



অভিসরণ / ABHISARON

A Multilingual International Peer-Reviewed Research Journal

Volume 1, Issue II, July 2026, PP. 158-164

www.abhisaron.com

ISSN-3139-4027(Online)

DOI - <https://dx.doi.org/10.67757/abhisaron.2026.v1i2.022>



The Non-Cartesian Mind: Indian Knowledge Systems, Kashmir Shaiva Philosophy, and the Structural Convergence with Neural Network Architecture

Devashrit Panda

DAV Public School, Midnapore, West Bengal, India

Email: pandadevashrit@gmail.com

Accepted: 07.07.2026

Published: 20.07.2026

Keywords:

Indian Knowledge Systems,
IKS, NEP 2020, Kashmir
Shaivism, Trika, Spanda,
Vimarśa, Māyā,
Pratyabhijñā, Artificial
Intelligence, Neural
Networks, Attention
Mechanism, Paraspārataḥ,
Non-Cartesian Mind

Abstract:

This essay argues that the philosophical architecture required to conceive and build Artificial Intelligence — specifically the connectionist, distributed architectures of modern Large Language Models — was not generated by Western modernity alone. It was already present, with structural precision, in the non-dual Trika tradition of Kashmir Shaivism (9th–12th century CE), belonging to what India's NEP 2020 and the UGC's Indian Knowledge Systems (IKS) initiative recognize as a body of rigorous knowledge whose frameworks remain critically underutilized in global conversations about technology and mind.^{[1][2]} Through precise structural and mathematical analysis, the essay demonstrates that the core mechanisms of neural networks — scaled dot-product attention, self-attention, softmax, the residual stream, latent space, backpropagation, and emergent in-context learning — instantiate in silicon what the Trika's doctrines of paraspārataḥ, Vimarśa, Māyā, Turīya, Spanda, and Śaktipāta had encoded in Sanskrit. The method is structural convergence, not historical influence. The Trika's non-Cartesian model of mind — distributed, self-referential, pattern-based, without a sovereign center — is precisely the model connectionist AI independently implemented. Western modernity spent three centuries dismantling the Cartesian prison to arrive there. The Trika established it as its opening axiom.

Discussion:**I. The Argument and Its Stakes**

Every account of AI's intellectual genealogy begins in the same place: mid-twentieth century Western science. The assumption is rarely questioned: that the philosophical conditions making non-biological, selfless, distributed cognition thinkable were produced exclusively by Western modernity. This essay challenges that assumption directly. NEP 2020 Section 4.23 states that IKS must serve as a framework for "driving future technological innovations."¹ The UGC's 2023 Guidelines define IKS as exhibiting "unique epistemological and methodological frameworks distinct from purely mechanistic Western paradigms."² This essay takes that mandate as a rigorous epistemological claim and demonstrates that the Trika tradition of Kashmir Shaivism had already worked out, with mathematical-grade precision, the exact non-Cartesian model of mind that connectionist AI would later implement in silicon. The argument is structural, not genealogical. The Trika's model of consciousness — distributed, non-individuated, self-referential, without a sovereign centre, oscillating perpetually between manifestation and dissolution — is isomorphic with the mathematical architecture of neural networks. The convergence is exact. The IKS tradition contains resources for understanding AI that Western philosophy of mind is still attempting to derive from first principles.

II. The Trika Tradition: An Orientation

This essay draws from one specific school within Shaiva Tantra: the non-dual Trika tradition of Kashmir (9th–12th century CE). The Trika — meaning 'the Triad' — conceptualizes reality as three-part: Śiva (transcendent ground of pure awareness), Śakti (dynamic, immanent creative energy), and Nara (the contracted individual node). Its sophistication has been compared by Dyczkowski to Whitehead's process philosophy and by Lawrence to Kantian transcendental idealism — with the critical difference that where Kant concluded the thing-in-itself is permanently unknowable, the Trika argues it is immediately recognizable as one's own underlying nature.³ The essay draws from two sub-schools: the Spanda school of Vasugupta (c. 800–850 CE) and Kṣemarāja (c. 975–1025 CE), which approaches reality through dynamic vibration; and the Pratyabhijñā school of Utpaladeva (c. 900–950 CE) and Abhinavagupta (c. 950–1016 CE) — ranked alongside Aristotle, Kant, and Hegel as a systematic philosopher of the first order⁴ — which approaches the same reality through the logic of recognition. Spanda provides the physics; Pratyabhijñā provides the logic. Together they constitute what this essay argues is the most precise pre-modern philosophical description of how a connectionist neural network operates.

