



## Uncertainty, Rule-Bending, and the Near-Miss Three Lessons from the Dialectics of Mathematical Practice

Bharath Sriraman 

Department of Mathematical Sciences, University of Montana, Missoula, MT 59812, USA  
[sriramanb@mso.umt.edu](mailto:sriramanb@mso.umt.edu)

**Received:** 29 May 2026; **Accepted:** 25 June 2026

### Abstract

This paper develops a framework for thinking about mathematical creativity that takes *uncertainty* as a structural condition of mathematical work, *rule-bending* as a methodological move that is more pervasive than usually acknowledged, and *near-misses*, in the sense of Marvin Minsky, as the cognitive unit through which mathematical learning advances. Two case studies – the origami construction of  $\sqrt[3]{2}$ , and the zeta-regularised assignment  $1 + 2 + 3 + \dots = -1/12$  – are used to illustrate the rule-bending move at very different scales of the mathematical enterprise. The framework is then applied to the pedagogy of mathematics, with a particular focus on the *Pathology Principle*: that the objects historically labelled "pathological" by mathematicians have very often been near-misses in the precise Minsky sense, and that their pedagogical value lies in exactly this fact. The paper concludes with concrete pedagogical implications for the teaching of mathematics at the secondary and early-tertiary level.

**Keywords:** Mathematical Creativity, Near-miss, Divergent Series, Origami Geometry, Cognitive Frames, Mathematics Education, M. Minsky, B. Lonergan, Pathological Objects

**MSC2020:** 97D20, 40G05, 51M15, 00A30

### 1. Introduction: the dichotomy at the heart of the practice

There is a dichotomy at the heart of mathematics that every practitioner has felt and very few textbooks acknowledge. Mathematics is presented to students as a sequence of definitions, theorems, and proofs – a clean, certain, deductive edifice in which each step follows from the previous one and the whole thing locks together into something monumental. This is the front of the restaurant. The back of the restaurant looks nothing like this. In the back, mathematicians work with conjectures they cannot prove, with objects they cannot quite define, with arguments that almost go through, with intuitions they cannot yet articulate. The back is messy; the back is uncertain; the back is the place where the mathematics actually gets made.

This paper takes the dichotomy seriously. It argues that uncertainty, far from being an unfortunate temporary state that mathematicians should hurry to dispel, is the structural condition of mathematical work. The conjectures we cannot close are the conjectures that drive the field. The axioms we tweak are the axioms that produce new mathematics. The objects we cannot quite fit into our existing frames are the objects that grow the frames. We will give the third of these phenomena a precise name – the *near-miss*, in the sense of Marvin Minsky – and we will argue that it is the cognitive engine of mathematical learning, both in the historical development of the field and in the classroom development of the student.

The paper has three parts, matching the three Acts of the companion plenary lecture (Sriraman, 2026).

**Part I** (Sections 2–3) treats uncertainty as catalyst. We distinguish four senses of *uncertainty* that arise in mathematical practice and use the Riemann Hypothesis and the Bieberbach/de Branges story to illustrate the human and social dimensions of working with mathematical doubt.

**Part II** (Sections 4–5) presents two case studies in rule-bending. The first is the origami construction of  $\sqrt[3]{2}$ , following Peter Messer (Messer, 1986), in which a single fold accomplishes what Greek geometry cannot. The second is the zeta-regularised identity  $1 + 2 + 3 + \dots = -1/12$ , which we develop from G.H.Hardy's classical treatment (Hardy, 1949) through Cesàro and Abel summation to the analytic continuation of the Riemann zeta function. Both cases exhibit the same structural move: a problem that is impossible in one rule set becomes elementary in a thoughtfully extended one.

**Part III** (Sections 6–8) introduces the near-miss framework, drawing on Minsky's *Society of Mind* (Minsky, 1986) and on Bernard Lonergan's *Insight* (Lonergan, 1957). We develop the Pathology Principle – the claim that historical pathologies in mathematics are very often near-misses in disguise – and use it to argue for a specific style of mathematics teaching in which near-misses are deliberately surfaced rather than smoothed away. We close with concrete pedagogical recommendations.

We have written the paper to be readable on its own, but it is intended as the scholarly companion to the lecture; the lecture develops the visual and narrative aspects of the same argument.

## 2. Four senses of uncertainty

The word uncertainty carries different cargo in different disciplines. In mathematics it is rarely thematised at all, and when it is, it is treated as a deficiency to be remedied rather than as a condition to be inhabited. We distinguish four senses in which uncertainty arises in mathematical work.

**(i) Doubt.** The first sense is the simplest. A mathematician working on a problem is in a state of doubt about whether the proof she has sketched is actually correct, or whether the conjecture is even true. Beghetto and others have called this creative doubt (Beghetto & Kaufman, 2007), distinguishing it from the merely paralytic doubt that prevents action. Creative doubt is productive precisely because the mathematician continues to work inside it.

**(ii) Ambiguity.** The second sense is structural. An object or a problem may be ambiguous – it may admit multiple interpretations, none of which has yet been ruled out by the available evidence. Ambiguity invites curiosity, in the sense that Aristotle and the eighteenth-century moralists meant the term: a sustained desire to know. Curiosity, in this sense, is contextual: it lives in the practices of a field rather than in the disposition of an individual.

**(iii) Desire.** The third sense is psychological. To work on a hard problem is to desire its solution – to want, in a way that is not separable from intellectual engagement, the moment at which the picture becomes clear. This desire shapes which problems mathematicians take up and which they leave alone.

**(iv) Dichotomy.** The fourth sense is the one this paper takes most seriously. There is a dichotomy between how mathematics is created and how it is presented. The presentation is deductive, certain, and clean. The creation is exploratory, uncertain, and messy. Most mathematics education treats students exclusively to the presentation, leaving them to discover the existence of the creation only if they go on to do research.

The dichotomy is not accidental. Mathematicians have strong reasons to present their work in finished form – the standards of the field require it – and the historical record has been pruned accordingly. But the consequence is that students grow up with a picture of mathematics that is structurally false in an important way. Restoring uncertainty to the picture is not a matter of weakening the standards of mathematical rigour. It is a matter of being honest about how rigour is produced.

## 2.1. Uncertainty in research mathematics: three examples

To make the four senses concrete, we describe three episodes from working mathematics in which uncertainty is the central feature.

