# 楊-米爾斯理論的源起、內容與發展

109022115 LI,QIAN-RUI (李謙睿)
*Physics Department, National Tsing Hua Univsersity*
(Dated: July 1, 2023)

本文是近代物理發展史的期末報告，我將基於楊振寧與羅伯特·勞倫斯·米爾斯 (Robert Laurence Mills) 在 1954 年發表的論文《Conservation of Isotopic Spin and Isotopic Gauge Invariance》，對楊-米爾斯理論的整個發展過程、其所處時代下的物理背景與前因後果進行簡單的統整和描述，包括了當時物理的發展方向、規範不變性與諾特定理的詳細概念、同位旋與 SU(2) 群的簡單介紹、楊-米爾斯理論的核心與其數學過程、後續對標準模型的影響等。然而作為一個大三的學生，由於對許多知識十分陌生，因此我將從比較基礎的部分把數學過程推演一遍，也許在教授眼中會略顯繁複冗長，但主要是為了讓我自己能更好的釐清觀念並把整個故事說清楚，這些過程也能同時作為我學習的一個紀錄。楊-米爾斯理論的數學框架影響了近代粒子、高能物理的發展，是十分重要的理論，我覺得作為一個物理系的學生花時間去理解該理論是十分值得的。

## I. 背景與概念介紹

### A. 物理學的發展

物理學的發展主線之一便是希望能夠用最簡潔與優美的方式來描述所有物理現象，把種種不同的現象用更深層次與接近本質的理論加以聯繫與統一。牛頓統一了地上與天上的力並告訴世人蘋果的掉落與行星和恆星間的運動皆遵循同一法則，馬克士威則統一了電與磁並進一步揭示了光的本質。因此到了 19 世紀末，人們可以把所有已知的物理現象歸結於引力或是電磁力，然而牛頓的重力理論框架確與馬克士威的電磁學卻有一些結構上根本的矛盾，但該問題很快的被愛因斯坦發現並被解決。愛因斯坦根據「相對性原理」，也就是物理定律應當適用於不同座標系，與「光速恆定」的事實提出了狹義相對論，其體現的是物理定律在勞倫茲變換下的不變性。電磁學很好的適應了狹義相對論的框架，而牛頓的各種力學原理則需要加以修改才能適應狹義相對論的框架，但對於重力，引力的本質狹義相對論卻無法良好的描述，也正是因為如此愛因斯坦接著提出了著名的廣義相對論。

在廣義相對論的時空觀下電磁學與重力理論不再有矛盾，因此似乎下一步就應該要能把電磁力與重力統一成一種力並誕生所謂的統一場論了! 然而愛因斯坦最終並沒有成功完成該工作，並且至今該問題仍是十分有挑戰性的課題。而在愛因斯坦提出廣義相對論之後，隨著人們科技與物理的持續發展並對原子核有更多了解後又冒出了兩個全新的力—弱力與強力。這下物理學變的更加複雜了，不但沒統一重力與電磁力，現在又冒出兩個連該如何描述都不太確定的力，那更不用說要統一這四個力了。面對如何描述弱力與強力的難題，楊-米爾斯的理論展現了其關鍵的地位。首先重力使用廣義相對論描述、電磁學使用馬克士威方程描述，而量子化後的電磁學可以用量子電動力學描述。如今我們知道強力便是使用楊-米爾斯理論來描述的，而弱力在與電磁力完全的統一後也是用楊-米爾斯理論描述，可見楊-米爾斯規範場理論的架構對後世物理發展的重要性。雖然楊-米爾斯的理論本身並沒有告訴我們強力和弱電力具體該怎麼樣，但該理論做為一個精妙的模型提供給後世科學家在研究強力和弱電力時的方向與方法，這就是楊-米爾斯的理論真正的價值。

### B. 物理學的研究方法

物理學的發展過程一直以來大部分都是先做實驗，然後物理學家根據實驗結果的數據來尋找隱藏在期中的規律與道理，並最終用數學來總結與描述這些規則和定律，直到有一個人將這個過程翻轉了，這個人便是愛因斯坦。

