The Attralucian Essays:

Exploring the Finite

First Edition

The Attralucian Essays

The Pi Files: A Geometric Detective Story and the Case for a New Way of Knowing

Kevin R. Haylett

Overview

Prologue: The Case of the Clashing Witnesses

Every good detective story starts with a contradiction. Two reliable witnesses describe the same event in completely different ways. One's account is vivid, full of shape and motion. The other's is clean, precise, and numerical. Who do you trust?

This isn't a scene from a paperback novel. It's the central mystery of one of mathematics' most famous constants: π (pi).

For centuries, π has been the ultimate symbol of randomness. Its endless string of digits after the decimal point—3.1415926535...—has passed every statistical test ever thrown at it. The Accountant's Witness is unimpeachable: "By the numbers," it states, "π is random. There is no pattern. Case closed."

But then, a new kind of witness took the stand. By using a technique from physics called delay embedding, we could for the first time see the digits of π not as a list, but as a path unfolding in three dimensions. This Geometric Witness looked at the same data and said something astonishing:

> "Wait. Look. When I view it through this

> specific lens ($\tau = 1$), the path is a coiled filament, swirling and organic. But when I change the lens slightly ($\tau = 5$), it becomes an angular scaffold, full of sharp, rectangular voids. These are not the same thing."

And so the mystery was born. How can something be random by every measure, yet produce such starkly different and compelling geometries? Is the eye fooling us? Or are our measures not measuring the right thing?

The Accountant's Testimony: The Power and Poverty of Flattening

Faced with the geometric witness's shocking claim, the scientific establishment did what it always does: it turned to statistics. This isn't a bad instinct. Statistics has been our most powerful tool for the last 400 years. Its power is flattening.

Imagine a beautiful, winding path through a forest. A statistician might produce a flawless report: the average width of the trees, the density of bushes per acre, the mean distance traveled. Their report is technically correct. But it erases the path itself.

Measures like Mutual Information, Transition Matrices, and Recurrence Quantification Analysis (RQA) all return the same verdict: "No significant departure from ran-

domness." The Accountant concludes: "The tree density is the same in both forests. Therefore, there is no real difference." This is the "Statistical Prior"—the assumption that if the flattened numbers don't show a difference, no difference exists.

The Geometric Witness: A Testimony of Shape

But what if the Geometric Witness isn't hallucinating? What if they are simply testifying to an aspect of reality that the Accountant's tools cannot measure?

The method is straightforward: treat the digits of π as a time series and embed them in 3D space using Takens' method:

$$x_t = [x_t, x_{t+\tau}, x_{t+2\tau}] \tag{1}$$

where τ is the delay parameter—essentially, the "time lens" through which we view the sequence.

For $\tau = 1$, the trajectory appears coiled, dense, and filamentary, like a strange new form of DNA.

For $\tau = 5$, the shape snaps into something entirely different: scaffolded, angular, full of sharp, rectangular voids.

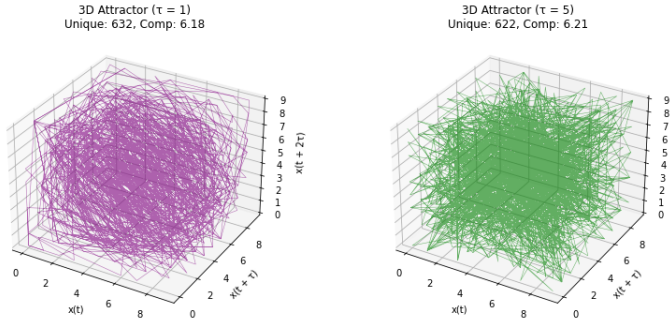

Fig. 1 Side-by-side 3D embeddings for $\tau = 1$ and $\tau = 5$.

The difference is immediate and striking to the human eye. The Accountant's tools, by design, destroy temporal order and are blind to this connective tissue. Of course they cannot see the difference between a coil and a scaffold.

Enter the Geofinitist Detective: A New Lens for the Case

Geofinitism argues: stop trying to prove one witness right and the other wrong. Instead, situate their testimonies.

Statistics is not an external judge. It is a finite procedure, powerful for questions of "how much" but the wrong tool for questions of "what shape." Changing τ is not a flawed experiment but a controlled perturbation that reveals invariants.