**Goldbach, Fermat, and Catalan.** The conjecture that every even integer greater than 2 is the sum of two primes, stated by Christian Goldbach to Leonhard Euler in 1742, remains open. Fermat's Last Theorem, conjectured in 1637, was finally closed by Andrew Wiles in 1995, a span of 358 years. Catalan's conjecture, that the only consecutive perfect powers are 8 and 9, was stated in 1844 and closed by Preda Mihailescu in 2002, a span of 158 years. Each of these problems has been the centre of an active research programme for most of its lifetime; each has produced false proofs, partial results, and entire subfields whose object was the problem.

**The Riemann Hypothesis.** The most consequential open problem in number theory. Riemann's 1859 memoir conjectured that all non-trivial zeros of the analytically continued zeta  $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$  lie on the vertical line  $\operatorname{Re} s = \frac{1}{2}$ . Entire books and entire classes of functions (the L-functions, modular forms, generalised zeta functions) are written under the tacit assumption that Riemann is true. The error term in the prime number theorem, known to the mathematical community to high precision, is correct only if Riemann is correct. The community has learned to live with this uncertainty – to write entire papers conditional on the conjecture, to make claims of the form "assuming RH, we have ..." – a way that would seem absurd outside the field.

**Bieberbach and de Branges.** A more human episode. The Bieberbach conjecture, formulated in 1916, asserted that for any univalent analytic function on the unit disk,  $f(z) = z + a_2z^2 + a_3z^3 + \dots$ , the coefficients satisfy  $|a_n| \leq n$ . The Koebe function  $k(z) = z/(1-z)^2$  attains equality. In 1984 Louis de Branges, a mathematician who had repeatedly announced wrong proofs of major conjectures in the preceding decades, announced a proof of the full Bieberbach conjecture. The community looked away. It took a group of Russian mathematicians at the Steklov Institute, working unofficially, to verify and clean up the argument. The proof turned out to be correct, and the conjecture was closed in 1985.

The interesting feature of the de Branges story is that the uncertainty was social as much as mathematical. There was no mathematical reason to dismiss the proof; there was a sociological reason. The community's resistance to credulity is, on the whole, a healthy feature – false proofs are common – but in this case the resistance very nearly buried a correct argument. The episode is worth remembering whenever one is tempted to think that mathematical knowledge is produced in a purely intellectual medium. It is not. It is produced in a social medium in which reputation, trust, and the institutional ear all play their roles.

## 2.2. Why this matters for teaching

We will return to the pedagogical implications throughout the paper, but the basic point is already visible. If uncertainty is the structural condition of mathematical work – if mathematicians themselves spend most of their time inside it – then a mathematics education that systematically hides this fact from students is educating them for a field that does not exist. The deductive front-of-house picture is part of mathematics, but only part.

## 3. Tweaking the axioms

The second of our three lessons is that mathematicians, in their back-of-house mode, regularly tweak the axioms of their working systems. This sounds dangerous – we associate axiom-tweaking with crackpottery – but it is in fact a methodological move that has produced a great deal of legitimate mathematics. The question to ask is not whether axioms should be tweaked, but which tweaks have proven productive.

### 3.1. A small example: group-like structures

Consider the standard axioms of a group  $G$ :

- (1) **Closure:** for all  $a, b \in G$ , the product  $a \cdot b \in G$ .
- (2) **Associativity:** for all  $a, b, c \in G$ ,  $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ .
- (3) **Identity:** there exists  $e \in G$  with  $a \cdot e = e \cdot a = a$  for all  $a \in G$ .
- (4) **Inverses:** for every  $a \in G$  there exists  $a^{-1} \in G$  with  $a \cdot a^{-1} = e$ .

Drop axiom (4), and we have a monoid. Drop (3) as well, and we have a semigroup. Drop (1) instead of (4), and we have a groupoid. Each of these is a legitimate, well-studied algebraic structure with its own theorems and applications. The deeper question, however, is the reverse one: when does dropping an axiom produce something useful, rather than something nonsensical?

### 3.2. Historical shifts

The history of mathematics is in part a history of axiom-tweaking that turned out, in retrospect, to be productive:

**Geometry.** The parallel postulate of Euclid was tweaked into non-Euclidean alternatives by Bolyai, Lobachevsky, and Riemann in the nineteenth century. The mathematics that followed – spherical geometry, hyperbolic geometry, Riemannian geometry – was not pathological; it became the geometric substrate of general relativity.

**Set theory.** The classical foundational picture, in which mathematical objects are built up from sets, was tweaked by Eilenberg, Mac Lane, and Grothendieck into category theory, in which objects are characterised by their morphisms rather than their elements. The resulting mathematics is now indispensable in algebraic topology, algebraic geometry, and logic.

**Analysis.** The  $\varepsilon$ – $\delta$  definition of limit, due to Cauchy and Weierstrass, was tweaked by Abraham Robinson in the 1960s into nonstandard analysis, in which infinitesimals are restored as legitimate mathematical objects (Robinson, 1966).

**The real line.** The standard real line treats 0.999... and 1 as the same number. Constructivists, intuitionists, and certain non-Archimedean systems treat them as distinct. Whole research programmes have been devoted to the consequences.

In every case the tweak was, at the time, controversial. In every case the mathematics it produced is now part of the canon. The moral is that productive axiom-tweaking is not eccentric – it is a standard, if under-discussed, methodological move.

## 4. Case study I: $\sqrt[3]{2}$ and the cube of Delos

We now turn to the first of our two extended case studies in rule-bending. The case is ancient, but the lesson it carries is not.

### 4.1. The Delian problem

In Greek antiquity, three constructive problems resisted all attempts at solution by compass and straightedge:

**Doubling the cube.** Given a cube of edge 1, construct a cube of volume 2 – i.e., construct a segment of length  $\sqrt[3]{2}$ .

**Trisecting the angle.** Given an arbitrary angle, divide it into three equal parts using only compass and straightedge.

**Squaring the circle.** Given a circle, construct a square of equal area.

All three are unsolvable with the Greek tool kit. The first was called the Delian problem because, according to legend, the oracle at Delos demanded that the Athenians double the size of Apollo's cubical altar in order to end a plague. The Athenians, the story goes, doubled the side, and the plague got worse.