重力理論的發展最早從開普勒藉由第谷眾多的觀察資料中歸納出了行星的三大運動定律，爾後牛頓再從其中慢慢推敲出引力和距離平方反比的關係，然而愛因斯坦的重力理論，也就是廣義相對論，卻是難以從當時的觀察資料中推敲出來的。可以看到當一個現象的背後的機制過於複雜時，我們是難以靠實驗與觀察的結果去歸納出理論的，愛因斯坦試圖從基本的原理與假設上出發，在要求物理定律在廣義座標中的不變性下，根據物理的直覺與邏輯推理最終得到的廣義相對論，其預言的許多現象在當時甚至許多物理學家連想都沒想到過，更不用說有相關的實驗能按圖索驥。

至此，物理學的發展有一個新的方式，也就是先從某種對稱性出發並藉由數學根據其基本性質推導或湊出方程式，實驗則從原來的歸納理論的基礎變成驗證理論的工具，而某種對稱與不變性也從理論的副產品變成了主宰理論的核心。正是在這樣嶄新的物理發展脈絡下才造就了楊-米爾斯理論的誕生。

### C. 規範不變性與諾特定理

所謂的規範不變性就是在規範變換下物理定律保持不變的特性，那麼規範變換又是什麼呢？我覺得可以把規範變換理解成一種對物理公式中某一變量加入新的自由度，並同時保持物理定律不被改變的一種數學操作。更實際的來說，決定一個物理定律與在此物理定律下物體的運動法則的東西，根據最小作用量原理，這個東西就是作用量 (action)，而決定作用量的是拉格朗日量 L(Lagrangian)，一般寫作

$$S = \int L dt \tag{1}$$

但在相對論場論中一個場是時空的函數，為了更廣泛的考慮一個場在時空中的物理行為，通常會使用拉格朗日密度 $\mathscr{L}$(Lagrangian density) 來描述定域場的作用量，此時拉格朗日密度 $\mathscr{L}$ 為場 $\phi$ 與場的微分項 $\partial_\mu \phi$ 的函數

$$S = \int_a^b \mathscr{L}\ dx\ dy\ dz\ d(ct)$$
$$= \int_a^b \mathscr{L}(\phi, \partial_\mu \phi)\ d^4x \tag{2}$$

因此現在可以清楚的說明所謂的規範不變性在場論中表示的便是對朗日密度 $\mathscr{L}$ 中的場進行變換卻又不改變 $\mathscr{L}$ 歐拉-拉格朗日方程結果、也就是不改變運動方程的操作。而諾特 (Noether)，一個傑出的女性物理學家，告訴了世人這樣操作性質背後的重要意義，也就是一個變換下不變的性質，或是說一個變換下的對稱性會對應到一個守恆量，這就是所謂的諾特定理。讓我們從最小作用量原理出發，當場 $\phi$ 任意改變 $\delta\phi$ 時，作用量也將發生改變

$$\delta S = \int_a^b d^4x \left[\frac{\partial \mathscr{L}}{\partial \phi}\delta\phi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta(\partial_\mu\phi)\right] \tag{3}$$

由於這裡不考慮座標本身有變化，也就是 $\delta(d^4x) = 0$，所以有

$$\delta(\partial_\mu \phi) = \partial_\mu(\delta\phi) \tag{4}$$

透過微積分的 product rule，作用量的變分可改寫為

$$\delta S = \int_a^b d^4x \left[\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right)\right]\delta\phi$$
$$+ \int_a^b d^4x\ \partial_\mu\left[\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta\phi\right] \tag{5}$$

可以看到第二項可藉由高斯定律把體積分換成面積分，這邊的體積是四維的體積，面積則為三維的面積，得

$$\delta S = \int_a^b d^4x \left[\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right)\right]\delta\phi$$
$$+ \oint_\sigma d\sigma\ \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta\phi \tag{6}$$