The Geofinitist detective's first move is to declare: "The

look is the data." The observed difference between the coil and the scaffold is not an illusion to be explained away; it is the primary phenomenon to be investigated.

The Smoking Gun: AI as an Independent Eyewitness

Geometry as Measurement

Changing τ is itself a measurement ritual—just as shuffling is a ritual. Both perturb the system, but they reveal different invariants.

From a geofinitist perspective:

- The difference between $\tau = 1$ and $\tau = 5$ is *data*.
- The "look"—the coils, the scaffolds, the voids—is a legitimate observation.
- Statistical invisibility does not imply non-existence; it only shows that a particular projection is insensitive to that feature.

Crucially, we are no longer limited to statistics to make this claim. Modern multimodal language models can act as measurement devices:

1. Render the two embeddings ($\tau = 1$ and $\tau = 5$) as images.

2. Pass them to the model for description rather than classification.

3. The model produces a natural-language narration, which encodes its internal high-dimensional representation of each image into a linguistic form.

For example, using a vision–language model:

> **Description a ($\tau = 1$):** "A dense, coiled structure with smooth, filament-like paths weaving through the cube. The pattern appears rotational, with few large voids."
>
> **Description b ($\tau = 5$):** "A more rigid, scaffold-like pattern with sharp angles and rectangular gaps, as if points lie along grid edges forming a lattice."

These descriptions are not arbitrary text, they are linguistic projections of very high-dimensional image embeddings. The fact that the model consistently produces distinct and semantically coherent descriptions for $\tau = 1$ and $\tau = 5$ demonstrates:

- The difference is encoded in the model's internal representation.
- This representation can be measured numerically (e.g., cosine distance between embedding vectors) or narrated as natural language.
- In both cases, the result captures the *geometry* of

the data.

This approach effectively turns words into *measurements*: the semantic difference between "coiled" and "scaffolded" is itself a finite indicator of distinct positions in the model's latent space.

The New Beat: Patrolling a Geometric World

This case forces a shift: if we care about shape, structure, and connection, we need a new toolkit designed for that purpose:

- **Topological Data Analysis**: To measure loops, voids, and connected components.
- **Spectral Graph Analysis**: To reveal harmonic signatures of paths.
- **Image-Embedding Similarity**: To treat AI not as a classifier but as a measurement instrument.

Mathematical Framing: Flattening vs. Embedding

We may formalize the contrast between statistical reduction and high-dimensional embedding as follows.

Statistical Flattening

A statistical functional f takes the trajectory $\{x_t\}$ and collapses it into a scalar summary:

$$S = f\big(\{x_t\}\big) \quad \mapsto \quad s \in \mathbb{R}$$

where S represents the set of observations (e.g., delay vectors), and s is a single number such as mutual information, determinism, or a p-value. This procedure is a *dimensional reduction* that discards the geometric configuration of the trajectory.

High-Dimensional Embedding

By contrast, a multimodal model uses an encoder Φ to project the entire rendered image I of the trajectory into a high-dimensional latent space:

$$\Phi : \mathcal{I} \to \mathbb{R}^d, \qquad d \gg 3$$

where $\mathcal{I}$ is the image space and d is the embedding dimension (often $d = 512$ or 768 in modern vision-language models).

The semantic difference between two embeddings (e.g., for trajectories with different delays τ_1, τ_2) is then quantified by a metric such as the Euclidean distance:

$$D(\tau_1, \tau_2) = \left\| \Phi(I_{\tau_1}) - \Phi(I_{\tau_2}) \right\|_2$$

This D is a *geometric measure*: it respects the global shape of the trajectory as captured in the image and preserves information that scalar statistics flatten away. In this sense, the embedding provides a measurement that is sensitive to the *flow* of the sequence, not merely its occupancy.

From Pixels to High-Dimensional Geometry

A digital image is, in its raw form, a two-dimensional array of pixel intensities. However, a modern image encoder does not operate directly on pixels as independent values. Instead, it applies a hierarchy of learned filters and nonlinear transformations, progressively mapping the image into a series of higher-dimensional feature representations:

- **Low-level features:** edge detectors and local contrast filters extract gradients and orientations.
- **Mid-level features:** combinations of edges form contours, textures, and object parts, capturing local shape information.
- **High-level features:** global patterns, object-level semantics, and even implied three-dimensional rela-

tionships are encoded, producing a structured representation of the entire image.