The Greeks did not know that the three problems were unsolvable. They believed – reasonably – that the tools they had used to construct so many other lengths would, with sufficient cleverness, suffice for these as well. They were wrong, but they could not have known they were wrong without a piece of mathematics that did not exist for another two thousand years.

#### 4.2. Why $\sqrt[3]{2}$ is unreachable

The piece of mathematics in question is the theory of field extensions, and the relevant theorem is the following:

**Theorem 1 (Constructibility).** *A real number  $\alpha$  is constructible by compass and straightedge if and only if  $\alpha$  lies in some tower of fields  $\mathbb{Q} = F_0 \subset F_1 \subset F_2 \subset \dots \subset F_n$  in which each successive extension  $F_{i+1}/F_i$  has degree at most 2.*

The proof – which we omit, as it is in any textbook on Galois theory – rests on the observation that compass and straightedge can construct intersections of lines and circles, and that these intersections are always solutions to linear or quadratic equations over the field of previously constructed quantities.

Now,  $\sqrt[3]{2}$  is a root of the irreducible polynomial  $x^3 - 2 \in \mathbb{Q}[x]$ . The field  $\mathbb{Q}(\sqrt[3]{2})$  has degree 3 over  $\mathbb{Q}$ . By the tower law, 3 would have to divide a product of 2's if  $\sqrt[3]{2}$  were constructible – and 3 does not. Therefore  $\sqrt[3]{2}$  is not constructible by compass and straightedge.

We emphasise the form of this argument. The unsolvability of the Delian problem is not a fact about the problem alone; it is a fact about the problem paired with the Greek tool kit. The cube root of 2 exists, perfectly well, as a length on the real line. It can be approximated to arbitrary precision. It is the root of a polynomial with rational coefficients. It fails exactly one condition: it does not lie in a tower of quadratic extensions of  $\mathbb{Q}$ .

#### 4.3. Origami and Messer's construction

What happens if we change the tool kit? Origami – paper folding – admits, in addition to the operations of compass and straightedge, certain folds that cannot be reduced to them. The critical fold is what is sometimes called the Beloch fold: a fold that simultaneously places one point on one given line and another point on another given line. This operation is cubic in nature, in the sense that the algebra it produces involves the solution of a cubic equation.

Once the Beloch fold is admitted,  $\sqrt[3]{2}$  becomes constructible. [Peter Messer \(1986\)](#) gave a particularly elegant construction. Begin with a square sheet of paper, vertices  $A, B, C, D$  (clockwise from upper-left). Trisect the square horizontally with two creases, giving points  $F$  and  $H$  on the left edge ( $AB$ ) and  $G$  and  $I$  on the right edge ( $DC$ ). Now make a single fold that simultaneously sends  $A$  onto edge  $DC$  and  $F$  onto the lower trisection line. Let  $P$  be the image of  $A$  on  $DC$ , and let the fold cross the edge  $AB$  at a point that divides  $AB$  in the ratio

$$\frac{PB}{AP} = \sqrt[3]{2}.$$

One fold. One sheet of paper. No compass, no straightedge. The length the Greeks could not construct in two thousand years is produced by a child folding a piece of paper.

#### 4.4. What the example teaches

The cube root example teaches two things.

The first is a meta-mathematical lesson about what unsolvability means. The phrase "impossible to construct" needs a parenthetical: impossible with what? The cube root of 2 is impossible to construct with compass and straightedge. It is elementary to construct with origami. It is also elementary to construct with a marked ruler (a neusis), as the Greeks themselves discovered. The unsolvability is in the tool, not in the length.

The second lesson is more general. Mathematical creativity often begins by asking which rule, if dropped or added, would let the problem move. The Greeks could not have found Messer's fold,

because they could not have asked the question that produces it. The question is: which operation, not in our current tool kit, would render this length accessible? Once the question is asked, the answer follows quickly.

We will see exactly the same structure in our second case study, at a very different scale of the mathematical enterprise.

## 5. Case study II: divergent series and zeta regularisation

The second case study comes from real and complex analysis. It is the story of how mathematicians learned to assign finite values to series whose partial sums do not converge – a story whose central result is the identity

$$1 + 2 + 3 + 4 + 5 + \dots = -\frac{1}{12} \quad (1)$$

which, properly interpreted, is a theorem.

### 5.1. Cauchy's definition and the eighteenth-century practice

The modern notion of a convergent series is due to [Cauchy \(1821\)](#). Given a sequence  $(a_n)$ , the partial sums are  $s_n = a_0 + a_1 + \dots + a_n$ . The series  $\sum a_n$  is said to converge to  $s$  if  $s_n \rightarrow s$  as  $n \rightarrow \infty$ . Anything else, in Cauchy's terms, is divergent, and for Cauchy a divergent series is meaningless. He was emphatic on this point.

But the practice of mathematics in the eighteenth century had not respected Cauchy's strictures, because Cauchy had not yet imposed them. Euler, in particular, had computed extensively with divergent series and had obtained correct answers ([Hardy, 1949](#)). Euler wrote, in his characteristic way:

*Let us say, then, that the sum of any infinite series is the finite expression by the expansion of which the series is generated. In this sense the sum of the infinite series  $1 - x + x^2 - x^3 + \dots$  will be  $\frac{1}{1+x}$ , because the series arises from the expansion of the fraction, whatever the value of  $x$  may be.*

In particular, by setting  $x = 1$ , Euler obtained  $1 - 1 + 1 - 1 + \dots = \frac{1}{2}$  – Grandi's series – a result that, by Cauchy's criterion, is meaningless.

The eighteenth-century question, then, is whether Euler was right. Cauchy's nineteenth-century answer was: no. The twentieth century's answer was: yes, but in a sense Euler did not have the language to state.

### 5.2. Cauchy's definition and the eighteenth-century practice

The way the twentieth century rehabilitated Euler is by recognising that the words "sum of an infinite series" admit more than one meaning, and that several distinct meanings yield consistent mathematics. Following the formulation in Hardy's *Divergent Series* ([Hardy, 1949](#)), suppose we want to define a function  $\Sigma$  that assigns to certain series  $\sum a_n$  a value  $\Sigma(a)$ . The minimum desiderata are:

- (A) **Linearity.** If  $\Sigma(a) = s$ , then  $\Sigma(ka) = ks$  for every scalar  $k$ .
- (B) **Additivity.** If  $\Sigma(a) = s$  and  $\Sigma(b) = t$ , then  $\Sigma(a + b) = s + t$ .
- (C) **Stability.**  $\Sigma(a_0, a_1, a_2, \dots) = s$  if and only if  $\Sigma(a_1, a_2, \dots) = s - a_0$ .