期中 $\sigma$ 是邊界的表面，對於「任意」擾動 $\delta\phi$ 我們要求其在邊界為零，也就是作用量在邊界的擾動為零，因此第二項對邊界的面積分為零，在套用最小作用量原理後得

$$\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right) = 0 \tag{7}$$

這便是經典的歐拉-拉格朗日方程，它描述了場的運動方程。從以上步驟可看到邊界項由於散度定理，總是可以在計算作用量的積分過程中視為零，因此不妨對 $\mathscr{L}$ 加入一個任意項，該項同樣為某個向量 $\Lambda^\mu(\phi, \partial_\mu\phi)$ 的散度，在平坦的時空下在如此變化儘管拉氏量不同但不改變運動方程、我們稱之為正則變換。

$$\mathscr{L}' = \mathscr{L} + \partial_\mu \Lambda^\mu \tag{8}$$

接下來為了找到諾特定理中的守恆量與對稱變換之間的關係，讓我們考慮在某種對稱的變換下，座標和場的變分為 $\delta x^\mu = x'^\mu - x^\mu$ 和 $\delta\phi = \phi'(x') - \phi(x)$，作用量的變分為

$$\delta S = \int_a^b (d^4x')\mathscr{L}'(x') - \int_a^b (d^4x)\mathscr{L}(x)$$
$$= \int_a^b \left(d^4x\left|\frac{\partial x'^\mu}{\partial x^\nu}\right|[\mathscr{L}(x) + \delta\mathscr{L}] - (d^4x)\mathscr{L}(x)\right)$$
$$= \int_a^b [d^4x + \delta(d^4x)][\mathscr{L}(x) + \delta\mathscr{L}] - (d^4x)\mathscr{L}(x)$$
$$= \int_a^b [\delta(d^4x)\mathscr{L} + d^4x(\delta\mathscr{L})] + O((\delta x)^2) \tag{9}$$

在考慮到 $\delta x$ 一次項下，積分座標的變換為

$$\delta(d^4x) = d^4x\left|\frac{\partial x'^\mu}{\partial x^\nu}\right| - d^4x$$
$$= d^4x\left|\delta_\nu^{\ \mu} + \partial_\nu(\delta x^\mu)\right| - d^4x$$
$$= d^4x(\partial_\mu \delta x^\mu) \tag{10}$$

拉格朗日密度 $\mathscr{L}$ 的變分除了本身的變化 $\mathscr{L} \to \mathscr{L}'$ 外，也有來自座標變換的變化 $\mathscr{L}'(x) \to \mathscr{L}'(x')$，因此考慮 $\delta = \delta_0 + \delta x^\mu \partial_\mu$，得

$$\delta\mathscr{L} = \mathscr{L}'(x') - \mathscr{L}(x)$$
$$= \delta x^\mu \partial_\mu \mathscr{L} + \delta_0 \mathscr{L}$$
$$= \delta x^\mu \partial_\mu \mathscr{L} + \left[\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right)\right]\delta_0\phi + \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta_0\phi\right)$$

又由於要求變換下不改變該系統物理法則，因此歐拉-拉格朗日方程成立下拉格朗日密度 $\mathscr{L}$ 的變分為

$$\delta\mathscr{L} = \delta x^\mu \partial_\mu \mathscr{L} + \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta_0\phi\right) \tag{11}$$

將 (11) 式和 (10) 式帶入 (9) 式可得

$$\delta S = \int_a^b d^4x\left[(\partial_\mu \delta x^\mu)\mathscr{L} + \delta x^\mu(\partial_\mu \mathscr{L}) + \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta_0\phi\right)\right]$$
$$= \int_a^b d^4x\ \partial_\mu\left[\delta x^\mu \mathscr{L} + \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta_0\phi\right)\right]$$

運用 $\delta = \delta_0 + \delta x^\rho \partial_\rho$ 關係將 $\delta_0$ 替換成 $\delta - \delta x^\rho \partial_\rho$ 後可得

$$\delta S = \int_a^b d^4x\ \partial_\mu\left[\delta x^\mu \mathscr{L} + \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}(\delta - \delta x^\rho \partial_\rho)\phi\right)\right]$$
$$= \int_a^b d^4x\ \partial_\mu\left[\left(\frac{\delta x^\mu}{\delta x^\rho}\mathscr{L} - \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\partial_\rho\phi\right)\delta x^\rho + \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right)\delta\phi\right]$$