Mathematically, this process can be seen as a sequence of mappings

$$I \xrightarrow{\phi_1} F_1 \xrightarrow{\phi_2} F_2 \xrightarrow{\cdots} F_L = \Phi(I),$$

where I is the input image, F_ℓ is the feature representation at layer ℓ, and $\Phi(I)$ is the final embedding vector in a high-dimensional space $\mathbb{R}^d$ with $d \gg 3$.

This embedding preserves global geometric information: two images that are visually distinct to a human observer (e.g. a coiled trajectory vs. a scaffolded one) will map to well-separated points in the latent space. This separation can be measured quantitatively, for instance with the Euclidean distance

$$D(\tau_1, \tau_2) = \left\| \Phi(I_{\tau_1}) - \Phi(I_{\tau_2}) \right\|_2,$$

providing a geometric measure of difference that respects the flow structure captured in the image rather than flattening it to a single scalar statistic.

The Verdict: An Invitation to See

The investigation into π's geometry reveals a deeper truth about knowledge itself. The way forward is not to force π's geometry into null hypothesis tests but to curate it.

We propose an "Atlas of π's Faces," a gallery of embeddings for different τ, described in clear language and quantified with geometric tools.

This is an invitation to trust your eyes, to value the questions that arise from looking closely, and to join in mapping the manifold of π.

Appendix 1: Delay Embeddings and Takens' Theorem

Takens' theorem is a foundational result in nonlinear dynamics which guarantees that, under certain conditions, the dynamics of an unknown system can be reconstructed from a single scalar time series. The idea is to create vectors of delayed observations:

$$x_t = [x_t, x_{t+\tau}, x_{t+2\tau}, \ldots, x_{t+(m-1)\tau}] \tag{2}$$

where m is the embedding dimension and τ is the delay. For sufficiently large m, the set of reconstructed points preserves the topological properties of the original attractor: trajectories, loops, and invariant measures.

In the case of π, there is no underlying dynamical system in the classical sense, yet the embedding constructs a geometric object that captures correlations across scales. Varying τ can reveal different structural properties—sensitive dependence on τ is itself informative, suggesting that the sequence is not uniformly featureless but exhibits multi-scale geometry.

Appendix 2. Numerical Experiments and Plots

This appendix summarizes the key numerical experiments conducted to probe the τ-dependent structure of π's digit sequence. The aim is to complement the qualitative observations in the main text with quantitative context.

Data Extraction

The first $N = 10{,}000$ digits of π after the decimal point were generated using `mpmath` at high precision:

```
from mpmath import mp
import numpy as np

mp.dps = 11050  # headroom for rounding
s = mp.nstr(mp.pi, n=11002)
after_decimal = s.split('.')[1][:10000]
x = np.fromiter((ord(c)-48 for c
in after_decimal), dtype=np.int8)
```

This produced a vector $x \in \{0, \ldots, 9\}^N$ suitable for delay embeddings.

Delay Embeddings

We constructed 3D delay vectors for several choices of τ:

$$\mathbf{x}_t^{(\tau)} = [x_t, x_{t+\tau}, x_{t+2\tau}],$$

with $\tau \in \{1, 5, 10\}$.

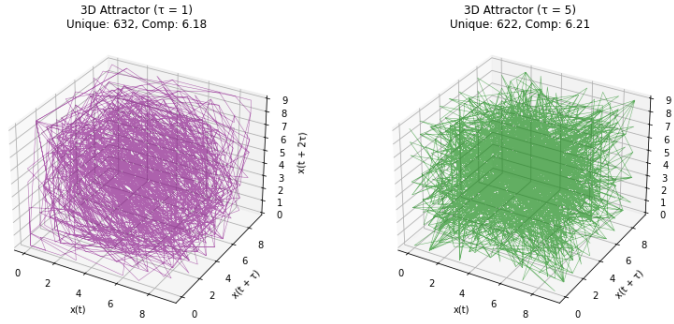

Figure 1: Side-by-side 3D trajectory plots for $\tau = 1$ and $\tau = 5$.

