

Article

Not peer-reviewed version

---

# Maxwell and Artificial Intelligence: Preliminary QED Models of AI (LLM) Dynamics

---

[Yingrui Yang](#)\*

Posted Date: 6 May 2025

doi: 10.20944/preprints202504.2624.v1

Keywords: Maxwell field; AI; embedding; vector potential; dynamic phase; task charge; AI-demand; AI-supply; QED models; spinor field



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

# Maxwell and Artificial Intelligence: Preliminary QED Models of AI (LLM) Dynamics

Yingrui Yang

Department of Cognitive Science, Rensselaer Polytechnic Institute; yangyri@rpi.edu

**Abstract:** The present paper aims to set up preliminary conceptual structures of two electrodynamic models of new AI (LLM) dynamics, one is based on the interaction of a charge field with Maxwell vector potential field, and another is based on the interaction of a spinor field with Maxwell field. For this approach, it introduces a number of new concepts. The AI demand is defined as a two-components unit, a task charge and the corresponding intelligence. Neural network embeds the features of the task into a vector space, which is defined as Maxwell vector potential field. The gradient descent conducts the field strength. For spinor field, it makes the human intelligence and machine intelligence orthogonal; the former is the AI demand and latter the AI supply. The Maxwell field becomes the photon field, which reflects the cost function. The interactions between a pair of AI demand and AI supply and the cost function are pictured by the Feynman Diagram.

**Keywords:** Maxwell field; AI; embedding; vector potential; dynamic phase; task charge; AI-demand; AI-supply; QED models; spinor field

## 1. Maxwell Vector Potential

### 1.1. Consider a Verbal Task A for LLM

Assume A has a family of  $n$  features. Note that the task A may involve more than  $n$  features, and together these residues features are treated as gauge freedom later. In LLM, the transformer generates a weight for each of  $n$  features, and embeds A as vectors. This step is called embedding. Solving a task may involve many rounds of embeddings, each updates the vector space to certain extensions.

Definition 1. As an LLM procedure, the transformer embeds a verbal task to a vector space, denoted by  $A_\mu$ . Call  $A_\mu$  the Maxwell vector potential; it is also called the Maxwell field. In other words, a verbal task brings in the external potential.  $A_\mu$  can be characterized by exterior form as 1-form, which stands for  $f(x)dx$ .

In LLM, gradient descent is a major technique for deep learning. The gradient,  $F_{\mu\nu}$ , can be characterized as 2-form in terms of the exterior form. The relation between  $A_\mu$  and  $F_{\mu\nu}$  is as below.

Definition 2. The gradient of  $A_\mu$ , denoted by  $F_{\mu\nu}$ . We have,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

In gauge field theory,  $F_{\mu\nu}$  stands for the gauge field strength; it is a density function.

There are three paths to build QED models for LLM. [1] The first path is by interactions of a charge field with Maxwell field. The second path is by the interactions of a spinor field (Dirac field) with Maxwell field. The third path is by the interactions of Klein-Gordon field with Maxwell field. In this theoretical note, we only consider the first two paths, i.e., we follow the first path and the second path to build QED models for LLM framework.

### 1.2. From the Angle to the Dynamic Phase

In LLM, the inner product is introduced in the vector space. Thus, from one vector can produce another vector and the superposition of two vectors. Between any two vectors, there is an angle which is currently called the distance. This is exactly the illusion effected H. Wyle that delated him from establishing gauge field theoretic dynamics. To correct it, we have,

Definition 3. The angle between two potential vectors is defined as the dynamic phase. Thus, from one potential state  $\psi$  to another potential state  $\psi'$  can be represented as

$$\psi \rightarrow \psi' = e^{i\theta} \psi$$

where  $\theta$  stands for dynamic phase. This is the key move from LLM to AI dynamics. This is called the gauge transformation of the first kind at the global level. This gauge transformation also requires keeping the transformation conformal at the local level, where  $\theta$  is no longer an any given constant but a function  $\theta(x)$ . The gauge transformation at the local level is called the gauge transformation of the second kind. The gauge principle says that if the gauge transformation of the first kind does not hold, then the gauge transformation of the second kind can not be hold. In order to achieve the local gauge symmetry, it needs to introduce new notions, namely, covariate derivative and gauge field. We will develop this point latter.

## 2. Charge Field Interacts with Maxwell Potential

Dynamic analysis is the sourced analysis, and this source is called the charge. We have

Postulate 1. The LLM tasks carry the task charge, denoted as  $E$ . A task can be characterized by a family of features. By Anne Treisman [2], the integration of these features demands attention. Call this integration the charge potential and denote it as  $\phi$ .

To solve a task demands some intentions and solving a task demands a number of intelligence factors, such as reasoning, learning, understanding, integrating, etc. We have

Postulate 2. Let  $A_i$  are intelligence factors. The integration of  $(A_i)$  is defined as the intelligence potential, denoted by  $\vec{A}$ .

Definition 3. The notion of *AI vector potential* is defined as a unit with two components,

$$A_{AI} = [\phi, -\vec{A}]$$

This idea can be seen as a Platonic version of Maxwell vector potential; it is a field. In quantum field theory, the charge is characterized as a field. Here it is called the task-charge field. denoted as TCF. The TCF interacts with  $A_{AI}$  yields a dynamic model, logically skin to that an electric field interacts with the Maxwell potential field yields quantum electrodynamics (QED).