最後考慮對稱的變換是連續的變換，可以透過一組連續的變換參數 $\omega^A$ 表達

$$\delta x^\rho = \frac{\delta x^\rho}{\delta \omega^A}\delta\omega^A \tag{12}$$

$$\delta \phi = \frac{\delta \phi}{\delta \omega^A}\delta\omega^A \tag{13}$$

最終得到

$$\begin{aligned}\delta S = &\int_a^b d^4x\, \partial_\mu \left[\left(\frac{\delta x^\mu}{\delta x^\rho}\mathscr{L} - \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\partial_\rho\phi\right)\frac{\delta x^\rho}{\delta\omega^A}\right]\delta\omega^A \\ &+ \int_a^b d^4x\, \partial_\mu \left[\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right)\frac{\delta \phi}{\delta\omega^A}\right]\delta\omega^A\end{aligned}$$

若對於所有 $\delta\omega^A$，$\delta S = 0$，則可找到對應的守恆流 $\partial_\mu j^\mu_A = 0$，守恆流 $j^\mu_A$ 為

$$j^\mu_A = -\left[\left(\frac{\delta x^\mu}{\delta x^\rho}\mathscr{L} - \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\partial_\rho\phi\right)\frac{\delta x^\rho}{\delta\omega^A} + \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right)\frac{\delta \phi}{\delta\omega^A}\right] \tag{14}$$

以上便是場論中的諾特定理，只要作用量在對稱變換下保持運動方程的不變、歐拉-拉格朗日方程保持成立，就會導致對應的守恆流 $j^\mu_A$，額外加入的負號是為了確保能量密度大於零。由於拉氏密度 $\mathscr{L}$ 在 (8) 式的正則變換下不改變作用量，因此守恆流的空間項會被影響而不唯一，但因為我們考慮的場 $\phi$ 在無限遠處是零或是滿足特定邊界條件，因此守恆流空間項的部分對作用量的貢獻為零，所以 $\partial_\mu j^\mu_A = 0$ 對全空間的積分後可得不隨時間變化的守恆荷 $Q_A$

$$Q_A = \int_{-\infty}^{\infty} d^3x\, j^0_A \tag{15}$$

有了以上的成果，我們可以在已知的理論中根據已知的守恆量去尋找相關的對稱性，例如電磁學中的電荷守恆對應到的對稱性是整體場的相位變換不變性。詳細來說，考慮能夠描述自旋 1/2 粒子的狄拉克方程，其對應的拉氏量 $\mathscr{L}$ 為 (令 $\hbar = 1, c = 1$)

$$\mathscr{L} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi \tag{16}$$

其中的 $\psi$ 是「一組」波函數作為元素構成的向量 (有兩個波函數，每個波函數用 2×1 的向量構成，所以總共有四個分量)，通常稱之為旋量 (spinor)，由於這裡要使用量子場論的概念，因此應該稱呼 $\psi$ 為一個旋量場，狄拉克方程則是描述了一個自旋 1/2 的場方程。可以很容易發現若是對 $\psi$ 進行全體的相位變換 $\psi \to \psi' = e^{i\theta}\psi$ 並且 $\overline{\psi} \to \overline{\psi'} = e^{-i\theta}\overline{\psi}$ 是不改變拉氏量 $\mathscr{L}$ 的，此時變換參數 $\omega^A = \theta$，其對應的守恆流為

$$j^\mu_\theta = \overline{\psi}\gamma^\mu\psi \tag{17}$$

守恆荷便是我們熟悉的總電荷，更廣義的說是費米子的總數，此詮釋可從對狄拉克場的量子化中得到。

$$Q_\theta = \int_{-\infty}^{\infty} d^3x\, \psi^\dagger\psi \tag{18}$$

現在有件有趣的事情，那就是如果我們將對場整體、全域的變換不變性推廣到侷域，並要求侷域的變換也有不變性，那麼結果會如何呢? 引發楊米爾斯理論的源頭便是來自於全域不變性推廣到侷域不變性所帶來的驚人結果。

### D. 全域對稱與侷域對稱

若把全域 (global) 的變換解釋為對一個團體進行全體一致的變換，那麼侷域 (local) 的變換就是一個群體中每個成員按各自的步調進行變換，彼此之間是不一致的。如果有一物體能夠經過侷域的變換仍保持不變的話，那麼該物體就是具備侷域對稱性的。回到對作用量與拉氏量的討論上，全域的變換就是所有的變換與時空座標無關，侷域變換則是變換本身是一個與時空座標有關的函數。一樣從電子的角度出發，從對描述電子的狄拉克場的全域相位變換的規範不變性中我們得知了電荷守恆的結果，那麼若是要求在侷域變換下也有規範不變性的話，我們是否會得到更多物理呢?