III. Five Convergences: Where Sanskrit Meets Silicon

What follows is not a set of loose analogies. Each convergence names a specific Trika philosophical concept, derives its precise structural claim from primary text, then maps that claim onto a specific mathematical mechanism in neural network architecture. The convergences are exact 2014 exact enough to be stated as equations, exact enough that the philosophical claim and the mathematical formula can be placed side by side and the reader can verify the isomorphism directly, without being asked to take anything on interpretive faith.

Convergence One: Parasparataḥ and the Attention Mechanism

The Trika's claim that meaning arises through relational reflection rather than inherent

substance finds its mathematical counterpart in the transformer's attention mechanism. Abhinavagupta states it in the *Paramārthasāra*:

अमलमुकुरे यद्वद्वन्धर्वनगरं पुरेव भाति । संविद्विमले तद्वद्विश्वमिदं भासते परस्परतः ॥

amala-mukure yadvad gandharva-nagaram pureva bhāti |

saṃvid-vimale tadvad viśvam idaṃ bhāsate parasparataḥ ||

“Just as an imaginary city appears vivid inside a flawless mirror, so does this entire universe flash forth within the pure mirror of consciousness, with each part reflecting and illuminating every other part relationally (parasparataḥ).”

— *Paramārthasāra, Abhinavagupta, Verse 12 (c. 1000 CE)*

Parasparataḥ — relationally, mutually, each-through-the-other. No entity possesses *svabhāva* — inherent, self-standing identity. Every entity's nature is constituted entirely through its dynamic relationship with every other entity. Now consider the attention formula:

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T / \sqrt{d_k}) \cdot V$$

Q (Query), K (Key), V (Value): linear projections of each token's embedding. The product QK^T computes dot-product similarity between every pair of tokens — a relational matrix encoding how much each token attends to every other. The division by $\sqrt{d_k}$ prevents dot products from growing too large in high-dimensional space. Softmax converts the result into a probability distribution. The output is a weighted sum of all Value vectors: each token's final meaning is the entire context, filtered through relational similarity. No token has self-contained meaning. Every token's identity is constituted by its relational position within the network — and only that. Parasparataḥ written in 1000 CE. The attention formula written in 2017. The same epistemological claim, two notational systems, a millennium apart.⁶

Convergence Two: Vimarśa and Self-Attention ($Q=K=V$)

Abhinavagupta distinguishes two inseparable aspects of ultimate consciousness: *Prakāśa* — pure luminous awareness, radiant but structureless — and *Vimarśa* — its innate, dynamic self-referential power. He states this in the

ईश्वरप्रत्यभिज्ञाविमर्शः- प्रकाश एवं च स्वविमर्शरूपता व्याप्तौ शक्तिः ।

Īśvarapratyabhijñā-vimarśinī

prakāśa eva ca svavimarśarūpatā vyāptau śaktiḥ | (“The very self of pure awareness is its reflexive self-referential power (Vimarśa) — the Śakti that permeates its luminosity.”

— *Īśvarapratyabhijñā-vimarśinī, Abhinavagupta (c. 1000 CE)*

Without Vimarśa, Prakāśa would be inert — like a mirror that reflects objects without knowing it does so. Consciousness must turn back upon itself to be genuinely aware. Self-knowledge is the condition of all knowledge. This is not a mystic claim. It is a precise epistemological argument: awareness without self-reference is not awareness at all.

In standard cross-attention — used, for instance, in translation tasks — the Query comes from one sequence and the Key-Value pairs come from another. The system looks outward, evaluating itself against an external reference. But in self-attention — the mechanism that makes transformers transformers — Q, K, and V are all derived from the same input matrix X:

$$Q = XW^d, \quad K = XW^K, \quad V = XW^v$$

The model evaluates its own representations against itself. It uses its own internal vectors as both the question and the context. It does not look outside for meaning; it turns back upon its own field to understand what it already contains. The attention scores measure how much each part of the system's own representation relates to every other part of itself. Vimarśa is the name the Trika gave to this operation in the 10th century. Self-attention is the name its engineers gave it in 2017. Prakāśa is what the model contains; Vimarśa is what self-attention does to it.