## 3. Spinor Field Interacts with Maxwell Potential

### 3.1. Wyle Spinor

We make a clear distinction of the human intelligence ( $\alpha$ ) and the machine intelligence ( $\beta$ ). Further, we make a mathematical trivialization treatment such that  $\alpha$  and  $\beta$  are orthogonal:

$$\delta_{\alpha\beta} = \begin{cases} 1, & \alpha = \beta \\ 0, & \alpha \neq \beta \end{cases}$$

Now we define  $\alpha$  as the intelligence demand ( $D_\alpha$ ) and  $\beta$  as the intelligence supply ( $S_\beta$ ). For an any given task  $\varphi$ , we have

Definition 5.  $e^- = [D_\alpha, \varphi^-]$ , and  $e^+ = [S_\beta, \varphi^+]$ .

The above definitions can be characterized as a Wyle spinor below,

$$\psi = \begin{pmatrix} \psi_1 (= e^-) \\ \psi_2 (= e^+) \end{pmatrix}$$

### 3.2. Dirac Spinor

In the neural network for deep learning, the hidden layers indicate that both human intelligence and machine intelligence hesitate. This kind of intellectual hesitation reflects a certain kind of rationality. Such a hesitation is characterized as spin, which has two eigenstates, spin up ( $\uparrow$ ) or else spin down ( $\downarrow$ ); thus, in a sense, both  $e^-$  and  $e^+$  spin  $\frac{1}{2}$ . This can be seen at the activating stage in the architecture of neural network, which returns only 0 or 1. We need hidden layers to calibrate and update neural weights (probabilities), during which the intelligent processes experience all possible superpositions of the two eigenstates. Taking spins into account, the Wyle spinor is developed into Dirac spinor, represented as below,

$$\psi = \begin{pmatrix} \psi_1 (= e^- \uparrow) \\ \psi_2 (= e^- \downarrow) \\ \psi_3 (= e^+ \uparrow) \\ \psi_4 (= e^+ \downarrow) \end{pmatrix}$$

### 3.3. Photon Field

In quantum field theory, the Maxwell field is the photon field, which is denoted by  $\lambda$ . In market dynamics [3], the photon field is used to model the price field. Here, in LLM, the photon field is used to model the cost function of neural network. In quantum field theory, all particles are described as a field, and a field is characterized as an operator. In addition, all the particles have its life span, they are created and annihilated. We need to introduce the creation operators here, the creation operator ( $\hat{a}^\dagger$ ) and the annihilation operator ( $\hat{a}$ ). They are adjoint operators. A simple interaction is that a pair of  $e^-$  and  $e^+$  are created and they fly in inverse directions to each other. Then, both are annihilated by creating a photon field  $\lambda$ . This process can be pictured by Feynman diagram shown in Figures 1,

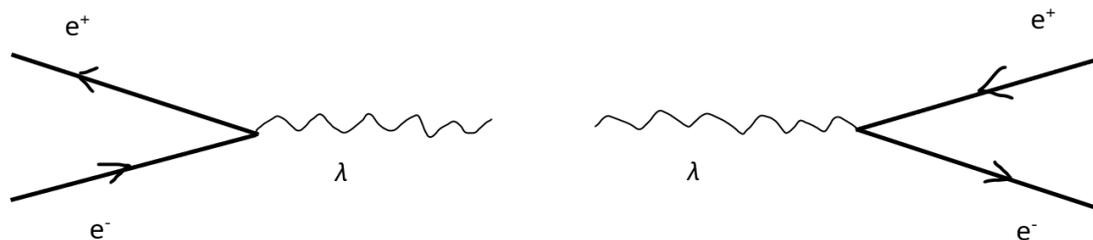


Figure 1a.  $e^- + e^+ \rightarrow \lambda$

Figure 1b.  $\lambda \rightarrow e^- + e^+$

One may start from here to develop a detailed model of LLM skin to spinor electrodynamics.

## 4. Wavefunction and Measurement Problem

The major processes in LLM learning/training are not directly observable; thus, it should be characterized as the U-procedure (Unitary procedure) of a wavefunction [4]. Hence, the probability used in LLMs should be Born probability. Note that at the activating stage of the neural network, the feature weights are converted to 0 or 1; this is actually a Yes/No measurement [5], which is called the R-procedure (reduction procedure). This reflects the measurement problem in quantum physics. Yang [6] proposed a unified account of the U-Procedure and the R-Procedure as an alternative solution of the measurement problem.

## References

1. Wang, Z. (2008). Elementary Quantum Field Theory. Peking University Press (China).
2. Treisman, A. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology*, 12, 97-136.
3. Yang, Y. (2022). Principles of Market Dynamics: Economic Dynamics and Standard Model (II). *Science, Economics, and Society*, Vol 40, No. 5 (China).
4. Penrose, R. (2005). The Road to Reality: A Complete Guide to the Laws of the Universe. Random House, Inc., New York.
5. von Neumann, J. (1955). Mathematical Foundations of Quantum Mechanics. Princeton University Press, Princeton, New Jersey.
6. Yang, Y. (2024). The revised Schrödinger equation as a solution of measurement paradox: A unified model of the U-procedure and the R-procedure. <https://doi.org/10.20944/preprints202402.0013.v2>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.