現在同樣使用狄拉克的拉氏量 (令 $\hbar = 1, c = 1$)

$$\mathscr{L} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi$$

如果旋量場 $\psi$ 的相位因子在不同時空點是不同的，即侷域的相位變換 $\psi \to e^{i\theta(x^\mu)}\psi$，那麼拉氏量能在該變換下還保持不變嗎？答案是否定的，因為對場進行時空的求導時顯然會多一項，即

$$\partial_\mu(e^{i\theta(x^\mu)}\psi) = i(\partial_\mu\theta)e^{i\theta}\psi + e^{i\theta}\partial_\mu\psi \tag{19}$$

因此

$$\begin{aligned}\mathscr{L}' &= \overline{\psi'}(i\gamma^\mu\partial_\mu - m)\psi' \\ &= e^{-i\theta(x^\mu)}\overline{\psi}(i\gamma^\mu\partial_\mu - m)e^{i\theta(x^\mu)}\psi \\ &= \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi - (\partial_\mu\theta)\overline{\psi}\gamma^\mu\psi \\ &= \mathscr{L} - (\partial_\mu\theta)\overline{\psi}\gamma^\mu\psi\end{aligned} \tag{20}$$

若我們令 $\lambda(x^\mu) = -\theta(x^\mu)/q$，則可以把上式改寫成

$$\mathscr{L} \to \mathscr{L} + (q\overline{\psi}\gamma^\mu\psi)\partial_\mu\lambda \tag{21}$$

其中 q 代表的是電荷。現在若我們希望拉氏量要在侷域變換下是不變的話，那我們就要想盡辦法把 (21) 式中多出的那一項給消掉，而為了達到此目的我們不得不在原本的拉氏量上加上一些額外的項，假設修改的拉氏量為

$$\mathscr{L} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi - (q\overline{\psi}\gamma^\mu\psi)A_\mu \tag{22}$$

其中 $A_\mu$ 是某個新的向量場，隨著 $\psi$ 的侷域相位變換按照以下的規則變換

$$A_\mu \rightarrow A_\mu + \partial_\mu \lambda \tag{23}$$

如此一來 $A_\mu$ 變換時的 $\partial_\mu \lambda$ 項彌補了 (21) 式中多出的項，如此一來便能保持拉氏量在侷域變換下是保持不變的。然而我們為了達到目的而多加進的這一項還無法說清楚整件事，因為可以看到我們多加進的這一項是向量場 $A_{mu}$ 與原本的旋量場 $\overline{\psi}\psi$ 的耦合和交互作用的項，那麼就如同 $\overline{\psi}\psi$ 有自己獨立的項一樣，向量場 $A_\mu$ 理應也要有自己獨立的項才說的通。我們不妨來猜猜看 $A_\mu$ 獨立的項該長什麼樣子，由於相對性原理拉氏量應當是一個勞倫茲不變量 (純量)，因此最簡單的猜測為

$$A^\mu A_\mu \tag{24}$$

但可惜的是在進行 (23) 式的侷域變換時它並不會保持不變，而為了同樣不要破壞侷域的對稱性，實際上為了保持侷域的規範不變性 $A_\mu$ 的獨立項長

$$(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) \equiv F^{\mu\nu}F_{\mu\nu} \tag{25}$$

到此為止，我們已經找到了 $A_\mu$ 場在拉氏量中應該呈現的樣子。我們不妨透過後見之明，將我們得到的結果與描述自旋均為為 1、質量為 m 的 Proca 場的作用量進行比較 (令 $\hbar = 1, c = 1$)

$$\mathscr{L} = \frac{-1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}m^2 A^\mu A_\mu \tag{26}$$