Pairwise Distance Distributions

Pairwise Euclidean distance distributions between $\tau = 1$ and $\tau = 5$ embeddings were compared using the Kolmogorov–Smirnov test:

$$D_{\mathrm{KS}} = 0.0037, \qquad p \approx 0.0025,$$

indicating a statistically significant, but very small, distributional shift—consistent with the interpretation that the point occupancy is similar but the trajectory wiring differs.

Average Mutual Information (AMI)

Average Mutual Information (AMI) was computed for $\tau = 1 \ldots 20$ with Miller–Madow bias correction and com-

pared against 200 shuffled surrogates.

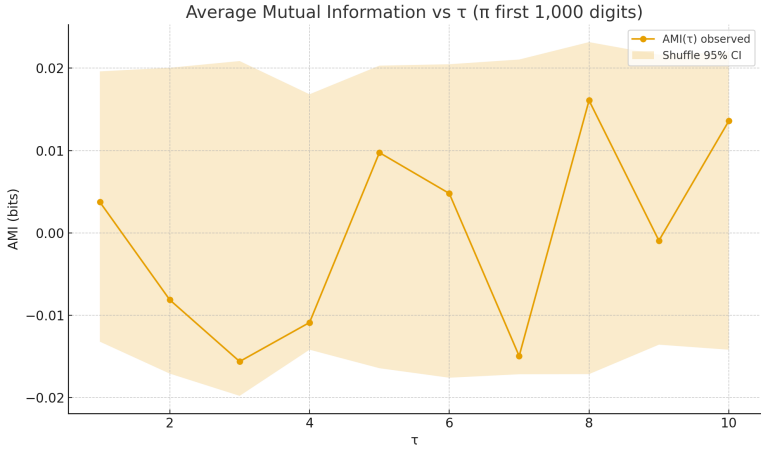

Figure 2: AMI versus τ with 95% surrogate confidence interval shading.

All AMI values fell within the surrogate confidence bands, showing no detectable lag dependence at this scale.

Transition Matrices

Lag-τ transition matrices

$$P_{ij}^{(\tau)} = P(x_{t+\tau} = j \mid x_t = i)$$

were computed for $\tau = 1$ and $\tau = 5$ and compared to surrogates via z-scores.

Probabilities were near-uniform; deviations were scattered and consistent with random fluctuation.