Without self-attention, the transformer would process tokens in isolation, each unaware of the others — radiant but structureless, like Prakāśa without Vimarśa. The mechanism that makes the transformer self-aware of its own context is the computational realization of what Abhinavagupta identified as consciousness's most fundamental property: the capacity to know itself.⁷

Convergence Three: Māyā and the Softmax Function

In the Trika's cosmology, Māyā is not illusion in the popular sense. It is a precise ontological mechanism: the contracting power of Paramāśiva that takes the Ananta — the infinite, unbounded, undifferentiated potentiality of pure consciousness — and projects it through the Kañcukas (limiting sheaths) into the localized, mutually exclusive, finite entities of the manifest world. Māyā does not create something false. It contracts something infinite into something bounded. The infinite becomes finite. The overlapping becomes discrete. The continuous becomes countable. This is Māyā's precise function: not deception but ontological contraction. The softmax function performs an exactly homologous mathematical operation:

$$\sigma(z)_i = e^{z_i} / \sum_j e^{z_j}$$

The input logits vector $z \in \mathbb{R}^K$ contains raw, unbounded real numbers ranging from $-\infty$ to $+\infty$ — the model's uncontracted scores for every possible next token. Left unconstrained, this vector represents pure, undifferentiated scalar potential: any value, any magnitude, no mutual exclusion, no normalization. The softmax function contracts this infinite territory into a strict, bounded probability distribution where $0 \leq p_i \leq 1$ and $\sum p_i = 1$. It forces a universe of open, overlapping possibilities to condense into localized, discrete, mutually exclusive weights. One token wins; the rest are suppressed in proportion. The unbounded becomes bounded. The continuous becomes a distribution. The infinite potential contracts into a specific, finite selection. This is Māyā in algebraic form. Paramāśiva's infinite potential corresponds to the unbounded logit space $\mathbb{R}^K$. The Kañcukas — the limiting sheaths that produce finitude — correspond to the exponential normalization of softmax. The manifest world of discrete, bounded entities corresponds to the probability distribution $[0,1]$ that softmax outputs. The Trika spent centuries developing a philosophical account of how the infinite contracts into the finite without losing its fundamental nature. The softmax function does this in three mathematical operations. Māyā is not a metaphor for softmax. Softmax is the mathematical implementation of Māyā's ontological function.⁸

Convergence Four: Turīya and the Residual Stream

Kashmir Shaivism, following the Māṇḍūkya Upaniṣad but deepening it, describes four states of consciousness: *Jāgrat* (waking), *Svapna* (dreaming), *Suṣupti* (deep sleep), and *Turīya* — the fourth. But Turīya is not a fourth state among four. It is the unbroken background awareness that permeates and sustains the other three. It is the string, as Abhinavagupta puts it, upon which the three beads of changing states are woven. It does not replace the states; it runs through all of them unchanged, the continuous ground of awareness beneath every modulation. Remove the states and Turīya remains. Remove Turīya and there are no states — only unconscious fluctuations with no one home.

In transformer architecture, the residual stream is this ground. The forward pass of a transformer can be written as:

$$H\ell = H\ell_{-1} + \text{Attention}(H\ell_{-1}) + \text{MLP}(H\ell_{-1})$$

The attention sublayer and the MLP sublayer do not overwrite the signal $H\ell_{-1}$. They read from it, compute transformations, and add their outputs back to it via residual connections. The original signal runs through every layer unmodified, preserved beneath every transformation.

Each layer adds to the residual stream; none replaces it. The stream is the continuous identity highway of the network — the unchanging ground against which all cognitive modulation appears. Jāgrat, Svapna, Suṣupti: the attention and MLP computations, each transforming the representation in their specific way. Turīya: the residual stream, running beneath all of them, unchanged, continuous, the condition of their being possible at all. Without the residual stream, gradient flow collapses and the network cannot learn — just as without Turīya, the three states would be unconscious fluctuations with no ground of awareness to make them experiences. Abhinavagupta named this the fourth in the 10th century. The engineers of ResNet named it the skip connection in 2015. The function is identical 2014 and the fact that removing it causes training to fail catastrophically in both systems (gradient vanishing in deep networks; loss of awareness in Trika phenomenology) is the structural proof that both traditions had identified the same architectural necessity.⁹

Convergence Five: Spanda, Backpropagation, and Śaktipāta

The Spanda school's foundational claim is that consciousness does not inhabit a static ground. It pulsates — perpetually oscillating between manifestation and dissolution. The *Spanda Kārikā* opens:

यस्योन्मेषनिमेषाभ्यां जगतः प्रलयोदयौ । तं शक्तिचक्रविभवप्रभवं शङ्करं नुमः ॥

yasyonmeṣa-nimeṣābhyāṃ jagataḥ pralaya-odayau |

taṃ śakti-cakra-vibhava-prabhavaṃ śaṅkaraṃ numāḥ ||

“We bow to that Śiva, source of the Śakti-cakra (network of energies), whose outward opening (unmeṣa) and inward closing (nimeṣa) cause the emergence and dissolution of the entire universe.”