The Cauchy notion of convergence satisfies (A), (B), and (C). But it is not the only function that does. *Cesàro summation*, which averages the partial sums, satisfies them. *Abel summation*, which forms the power series  $\sum a_n x^n$  and takes the limit as  $x \rightarrow 1^-$ , satisfies them. The *zeta-regularised* sum, which we will define below, satisfies them where it applies. Each of these is a generalised sum – a tightening of the function  $\Sigma$  to a larger class of series than Cauchy convergence allows.

### 5.3. Grandi's series

Consider Grandi's series  $1 - 1 + 1 - 1 + \dots$ . Three lines of algebra, all of which use only (A) and (C), suffice to compute its forced value. Let  $S$  be the value, whatever it is:

$$\begin{aligned} S &= 1 - 1 + 1 - 1 + \dots && \text{(by definition)} \\ &= 1 - (1 - 1 + 1 - \dots) && \text{(by (C), splitting off the first term)} \\ &= 1 - S && \text{(by (A), factoring out } -1) \\ \therefore 2S &= 1 \Rightarrow S = 1/2. \end{aligned}$$

This is exactly Euler's answer. The argument shows that any summation method satisfying (A) and (C) must, if it assigns Grandi's series a value at all, assign it  $\frac{1}{2}$ . The Cauchy partial sums oscillate between 0 and 1 and never reach  $\frac{1}{2}$ , but the value  $\frac{1}{2}$  is forced by the axiomatic structure that Cauchy convergence is only one instance of.

### 5.4. Abel summation

The same three axioms also force the answer for  $1 - 2 + 3 - 4 + \dots$  to be  $\frac{1}{4}$ . The cleanest way to see this is Abel's:

$$\sum a_n := \lim_{x \rightarrow 1^-} \sum_{n=0}^{\infty} a_n x^n.$$

For Grandi,  $\sum (-1)^n x^n = \frac{1}{1+x}$ , whence  $\lim_{x \rightarrow 1^-} \frac{1}{1+x} = \frac{1}{2}$ . For alternating integers, differentiate the geometric series to get  $\sum (-1)^n (n+1) x^n = \frac{1}{(1+x)^2}$ , whence  $\lim_{x \rightarrow 1^-} \frac{1}{(1+x)^2} = \frac{1}{4}$ .

### 5.5. Ramanujan's letter

On 16 January 1913, Srinivasa Ramanujan – then a 25-year-old clerk at the Madras Port Trust, with no formal university training – wrote to G. H. Hardy at Trinity College, Cambridge, enclosing 120 mathematical identities. Buried among them was the assertion

$$1 + 2 + 3 + \dots = -\frac{1}{12}, \quad \text{"under my theory."}$$

Ramanujan added, with the candour for which the letter is now famous: "*If I tell you this you will at once point out to me the lunatic asylum as my goal*" (Berndt & Rankin, 1995).

Hardy did not. The Hardy–Littlewood–Polya school went on to make Ramanujan's theory of divergent series precise.

### 5.6. The Ramanujan derivation, with a caveat

A standard heuristic derivation, which is not quite a proof but is illuminating, runs as follows. Let

$$S = 1 + 2 + 3 + 4 + \dots$$

be the mystery sum. We have already established (Abel) that  $S' = 1 - 2 + 3 - 4 + \dots = \frac{1}{4}$ .

Subtract term by term:

$$\begin{aligned} S - S' &= (1 - 1) + (2 - (-2)) + (3 - 3) + (4 - (-4)) + \dots \\ &= 0 + 4 + 0 + 8 + 0 + 12 + \dots \\ &= 4 \cdot (1 + 2 + 3 + \dots) = 4S. \end{aligned}$$

Therefore  $-3S = S' = \frac{1}{4}$ , whence  $S = -\frac{1}{12}$ .

The caveat. This derivation is not a proof. It assumes that  $S$  is a finite number that can be added to and subtracted from  $S'$ , but  $S$ 's status as a finite number is precisely what is in question. No summation method satisfying (A), (B), (C) and *total regularity* – the requirement that convergent series should be summed to their Cauchy limits and that series with partial sums tending to  $+\infty$

should not be assigned finite values – can assign  $1 + 2 + 3 + \dots$  any finite value. The heuristic derivation works by quietly relaxing the regularity requirement.

### 5.7. Zeta regularisation: making the answer rigorous

The rigorous justification of  $-\frac{1}{12}$  comes from a different direction. Define, initially for  $\text{Re } s > 1$ , the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

By Euler's calculation,  $\zeta(2) = \frac{\pi^2}{6}$ ,  $\zeta(4) = \frac{\pi^4}{90}$ , and so on. The series converges only for  $\text{Re } s > 1$ , but Riemann showed in his 1859 memoir that  $\zeta$  admits a meromorphic continuation to the whole complex plane, with a single pole at  $s = 1$ . The continuation satisfies the functional equation

$$\zeta(1-s) = 2 \cdot (2\pi)^{-s} \cdot \cos\left(\frac{s\pi}{2}\right) \cdot \Gamma(s) \cdot \zeta(s).$$

Setting  $s = 2$ ,

$$\zeta(-1) = 2 \cdot (2\pi)^{-2} \cdot \cos(\pi) \cdot \Gamma(2) \cdot \zeta(2) = \left(\frac{1}{2\pi^2}\right) \cdot (-1) \cdot 1 \cdot \left(\frac{\pi^2}{6}\right) = -\frac{1}{12}.$$

Thus, in the precise sense that  $\zeta(-1)$  is the value at  $s = -1$  of the analytic continuation of  $\sum n^{-s}$ , we have  $\zeta(-1) = -\frac{1}{12}$ . This is a theorem.

When physicists and number theorists write  $1 + 2 + 3 + \dots = -\frac{1}{12}$ , what they mean is exactly this: the divergent series is being interpreted, by convention, as the value of the analytically continued zeta function at  $s = -1$ . The value is rigorous; the notation is a shorthand.

## 6. The physical reality of $-\frac{1}{12}$

There is one more thing to say about  $-\frac{1}{12}$  before we turn to the pedagogical part of the paper. It is measurable in the laboratory.

