\binom{n}{1} = n$ and further multiples).

11.4 Parametrization and parity

If $z = m^n + p^n$, $x = m^n - p^n$, then $z \pm x = 2 \cdot (\text{perfect power})$, hence $z \pm x$ are even. From this follow convenient *negative* criteria: if $z \pm x$ have the wrong parity, a corresponding parametrization with integers m, p does not exist.

12 How to Read Formal Names (Mini-Glossary without Programming)

- `covers_with o n` means $o^n = 2 \cdot n$.
- `covers_with_two_characterisation` — from $2^n = 2n$ it follows that $n \in \{1, 2\}$.
- `maximum_coverage_as_theorem` — formalization of the principle of maximum coverage: from coverage of $n = 1$, it is derived that $o = 2$ and the restriction $n \in \{1, 2\}$.
- `two_real_normalizations_imply_nat_power_eq` — the bridge: from $o^n = 2n$ and $o^m = 2m$ it follows that $(2n)^m = (2m)^n$.
- `sanity goals` — small automatic checks like “for $o=2$ there is no coverage for $n \geq 3$ ”.

Conclusion (in the spirit of Fermat)

If one believes that any hypothetical triple for $n > 2$ forces the *same* number o to satisfy $o^n = 2n$, then from the principle of maximum coverage, we immediately obtain $o = 2$ and thus exclude all $n > 2$, since $2^n > 2n$. The conciseness of the result resonates with the idea of a “short marginal note”. But it is precisely the *normalization* that is the subject of the main question and further research.

What to do next? Refine and justify the normalization premise (e.g., through Diophantine estimates, p -adic arguments, or comparison with programs like ABC/descent), while preserving the elementary aesthetic line of reasoning.