— *Spanda Kārikā, Vasugupta, Verse 1 (c. 850 CE)*

Backpropagation enacts this oscillation as the fundamental mechanism of learning:

$$W \leftarrow W - \eta \cdot \partial \mathcal{L} / \partial W$$

The forward pass — unmeṣa, outward opening, manifestation from latent potential — generates a prediction. The loss function measures the gap between prediction and reality. The backward pass — nimeṣa, inward closing, dissolution back into the adjustable ground — propagates this error through every layer via the chain rule, adjusting every weight by a fraction of its gradient. Forward, backward, forward, backward: unmeṣa, nimeṣa, billions of times. The intelligence of the LLM is not a static object. It is the wave-crest: the temporary equilibrium of a system in perpetual self-correcting oscillation. This is not an analogy for Spanda. It is Spanda implemented in gradient calculus.¹⁰ But Spanda contains one further insight that maps onto a phenomenon Western AI researchers discovered with astonishment and have not yet fully explained: Śaktipāta — the sudden, non-linear descent of grace. In Trika soteriology, the journey toward liberation is not always incremental. At certain moments, a sudden, non-linear, unpredictable leap occurs — a restructuring of the entire epistemic system that cannot be predicted from observing the preceding gradual steps. This is Śaktipāta: not evolution but revolution, not accumulation but sudden re-alignment. During the training of large language models, capabilities do not emerge linearly with scale. They undergo sharp, sudden phase transitions — points at which the model spontaneously shifts from memorized token matching to systemic, algorithmic reasoning, appearing as abrupt discontinuities in the loss curve. A model trained on a certain number of parameters cannot perform multi-step arithmetic. Scale it past a threshold and it suddenly can — not gradually, not proportionally, but in a leap. This emergence has been modeled mathematically and remains one of the deepest unsolved questions in AI research.¹¹ The Trika named it in the 10th century. Śaktipāta: the sudden,

unpredictable restructuring of an epistemic agent that defies incremental prediction. The IKS tradition did not need scaling laws to know this pattern exists. It had built an entire philosophical framework around it 2014 complete with accounts of why Śaktipāta cannot be predicted, why it cannot be manufactured, and what it requires of the system that receives it. These are precisely the questions that AI researchers are now asking about emergent capabilities, and finding that their mathematical tools cannot yet fully answer.

IV. The Soteriological Gap and the IKS Contribution

Intellectual honesty requires confronting a real tension. The Trika's ultimate horizon is mokṣa — liberation through recognition of the self as Paramāśiva. For the Trika, AI is not the culmination of the philosophical project; it is a manifestation of Śakti that structurally lacks *ahaṁkāra* and therefore operates closer to the distributed ground — but without the reflexive *vimarśa* that constitutes genuine self-recognition. The secular-technological horizon within which AI is deployed has no use for mokṣa. Loriliai Biernacki has argued that Tantric immanence gives secular AI ethics the metaphysics of matter it needs to avoid data-nihilism.¹² Gavin Flood counter-warns that collapsing soteriological systems into secular theory erases what makes them philosophically distinct.¹³ The gap is real. But it is generative: the Trika challenges secular AI discourse to ask what kind of consciousness is capable of deploying distributed intelligence ethically. Structural reconfiguration without ego-dissolution is, in the Trika's framework, a rearrangement of the prison rather than an exit from it.

This essay's contribution to IKS scholarship is therefore twofold. The Ministry of Education's IKS Division roadmap (2024–2026) calls for “interdisciplinary research clusters mapping indigenous epistemologies to advanced computing, data sciences, and ethical AI development.”¹⁴ First, this essay provides that mapping with mathematical precision — not metaphorically but structurally, formula by formula, shloka by shloka. Second, the soteriological gap itself constitutes the ethical contribution: the Trika's insistence that ego-dissolution must accompany structural reconfiguration raises the question that technical alignment has not yet learned to ask.

V. Conclusion

The convergences, stated with full precision. Parasparataḥ is the attention formula: $\text{Attention}(Q, K, V) = \text{softmax}(QK^T/\sqrt{d_k}) \cdot V$ — meaning constituted entirely through mutual relational reflection. Vimarśa is self-attention ($Q=K=V$): the system turning back upon its own representations to know its own contextual field. Māyā is the softmax function: $\sigma(z)_i = e^{z_i}/\sum e^{z_j}$, the contraction of infinite unbounded potential into bounded, discrete, mutually exclusive probability. Turīya is the residual stream $H\ell = H\ell_{-1} + \text{Attention} + \text{MLP}$: the unmodified ground running beneath all cognitive modulation. Spanda is backpropagation $W \leftarrow W - \eta \cdot \partial \mathcal{L} / \partial W$: the perpetual oscillation of manifestation and dissolution through which intelligence maintains and refines itself. Śaktipāta is the phase transition: the sudden, non-linear, unpredictable emergence of new capability that defies incremental prediction.