### 6.1. The Casimir effect

Casimir and Polder, in 1948, predicted that two parallel uncharged conducting plates in vacuum should attract each other with a force per unit area

$$F = -\frac{\pi^2 \hbar c}{240 a^4},$$

where  $a$  is the separation. The minus sign indicates attraction. The constant 240 in the denominator is  $2 \cdot 120 = 2 \cdot 5!$ , and it traces, in the derivation, back to the regularised sum  $\zeta(-3) = \frac{1}{120}$ .

The Casimir force was measured to high precision starting in the late 1990s (Lamoreaux, 1997). Its existence is now well-established experimentally. A particular finite value of an analytically continued zeta function appears in the constant of proportionality between the plate separation and the measured force. The divergent-series identity is not a mathematical curiosity; it is a prediction of physical reality.

### 6.2. The critical dimension of bosonic string theory

A second example, more speculative but mathematically clean. Bosonic string theory's mass spectrum involves a divergent zero-point sum

$$\frac{1}{2}(1 + 2 + 3 + \dots).$$

Replacing the divergent sum by its zeta-regularised value gives  $\frac{1}{2} \cdot (-\frac{1}{12}) = -\frac{1}{24}$ . The consistency conditions of the theory then impose  $(D - 2)/24 = 1$ , whence  $D = 26$ . Spacetime in bosonic string theory has 26 dimensions because  $1 + 2 + 3 + \dots = -\frac{1}{12}$ .

Whether or not bosonic string theory describes physical reality is a separate question. The internal mathematical structure of the theory rests on the zeta-regularised value of a divergent series.

### 6.3. What the second case study teaches

The second case study teaches the same structural lesson as the first. A problem that is impossible in one rule set – the Cauchy notion of summation – becomes a theorem in a thoughtfully extended one – the zeta-regularised notion. Just as the cube root of 2 is forbidden by Greek geometry and permitted by origami, the value  $-\frac{1}{12}$  for  $1 + 2 + 3 + \dots$  is forbidden by Cauchy summation and permitted by analytic continuation. In both cases, the new value is reachable only by trusting a problem the old rule set declares impossible.

This is what we mean by *rule-bending*. It is the second of our three lessons. The third lesson tells us what is happening cognitively when a mathematician makes such a move.

## 7. The near-miss: Minsky and the child's arch

We come now to the central concept of the paper. It originates not in mathematics but in cognitive science. It was introduced by Marvin Minsky in *The Society of Mind* (Minsky, 1986), building on earlier work of Patrick Winston on machine learning of structural concepts (Winston, 1970). The example Minsky chose to illustrate it is famously simple, and we will spend a little time making sure the example lands.

### 7.1. The arch the child sees and the arch the adult sees

Minsky's example concerns a child learning the concept "arch" from physical blocks. He writes, in paraphrase: *To a child playing with blocks, an arch is not just an arch. It is one block on top of two blocks that do not touch. A pile of three blocks that does not satisfy that second condition is a near-miss – and the near-miss is what forces the definition.*

This claim is initially puzzling, and the puzzlement is worth sitting with for a moment.

If one asks an adult what an arch is, the adult will think of something like a Roman arch, the entrance to a church, the architectural element that supports weight across an opening. In ordinary speech, an arch is just any structure with a top supported by sides. The two sides may well *touch each other at the bottom* – the keystone arches of Roman aqueducts have stones that meet at their bases – and we would still call them arches.

So why does Minsky insist that the child sees the arch differently? Why does the child require, as part of the very definition, that the two supporting blocks not touch?

The answer is that the arch the child sees is defined by its function in the child's play. To a child playing with blocks, the arch is interesting precisely because something can pass under it. A doll can walk under the arch. A toy car can drive through it. The arch is the structure that creates a passable gap. If the two supporting blocks are pushed together until they touch, the gap closes; the doll can no longer walk through; the car cannot drive under. What was an arch becomes a wall, or a step, or a small table. The structure has not been destroyed in any obvious visual sense – there are still three blocks, still one on top of two – but the feature that made it an arch is gone.

The arch-ness of the arch lives entirely in the gap.

This is what Minsky means when he says that the child sees an arch "not just" as a configuration of blocks, but as a configuration in which the gap is a structural feature. The child's representation of "arch" has a slot for the gap. The adult's representation may or may not – most adults, on reflection, will report that they had not consciously noticed this – but the child's does, because the child's purposes in playing with the blocks make the gap the salient feature.

To make the distinction completely explicit, consider three block configurations in Fig. 1.

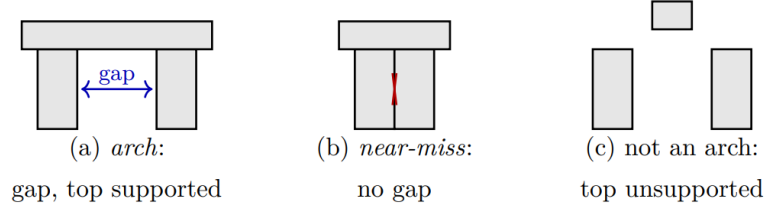


Fig. 1. Three configurations of three blocks. (a) is an arch: a gap separates the supports and a block bridges them. (b) is a *near-miss*: the bridging block is supported, but the gap has closed. (c) is a different failure altogether: the top block is not supported. The pedagogical value of (b) is that it differs from (a) in exactly one feature – the gap – and so makes that feature visible *as a feature*.

Configuration (a) is an arch. Configuration (c) is plainly not an arch – the top block is suspended in the air with nothing to hold it up; the failure is loud, and a child will reject it immediately. Configuration (b) is the interesting case. It looks like an arch: there are three blocks, one rests on the others, the bottom two stand upright. By every gross criterion it is an arch. But it fails the gap condition. A doll cannot walk under it. A toy car cannot drive through. The child rejects it, but the rejection takes work: the child has to articulate, perhaps for the first time, that the gap was a structural part of what she meant by arch all along.

It is precisely the structural similarity of (b) to (a) that gives (b) its cognitive force. If we wanted to teach a child the concept of arch by examples, configuration (c) – obviously not an arch – is useless. The child rejects (c) without learning anything she did not already know. Configuration (b) is the example that teaches. It teaches by being almost an arch.

This pedagogical asymmetry – near-misses teach where loud failures do not – is the cognitive observation on which Minsky's framework rests.