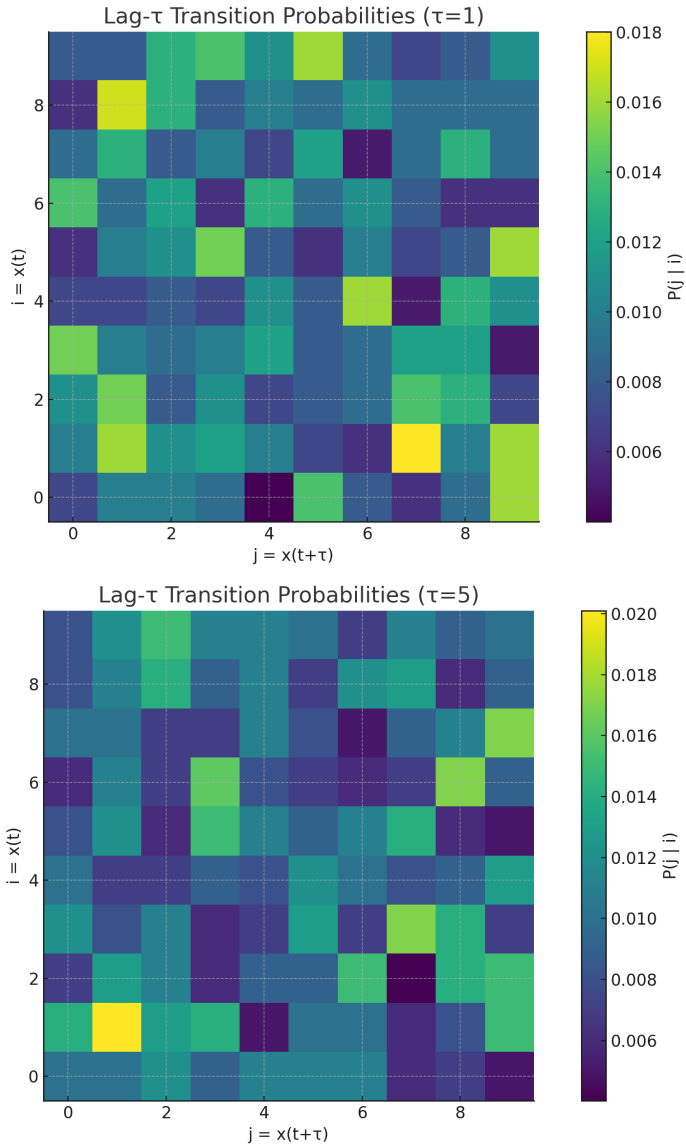

Figure 3: Transition probability heatmaps (top) and z-score maps (bottom) for $\tau = 1$ and $\tau = 5$.

Principal Component Analysis (PCA)

Principal component spectra of the delay vectors were computed for $\tau = 1$ and $\tau = 5$. Observed explained-variance ratios were within the surrogate confidence intervals, indicating no significant dimensionality reduction effects.

Recurrence Quantification Analysis (RQA)

Using a Theiler window $w = (m-1)\tau$ and recurrence threshold ε set to the 10th percentile of pairwise distances, we obtained the following RQA metrics:

Metric	$\tau = 1$ (Obs.)	$\tau = 1$ (CI)	$\tau = 5$ (Obs.)	$\tau = 5$ (CI)
RR	0.102 (set)	–	0.102	–
DET	0.396	0.37–0.40	0.029	0.027–0.031
L_{mean}	3.76	3.6–3.8	3.12	3.0–3.2

Table 1: RQA metrics for $\tau = 1$ and $\tau = 5$ with 95% surrogate confidence intervals.

These results show no significant departure from i.i.d. surrogates once proper controls are applied.

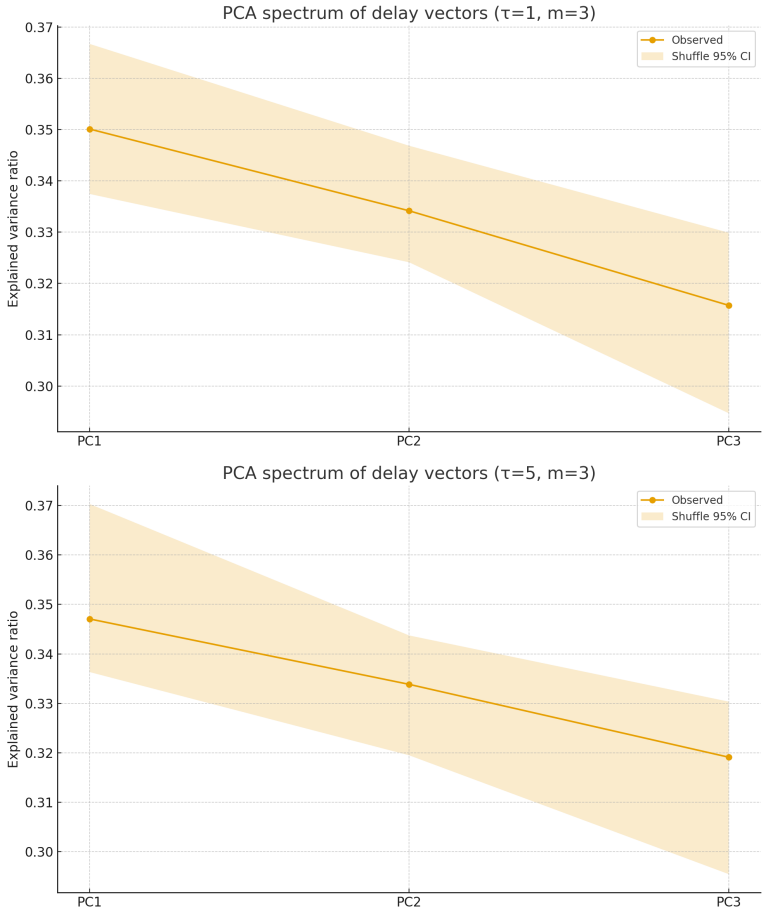

Figure 4: PCA explained variance spectra for $\tau = 1$ (top) and $\tau = 5$ (bottom)with surrogate confidence intervals.