從前面的討論中已經知道 $A^\mu A_\mu$ 項因為不符侷域的規範不變性而不應該存在，所以這就暗示了我們 $A^\mu$ 場描述的粒子是無質量的 (m=0)。如此一來完整的考慮符合侷域規範不變性的拉氏量為

$$\mathscr{L} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi - \left[\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}\right] - (q\overline{\psi}\gamma^\mu\psi)A_\mu \tag{27}$$

之所以說這是後見之明的原因是因為其實我們其實早就知道 $A^\mu$ 場不是別人就是電磁學中的電磁勢。所以看看在要求描述電子的狄拉克場服從侷域規範不變性的條件下，我們直接得到了量子電動力學的拉氏量，描述光子的 $F^{\mu\nu}F_{\mu\nu}$ 項和描述光子與電子交互作用的 $(q\overline{\psi}\gamma^\mu\psi)A_\mu$ 項代入歐拉-拉格朗日方程後便可得到馬克士威的電磁場方程，此時的電流密度為

$$J^\mu = q(\overline{\psi}\gamma^\mu\psi) \tag{28}$$

若是再引入規範協變求導的概念，也就是當我們面對 (19) 式中全域和侷域求導的差別時

$$\partial_\mu \psi \rightarrow e^{-iq\lambda}[\partial_\mu - iq(\partial_\mu \lambda)]\psi \tag{29}$$

把求導 $(\partial_\mu)$ 換成規範協變求導 $(D_\mu)$

$$D_\mu = \partial_\mu + iqA_\mu \tag{30}$$

如此便可抵消掉侷域變換時 (29) 式中多出的 $-iq(\partial_\mu \lambda)$ 使變換前後

$$D_\mu \psi \rightarrow e^{-iq\lambda} D_\mu \psi \tag{31}$$

如此一來完整的拉氏量可優雅的寫成

$$\mathscr{L} = \overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \left[\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}\right] \tag{32}$$

我們將其稱呼為滿足 U(1) 群的全域和侷域規範不變性，U(1) 的意思是把 $e^{i\theta}$ 視為 **U**nitary 且大小為 **1×1** 的矩陣，其性質滿足群的四個公理所以稱之為 U(1) 群。雖然我們早就知道馬克士威的電磁場方程，乍看之下好像沒給出什麼新東西，不過是用一些特別的方法在電磁理論中翻來覆去而已，然而事實上這卻是突破性的成就，因為我們僅僅透過要求電子的自旋場從全域的 U(1) 對稱進一步到侷域的 U(1) 對稱性便得到了電子與光子的交互關係，進而得到了整個電磁理論。這顯示了對稱性決定了相互作用，楊振寧正是受此啟發並將之推廣才有了之後的楊-米爾斯理論。現在我們離進到核心主題只剩一步之遙了，剩下的問題是楊振寧當年想把透過規範對稱性找到新的描述交互作用的這套方法用在哪呢？那麼首先要找到這個對稱性當然還得靠諾特定理，也就是去尋找基本交互作用中有沒有什麼特別的守恆量，這就是我們下一步的工作。

### E. 同位旋 (Isospin) 與 SU(2) 群

《Conservation of Isotopic Spin and Isotopic Gauge Invariance》論文一開始便在談論同位旋的概念，該概念源自海森堡在 1932 年的想法，主要是為了解釋為什麼質子和中子如此相像的原因，除了電性不同和質量上差了那麼一點點外質子和中子的各種特性可以說是幾乎一模一樣，尤其是在強交互作用下更是完全一模一樣，質子質子、中子質子、中子中子之間的強相互作用是沒有任何差別的。因此海森堡就假設中子和質子應該是「同一個東西」的兩種「不相同的態」，而這兩種態個藉由在一個抽象的「同位旋空間」中的旋轉互相轉換。由於在強相互作用中這兩種不同的態是完全相同的，也就是對於強相互作用在「同位旋空間」中的旋轉變換下是不變的，也就是強相互作用中同位旋守恆。所以就如同在電磁學中做的事情，首先要找到能夠描述對稱的變換方式，在電磁學中是 U(1) 的變換，而在「同位旋空間」中的「同位旋二重態」之間的轉換就如同自旋一樣是符合 SU(2) 群的變換，SU(2) 特殊么正群的 U(2) 的意思與 U(1) 一樣，代表 **U**nitary 且大小為 **2×2** 的矩陣。而 S 為特殊的意思，代表的是這些 U(2) 的矩陣的行列式值都要是