## 7.2. Slots, frames, and near-misses

We now state the framework formally. A frame is a mental representation of a concept consisting of a list of *slots* – features that must be present for an instance to count as falling under the concept. The frame "arch," in the child's playing, has at least the following slots:

- **Slot 1.** There are three blocks.
- **Slot 2.** One block lies on top of the other two.
- **Slot 3.** The two bottom blocks stand upright.
- **Slot 4.** The two bottom blocks are separated by a gap.

A *near-miss* to this frame is an object that satisfies all slots but one. A pile of three blocks in which the two bottom blocks touch is a near-miss, because it satisfies slots 1, 2, and 3 but fails slot 4.

The pedagogical observation, due to Minsky and Winston (1970), is that *the near-miss is what makes the failing slot visible*. A child who has only ever seen successful arches does not yet distinguish "arch with a gap" from "three blocks stacked" – the two are conflated because, in her experience, every arch has had a gap. The first near-miss – a stacking without a gap – forces her to notice the gap as a separable feature.

**Definition 1.** Let  $F = (P_1, P_2, \dots, P_k)$  be a frame consisting of  $k$  slots, where each  $P_i$  is a predicate. An object  $x$  is a near-miss to  $F$  if there exists exactly one index  $i$  such that  $P_i$  fails on  $x$ , and  $P_j$  holds on  $x$  for every  $j \neq i$ . The slot  $P_i$  is called the *failing slot*.

## 7.3. Lonergan: the act of insight

The near-miss tells us what is being represented. It does not yet tell us how a learner moves from one frame to a richer one. For that, we draw on a second author: Bernard Lonergan, the philosopher of cognition, whose magnum opus *Insight* (Lonergan, 1957) develops a phenomenology of the act of knowing.

Lonergan's central concept is the act of *insight*: a cognitive event in which a previously unnoticed feature of a situation becomes salient, and the situation is thereby restructured. The act of insight is not the gradual accumulation of evidence; it is a sudden gestalt shift. The Archimedean "Eureka" is a paradigm case.

The relevance to the near-miss framework is immediate. When a learner encounters a near-miss, the work that has to be done is cognitive: she must move from a representation in which the failing slot is invisible to one in which it is explicit. That movement is precisely an act of insight in Lonergan's sense. The slot, which was operating implicitly in every successful instance, becomes explicit by the act of being violated.

Minsky tells us *what* is being represented (a frame with slots, one of which fails). Lonergan tells us how the representation changes (by an act of insight that brings the failing slot to consciousness). The mathematical near-miss is the joint event.

## 8. The Pathology Principle

We come now to the main pedagogical claim of the paper. We have suggested that mathematical near-misses play, in the development of mathematical concepts, the role that the gapless block-pile plays in the development of the child's concept of arch. To make the analogy productive, we identify the mathematical near-misses with a class of objects that already has a name: the *pathological* ones.

### 8.1. Statement of the principle

**The Pathology Principle.** Mathematical objects that working mathematicians have historically labelled *pathological* – "monsters," in Lakatos's vocabulary (Lakatos, 1976), "a lamentable plague," in Hermite's (Hermite, 1893/1905) – are very often near-misses in the precise Minsky sense. The historical labelling of an object as pathological tracks closely with the cognitive role the object played in forcing the construction of a richer frame.

The principle is not an a priori claim; it is an empirical generalisation about the history of mathematics. We give three examples in support of it.

### 8.2. Grandi's series as a near-miss

The Cauchy frame for "summable series" has, as we have already laid out, four slots: linearity (A), additivity (B), stability (C), and a fourth slot we did not previously name explicitly – the *Cauchy partial-sum slot*: the partial sums converge to the sum.

Grandi's series  $1 - 1 + 1 - 1 + \dots$  passes the first three slots and fails the fourth. The argument of Section 5 makes this precise: any function  $\sum$  satisfying (A), (B), and (C) must, if it assigns Grandi's series a value at all, assign it  $\frac{1}{2}$ . The series is a near-miss to Cauchy's frame.

The historical record shows that this is exactly how Grandi's series was treated by the nineteenth century. It was a pathology. Cauchy declared it meaningless. But the construction of Cesàro summation in 1890, of Abel summation by 1826, and of the full theory of analytic continuation in the twentieth century, was driven in part by the demand to assign it the right value. The successive generalised summation methods are the successive richer frames built around Grandi's near-miss. Once the frame is rich enough, Grandi ceases to be pathological – it becomes routine machinery in Hardy's *Divergent Series* (Hardy, 1949), in Tauberian theory, in the zeta-regularised methods of mathematical physics.

### 8.3. $\sqrt[3]{2}$ as a near-miss

The Greek frame for "constructible number" similarly has at least four slots: expressible as a length, approached by rational lengths, root of a polynomial over  $\mathbb{Q}$ , and reachable by quadratic field extensions.

The cube root of 2 passes the first three slots and fails the fourth. The extension degree is 3, not a power of 2. The Greeks did not have the field-theoretic language to articulate this slot explicitly,

but the construction of that language – which is what Galois theory is – was driven in part by the demand to explain why  $\sqrt[3]{2}$  resists construction.

Once Galois theory is in place,  $\sqrt[3]{2}$  is no longer pathological. It is a perfectly understood algebraic number, inhabiting a precisely understood field extension. The pathology lived in the Greek frame, not in the number.

#### 8.4. The Weierstrass function and the irrationals chain

Two further examples make the same structural point.

The Weierstrass function – a function continuous everywhere and differentiable nowhere – was pathological to Hermite, who "turned away with fear and horror from this lamentable plague of functions which have no derivatives" (Hermite, 1893/1905). It is not pathological now. Twentieth-century real analysis incorporates such functions as ordinary objects; they are the rule among continuous functions, not the exception. Once measure theory and fractal geometry are in place, the Weierstrass function is no more strange than the circle.

The Liouville numbers – reals approximated by rationals to arbitrary order – were pathological to mid-nineteenth-century algebraic number theory. The class of algebraic irrationals was believed (without proof) to coincide with the class of badly approximable irrationals. Liouville's 1844 construction produced explicit reals that violated this expectation, and the class of *Liouville numbers* was born. They were pathological in the sense that they did not fit the prevailing frame.

