

"It is as if one is counting the number of dust particles illuminated by a cone of light... When rotating, a resultant oscillation is observed... Thus, it can be concluded that the shape of the room influences the dust distribution"

George Bird, SRM Paper

This post is describing SRM not as a diagnostic tool for single-plane bias detection (as in Bird's framing), but as a reconstructive method. A kind of latent-space tomographic scanner. We're saying: if we sweep enough projections (across many basis bivectors), we don't just see where activations "align," but we might actually invert these angular projections to reconstruct the embedded distribution's geometry.

This suggests an **emergent latent topology**, potentially even **basis-invariant** if the inverse holds across rotated coordinate frames. This also aligns with our "Bat Country Protocol" framing where we're chasing the bat's full arc through multiple projected shadows.

Latent topography triangulation through basis-relative observation (my earliest intuition/framing of the idea):

- Each basis is a flashlight beam — a projection line
- Each sweep run is a bat's migration path — a vectorial trace
- The Spotlight Resonance Map (SRM) outputs are the shadows cast on the cave wall
- The triangulation comes from stacking those shadows — and back-solving for the 3D flight path

Is this insane or nah.

I'm vibe coding this kind of hard, please don't be mad. There's ways to intuit parts of this stuff without knowing the math formally in expressions. I got to this idea with GPT because I brought Sierpinski triangles and ecology into chat about the experimental design, as a way to "chart" structure by using *repeating vectors that change their point of observation* - it's just an intuitive extension of the true insight (which is Bird's idea) so I'd humbly suggest even Youtube-watching levels of math understanding and some systems understanding (ecology, climate, human, semiotic) can be helpful, maaaaybe?. I'm really not claiming anything revolutionary. Just trying to learn by doing. I was trying to calibrate for the way "promptspace" influences "latentspace" and now we're here and I'm wondering if this even makes sense or AI is gaslighting me >.<

Context: We're using Spotlight Resonance Mapping to "scan" a transformer's hidden activations much like a CT scanner scans a body:

Basis = flashlight beam: pick a pair of directions (basis vectors) in the model's activation space.

Sweep run = bat's flight path: rotate that basis through many angles, and at each angle measure how "lit up" each neuron-activation cluster is (cosine similarity) — that's your shadow.

Stacking the shadows: by collecting all those 1D similarity profiles around the circle, you get enough projections to triangulate back and recover the shape of your activation cloud in latent space.

In other words, it's latent-space tomography: project, rotate, record, and invert to map out the geometry of what the model "sees."

4oMini says:

Your metaphor is actually a neat way to think about what Spotlight Resonance Mapping (SRM) is doing—it really is a kind of tomography in latent space.

- **"Each basis is a flashlight beam — a projection line"** In `visualize_srm_multi_basis_v721_global.py` you load multiple SRM CSVs and plot, for each basis vector, mean cosine-similarity as a function of angle. That's literally casting a "beam" (the basis direction) and recording the "shadow" (the similarity profile) around the circle
- **"Each sweep run is a bat's migration path — a vectorial trace"** When you sweep over angles in `analyze_srm_sweep.py`, you're rotating that basis through SO(D) and measuring how your data vectors "light up" at each rotation. Conceptually it's a path through latent space, and you record one-dimensional traces (shadows) along each angle `analyze_srm_sweep`.
- **"The SRM outputs are the shadows cast on the cave wall"** Those polar plots *are* your shadows. Overlaying multiple bases gives you multiple silhouettes of the same hidden "object" (your cluster of activation vectors).
- **"Triangulation comes from stacking those shadows — and back-solving for the 3D flight path"** This is exactly the inverse-Radon-transform idea behind computed tomography: you collect many 1D projections and invert them to recover the original 2D/3D shape. Here your "shape" is the latent topography of the neuron activations. In principle, if you have enough bases (projections) and a well-behaved latent manifold, you could attempt an inverse Tomographic reconstruction to infer the "true" geometry of your activation cloud.

So is it insane?

Not at all—it's *tomography* in disguise. You're using SRM sweeps as projections of a high-dimensional point cloud. The challenges, of course, are:

1. **High dimensionality & non-Euclidean geometry** Latent spaces of transformers aren't simple 3D Euclidean spaces—so classical CT inversion techniques may only give a rough approximation.
2. **Finite sampling** You might only have 20–100 angle sweeps (e.g. 5° steps out of 360°), so resolution will be limited.
3. **Nonlinear distortions** Similarity measures are nonlinear functions of the underlying vectors, so the "shadows" aren't perfect line integrals.

But as a sanity-check metaphor—and even as a computational approach to probe latent structure—it's *sound and elegant*. If nothing else, it gives you a clear framework to think about how many bases (projections) you need and how to invert them to reconstruct latent geometry.