1。在找到變換型式後的下一步就是要求除了全域變換下規範不變外也要符合侷域的規範不變性，換句話說就是這個 SU(2) 的變換是時空的函數，在不同時間與空間的變換都不一樣。在電磁學中為了滿足 U(1) 侷域規範不變性而引入了一個向量場 $A_\mu$，期變換符合

$$A'_\mu = A_\mu + \partial_\mu \lambda$$

那麼相似的應該也存在著一個場 $B_\mu$，其變換規則能夠抵銷掉 SU(2) 侷域變換下多出的項，使之保持規範不變性並用 $B_\mu$ 建構出對應的 $F_{\mu\nu}$。然而這部分原始論文中並沒有詳細的數學過程而是透過文字描述並直接給出幾個階段的結果，實際上這一部分的數學可沒那麼簡單，在電磁理論中的 U(1) 群是阿貝爾群，代表其群乘法作用順序是不影響結果的，也就是 $e^{i\alpha}e^{i\beta} = e^{i\beta}e^{i\alpha}$，然而 SU(2) 群中這些 2×2 的矩陣們就沒這種好性質了，這種不滿足交換率的群就是非阿貝爾群，也就是因為 SU(2) 的非阿貝爾群性質使得這一過程複雜了許多。楊振寧得力於其父親的影響，在 1950 年代對群論相關的數學知識可說是超越了當時大部分的物理學家，這部分的數學對他自然不是問題，但顯然這對我是一個問題，因此下部分我想從頭開始把詳細的過程理一遍。

## II. 核心概念分析與討論

如同上部分提到的，我們想知道的是質子和中子在強力的視角下，同位旋空間中旋轉變換的對稱性、或是說 SU(2) 非阿貝爾群的規範不變性。由於這是一個討論兩個自旋皆為 1/2 粒子 (質子或中子) 的問題，因此先假設有兩個沒有交互作用且自旋皆為 1/2 的旋量場 $\psi_1$ 和 $\psi_2$，根據狄拉克方程其拉氏量可寫成 (令 $\hbar = 1, c = 1$)

$$\begin{aligned}\mathscr{L} &= [i\overline{\psi_1}\gamma^\mu\partial_\mu\psi_1 - m_1\overline{\psi_1}\psi_1] \\ &\quad + [i\overline{\psi_2}\gamma^\mu\partial_\mu\psi_2 - m_2\overline{\psi_2}\psi_2]\end{aligned} \tag{33}$$

以上式子可以寫成以下形式

$$\mathscr{L} = [\overline{\psi_1}, \overline{\psi_2}]\,[i\gamma^\mu\partial_\mu - \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}]\begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \tag{34}$$

若是令一個新的旋量場 $\psi$ 為

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \tag{35}$$

該旋量場 $\psi$ 中的 $\psi_1$ 和 $\psi_2$ 各自有 4 個元素，並在經過 $\partial_\mu$ 作用後各自都跟 $\gamma^\mu$ 的四個元素縮併 ($\gamma^\mu$ 的每個元素都是 4×4 的矩陣)，如果又同時假設兩個粒子質量相同，則可把拉氏量寫成

$$\mathscr{L} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi \tag{36}$$

可以看到雖然 (36) 式與原本只考慮一個自旋 1/2 場的拉氏量幾乎一模一樣，但實際上內部結構卻大不相同，(36) 式中的 $\psi$ 就像俄羅斯娃娃般套了三層向量總共有八個元素，其中以四個元素拆成兩份各自描述不同粒子，之所以搞的這麼麻煩就是因為同位旋的概念中質子和中子是同位旋二重態，就像自旋上和自旋下般要用一個 $\psi$ 來打包 $\psi_1$ 和 $\psi_2$ 並描述這兩個粒子在同位旋空間中形成的某個態。在此基礎上我們將 SU(2) 群中的幺正矩陣 U 作用在 $\psi$ 上對其進行變換，由於幺正矩陣的特性我們可以得知

$$\overline{\psi'}\psi' = (\overline{\psi}U^\dagger)(U\psi) = \overline{\psi}\psi \tag{37}$$