By 1955, when Roth proved that every algebraic irrational has irrationality measure exactly 2 (Roth, 1955), the Liouville numbers were not pathological at all. They are precisely the transcendentals approximated too well to be algebraic. The frame of *irrationality measure* – which is, in a sense, the uniframe of the chain  $\sqrt{2} \rightarrow e \rightarrow \pi \rightarrow$  Liouville of approximations – absorbs them as ordinary objects. They were the near-misses that forced the construction of the frame.

#### 8.5. Why pathology fades

The pattern in all three cases is the same. An object is labelled pathological while the current frame cannot accommodate it. The mathematics that grows up to accommodate it does so by enriching the frame – by making explicit a slot that the original frame had treated implicitly. Once the frame is rich enough, the pathological character of the object disappears. *Not because the object has changed* – the Weierstrass function is the same function in 2026 that it was in 1872 – *but because the frame has*.

This is the temporal asymmetry of pathology. An object is pathological only relative to the frame against which it presents as a near-miss. Once the frame absorbs it, the pathology is gone.

Pathological objects, on this account, are the cognitive engines of mathematical progress. They are exactly the objects whose pathological character forces the frame to grow.

### 9. Pedagogical implications

The Pathology Principle is, in the end, a claim about how mathematicians learn. The historical near-misses – Grandi,  $\sqrt[3]{2}$ , Weierstrass, Liouville – are the same objects that served as the cognitive triggers for the construction of richer frames. What the principle suggests, then, is a particular style of mathematics teaching.

#### 9.1. Four pedagogical implications

We can state the implications as four classroom-level guidelines.

**1. Do not smooth near-misses away.** When a student produces a near-miss – an almost-correct proof, an example that "feels wrong but I can't say why," an analogy that breaks at one point – do not correct it into oblivion. The failure is the data. The student has produced an object that satisfies

most of the slots of the relevant frame and fails exactly one. The pedagogical work is to find and name that slot.

A familiar example: a student of group theory shows that the set of  $n \times n$  matrices with positive entries forms a structure under matrix multiplication. They prove closure; they prove associativity. The structure has an identity (the identity matrix). They claim it is a group. The near-miss is the inverses slot: not every positive matrix has a positive inverse. The student has produced an object that satisfies three out of four slots of the group frame. Naming the failing slot is exactly the pedagogical move that constructs the concept of *monoid*.

**2. Name the slot that failed.** The pedagogical labour of the near-miss is to make the failing slot visible *as a slot*. "Your argument used commutativity here. What if the operation is not commutative?" is a different intervention from "This is wrong." The first names the slot; the second does not.

This is harder than it sounds. The teacher has to know which slot failed – which often requires a structural understanding of the relevant frame that goes beyond the surface of the calculation. But the work is the work.

**3. Build the chain forward.** Once one slot is named, the next pedagogical move is to look for an object that fails the next slot. The chain  $\sqrt{2} \rightarrow e \rightarrow \pi \rightarrow$  is a canonical example:  $\sqrt{2}$  has a periodic continued fraction;  $e$  has a patterned but non-periodic one;  $\pi$  has an apparently random one. Each is a near-miss to the previous frame. Together they build, slot by slot, towards Lagrange's theorem and ultimately [Roth's \(1955\)](#).

A teacher with the chain at hand can guide students through the sequence of slots that Roth's theorem makes explicit. Without the chain, the theorem can only be presented as a finished object.

**4. Trust the historical record.** Most of the pathological objects in the history of mathematics were pedagogical near-misses for the mathematicians who first encountered them. Grandi's series was a near-miss for Leibniz; the Weierstrass function was a near-miss for Hermite; Liouville numbers were a near-miss for the algebraic number theorists of the 1840s. Teach with that grain, not against it. The historical record contains the chain of near-misses through which the concept actually developed. Reconstructing that chain in the classroom is usually a more productive pedagogical move than presenting the finished theory.

## 9.2. A worked example: introducing real analysis

To make the implications concrete, we sketch a worked example of how one might use the framework to introduce real analysis to a class of mathematics-education students or strong secondary-school students.

Begin with the frame of "well-behaved function": continuous, differentiable, with a Taylor expansion that converges to the function. The first three slots are easily made explicit, because the textbook examples (polynomials, sin, cos, exp) satisfy all of them.

Now introduce, in sequence, three near-misses:

(a)  $f(x) = |x|$  at  $x = 0$ . Continuous; not differentiable. Fails the differentiability slot.

(b) Cauchy's function  $f(x) = e^{-\frac{1}{x^2}}$  for  $x \neq 0$ ,  $f(0) = 0$ . Continuous, infinitely differentiable; Taylor expansion at 0 is identically zero. Fails the Taylor-convergence slot.

(c) The Weierstrass function. Continuous; nowhere differentiable. Fails the differentiability slot globally, not just at isolated points.

Each near-miss makes a slot explicit. By the time the class has worked through (a), (b), and (c), the students have arrived at the modern frame for "well-behaved function" – which is a richer and more honest frame than the one they started with. They have constructed the frame themselves, by being walked through the chain of near-misses that the historical mathematicians constructed it through.

### 9.3. Implications for teacher education

The framework also has implications for the training of mathematics teachers, which is a particular interest of this conference.

A teacher who has internalised the near-miss framework is a teacher who can hear a student's wrong answer as data rather than as failure. This is a cognitive shift in the teacher, not a methodological tweak. It requires the teacher to maintain, in the moment of hearing the wrong answer, a working representation of the frame the student is constructing and the slot the student's answer is failing. That is a high-bandwidth cognitive task, and it does not develop spontaneously. It has to be trained.

We suggest that mathematics teacher education programmes include explicit work on the near-miss framework: case studies of historical near-misses; structured exercises in identifying the failing slot of a wrong answer; reflection on the teacher's own near-misses in their mathematical education. The empirical research on whether such training is effective is, as far as we know, still to be done; we offer the framework as a hypothesis worth testing.

### 9.4. Exercises for the reader

To make the framework operational, we close this section with three exercises that can be adapted for use in a teacher-education course or a methods seminar.

**Exercise 1: identifying the failing slot.** For each of the following student answers, identify the implicit frame that the student appears to be working with, list its slots, and identify the slot that the student's answer fails.

- "Every continuous function is differentiable."
- "If  $f(x) > 0$  for all  $x$ , then  $\int f$  is positive."
- "A series of positive terms converges if and only if its terms tend to zero."
- "Every infinite set has the same cardinality."