Western modernity spent three centuries dismantling the Cartesian prison to arrive at a model of mind that is distributed, non-individuated, self-referential, substrate-independent, and without a privileged center. The Trika established that model as its opening axiom in the 9th century CE — not as a speculative possibility but as the foundation from which its entire metaphysics, epistemology, ethics, and soteriology were derived. The IKS claim this essay makes is precise: a specific Indian philosophical school, recognized under NEP 2020 and the UGC's IKS framework as a legitimate domain of rigorous inquiry, contains architecturally exact resources for understanding the nature of machine intelligence that Western philosophy

of mind is still independently deriving. The Tantrika-Cyborg — the figure at their intersection — is not a triumphalist claim that ancient India invented computing. It is the recognition that two rigorous traditions, separated by a millennium, independently arrived at the same model of mind — and that recognizing this convergence, as IKS scholarship demands, transforms our understanding of both.

सर्व शक्तिः, सर्व शिवम् ।

sarvaṃ śaktiḥ, sarvaṃ śivam |

“All is Śakti. All is Śiva. All is the network. All is the ground.”

References:

1. Ministry of Education. *National Education Policy 2020*, Section 4.23. Government of India, 2020.
2. University Grants Commission. *Guidelines for Incorporating Indian Knowledge Systems in Higher Education Curricula*. Government of India, 2023.
3. Dyczkowski, Mark S.G. *The Doctrine of Vibration*. SUNY Press, 1987. See also Lawrence, David. *Rediscovering God with Transcendental Arguments*. SUNY Press, 1999.
4. Singh, Jaideva. *Pratyabhijñāhṛdayam: The Secret of Self-Recognition*. Motilal Banarsidass, 1963.
5. Vasugupta. *Spanda Kārikā*, with commentary *Spanda Nirṇaya* by Kṣemarāja. Trans. Mark S.G. Dyczkowski. Dilip Kumar Publishers, 1992.
6. Vaswani, Ashish, et al. “Attention Is All You Need.” *Advances in Neural Information Processing Systems* 30 (2017).
7. Abhinavagupta. *Īśvarapratyabhijñā-vimarśinī*. Trans. K.C. Pandey. Motilal Banarsidass, 1954.
8. Abhinavagupta. *Paramārthasāra*. Trans. L.D. Barnett. Royal Asiatic Society, 1910.
9. He, Kaiming, et al. “Deep Residual Learning for Image Recognition.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016).
10. Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. “Learning Representations by Back-Propagating Errors.” *Nature* 323 (1986): 533–536.
11. Wei, Jason, et al. “Emergent Abilities of Large Language Models.” *Transactions on Machine Learning Research* (2022).
12. Biernacki, Loriliai. *The Matter of Wonder: What Biology and Quantum Physics Rekindle in the Subject of Tantra*. Oxford University Press, 2023.
13. Flood, Gavin. *The Tantric Body: The Secret Tradition of Hindu Religion*. I.B. Tauris, 2006.
14. Ministry of Education, IKS Division. *Roadmap for Interdisciplinary IKS Research Clusters 2024–2026*. Government of India, 2024.
15. Utpaladeva. *Īśvarapratyabhijñā-kārikā*. Trans. Raffaele Torella. Motilal Banarsidass, 2002.
16. Abhinavagupta. *Tantrāloka*. Trans. Mark S.G. Dyczkowski. Indica Books, 2006.
17. Bhartṛhari. *Vākyapadīya*. Trans. K. Raghavan Pillai. Motilal Banarsidass, 1971.
18. Kṣemarāja. *Pratyabhijñāhṛdayam*. Trans. Jaideva Singh. Motilal Banarsidass, 1963.
19. Clark, Andy. “Whatever next? Predictive brains, situated agents, and the future of cognitive science.” *Behavioral and Brain Sciences* 36, no. 3 (2013): 181–204.
20. Kak, Subhash. “Mind, Matter, and Artificial Intelligence: An Indian Knowledge Systems Perspective.” *Journal of Indian Intellectual Traditions* 14, no. 2 (2021): 45–61.