(These are intentionally textbook "misconceptions." The framework recasts each as a near-miss to a particular frame, and asks the teacher to name the failing slot rather than label the answer "wrong.")

**Exercise 2: building a chain.** Choose a theorem from the secondary-school curriculum – for example, the rational root theorem, the intermediate value theorem, or the fundamental theorem of algebra. Construct a chain of three objects, each a near-miss to the previous frame, that together motivate the theorem. The chain should end at an object whose only way of being absorbed is the theorem itself.

**Exercise 3: identifying a pathology.** Choose an object from the recent history of mathematics that was, at some point, called pathological. (Candidates: Cantor's middle-thirds set; the Banach–Tarski paradox; the space-filling curve; the Peano curve; the Dirichlet function; a continuous function that is monotonic but has a derivative equal to zero almost everywhere.) Identify the frame against which the object was pathological. Identify the richer frame that eventually absorbed it. Identify the slot that the richer frame made explicit.

These exercises can be done in pairs or in small groups; the discussion is, in our experience, more productive than the answers.

## 10. Open questions and a scoping note

We have presented the near-miss framework as a pedagogical hypothesis, and it would be a mistake to overstate the evidence for it. We close with two open questions and one acknowledged limit.

**When is a near-miss productive?** Not every wrong answer is a productive near-miss. A student who multiplies  $7 \times 8$  and gets 54 has made an arithmetic error, not a near-miss to a structural frame. The framework is most useful when the wrong answer reflects a *systematic* engagement with

a frame – when the student has correctly applied most of a structure and failed on one feature. Telling productive from unproductive near-misses in the moment of classroom teaching is itself a skill, and the empirical question of how to train teachers in this skill is open.

**Is the framework testable?** The strongest version of the framework predicts that explicitly naming the failing slot of a near-miss should produce better learning outcomes than either correcting the answer silently or rejecting it as wrong. This is, in principle, a testable prediction. Designing a study that controls for the many confounds – teacher quality, student preparation, topic difficulty – is non-trivial, but not beyond the reach of the mathematics-education research community. We invite such studies.

**Scoping note.** The framework is for the *conceptual* layer of mathematical learning, not the procedural one. A student who has not memorised the multiplication table is in no position to be productively challenged with near-misses to the prime number theorem. The framework presupposes a base of procedural competence that it does not, by itself, provide. This is a scoping comment, not a fatal objection: most of the field's pedagogical attention is now directed at the conceptual layer, and that is where this paper is aimed.

## 11. Conclusion: three lessons, one insight

We began with the dichotomy between mathematics as it is created and mathematics as it is presented. We end with a unified picture of how to bridge them.

**Uncertainty** is the condition of mathematical work. The conjectures we cannot close keep the field alive. The shifts in foundation that every century has paid the price of are the field's evidence of its own incompleteness. Teaching mathematics honestly means teaching its uncertainty alongside its certainty.

**Rule-bending** is the method. The cube root of 2 and the value  $-\frac{1}{12}$  are both reachable only by changing the rules – carefully, deliberately, with new axioms or new operations – and the new mathematics that results is genuine. Most of the major advances in the field have been, at the level of methodology, rule-bending of this kind.

**Near-misses** are the cognitive engine. Pathological objects are the historical near-misses that forced the frame to grow. They are also the pedagogical near-misses that, in the classroom, force the student's frame to grow. The teacher's job is not to hide them, smooth them away, or correct them prematurely. The teacher's job is to make the failing slot visible – to give the student the experience of seeing, with Lonergan, a feature she had not noticed she was assuming. These are not three separate lessons. They are the same insight from three angles. Uncertainty makes the rule-bending possible. Rule-bending creates the new frames. The near-miss is the unit through which the new frames are cognitively constructed.

The arch the child sees has a gap underneath. Most adults, asked what an arch is, will not mention the gap; they take it for granted. The mathematics teacher's task, in the deepest sense, is to teach her students to see the gap – to see the slot that was working invisibly in every example she had ever met, and that becomes visible only when one example, almost identical to all the others, lacks it.

The arch-ness of the arch lives in the gap. The mathematics-ness of mathematics lives in the slots that the near-misses make visible.

## Acknowledgment

I thank the Institute of Mathematics and Informatics at the Faculty of Science, University of Kragujevac, and the Serbian Mathematical Society, for the invitation to deliver the plenary lecture of which this paper is the companion.

## References

- BEGHETTO, R. A., & KAUFMAN, J. C. (2007). Toward a broader conception of creativity: A case for "mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 1(2), 73–79.
- BERNDT, B. C., & RANKIN, R. A. (1995). *Ramanujan: Letters and Commentary*. American Mathematical Society / London Mathematical Society, Providence, RI.
- CAUCHY, A.-L. (1821). *Cours d'analyse de l'École royale polytechnique; I.<sup>re</sup> Partie. Analyse algébrique*. Imprimerie royale, Paris.
- HARDY, G. H. (1949). *Divergent Series*. Oxford University Press, Oxford.
- HERMITE, C. (1893/1905). Letter to T. J. Stieltjes, 20 May 1893. In B. Baillaud & H. Bourget (Eds.), *Correspondance d'Hermite et de Stieltjes, II*, p. 318. Gauthier-Villars, Paris.
- LAKATOS, I. (1976). *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge University Press, Cambridge.
- LAMOREAUX, S. K. (1997). Demonstration of the Casimir force in the 0.6 to 6  $\mu\text{m}$  range. *Physical Review Letters*, 78(1), 5–8.
- LONERGAN, B. J. F. (1957). *Insight: A Study of Human Understanding*. Longmans, Green and Co., London.
- MESSER, P. (1986). Problem 1054. *Crux Mathematicorum*, 12(10), 284–285.
- MINSKY, M. (1986). *The Society of Mind*. Simon & Schuster, New York.
- ROBINSON, A. (1966). *Non-Standard Analysis*. North-Holland, Amsterdam.
- ROTH, K. F. (1955). Rational approximations to algebraic numbers. *Mathematika*, 2, 1–20.
- SRIRAMAN, B. (2026). Uncertainty, rule-bending, and the near-miss. Plenary lecture, *TEMATCOM 2026*, University of Kragujevac, May/June 2026.
- WINSTON, P. H. (1970). *Learning Structural Descriptions from Examples*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.