因此 $\overline{\psi}\psi$ 在變換下是不變的，也就是在 SU(2) 全域變換下拉氏量是不變的。現在如同在電磁學中做的一樣要把全域規範不變推展到侷域的規範不變，這裡首先要將特殊幺正矩陣 U 換一種方式寫法，SU(2) 群的生成元維度為三，其生成元正是三個包立矩陣 $\sigma_1$、$\sigma_2$、$\sigma_3$，由三個包立矩陣作為元素構成的向量記做 $\sigma$，又考慮一個由三個任意實數夠成的向量 a，滿足 SU(2) 的矩陣可寫成

$$U = e^{i\sigma\cdot a} \tag{38}$$

為了推廣到侷域的規範不變性，令向量 a 是一個與時空相關的參數，並比照之前的作法令一個參數 $\lambda(x^\mu) = -a(x^\mu)/q$，其中 q 是類似電荷的某個常數，$\lambda$ 是一個三個分量的向量，如此一來侷域的 SU(2) 變換可寫成

$$\psi' = S\psi = e^{-iq\sigma\cdot\lambda(x^\mu)}\psi \tag{39}$$

如同 (19) 式一樣，在此變換下 $\mathscr{L}$ 中在求導時會多出一項

$$\partial_\mu(\psi') = \partial_\mu(S\psi) = (\partial_\mu S)\psi + S\partial_\mu\psi \tag{40}$$

式如同 (30) 式，設協變求導 $D_\mu$ 並替換掉 $\partial_\mu$

$$D_\mu = \partial_\mu + iq\sigma\cdot A_\mu \tag{41}$$

這裡使用 $A_\mu$ 只是為了類比於電磁學中的過程而使用相似的符號，實際上這裡的 $A_\mu$ 與電磁勢完全不同，是四個 1×3 的向量 $A_\mu = [A_{1\mu}, A_{2\mu}, A_{3\mu}]$，且變換規則也是未知。為了符合原始論文內容，我們令 $\epsilon = -q$，而 $B_\mu$ 與 $A_\mu$ 的關係為

$$B_\mu = \sigma\cdot A_\mu \tag{42}$$

如此一來協變求導 $D_\mu$ 即可寫成與論文中相同的形式

$$D_\mu = \partial_\mu - i\epsilon B_\mu \tag{43}$$

可以看到 $B_\mu$ 都是 2×2 的矩陣而 $A_\mu$ 則都是向量，由於之後的許多計算過於複雜，用 $A_\mu$ 來算會相對簡單一點。或是我們可以遵照論文中的說法，假設一個在同位旋空間中旋轉的「同位旋角動量」$T$，其與原本包立矩陣的關係為 $T = \sigma/2$，如此一來 $B_\mu$ 和 $A_\mu$ 關係可寫成

$$\epsilon B_\mu = 2(-q)\frac{\sigma}{2} \cdot A_\mu = 2\epsilon T \cdot A_\mu \tag{44}$$

同樣為了符合論文中的符號，也為了與電磁學中的符號做出區別，我們幫 $A_\mu$ 取一個新的名子叫 $b_\mu$，所以 $b_\mu$ 和 $B_\mu$ 的關係為

$$B_\mu = 2b_\mu \cdot T \tag{45}$$

如此一來協變求導 $D_\mu$ 即可寫成

$$D_\mu = \partial_\mu - 2i\epsilon b_\mu \cdot T \tag{46}$$

現在為了使 (36) 式保持規範不變性，即要求 (36) 式中

$$i\overline{\psi'}\gamma^\mu D'_\mu \psi' = i\overline{\psi}S^\dagger \gamma^\mu D'_\mu S\psi = i\overline{\psi}\gamma^\mu D_\mu \psi \tag{47}$$

因此可得 $D_\mu$ 的變換需滿足的條件為

$$S^\dagger D'_\mu \psi' = D_\mu \psi \tag{48}$$

又因為 S 是么正矩陣，$S^\dagger = S^{-1}$，所以現在 $D_\mu$ 的變換需滿足的條件可寫成

$$\begin{aligned} &D'_\mu \psi' = SD_\mu \psi \\ \Rightarrow &(\partial_\mu - i\epsilon B'_\mu)\psi' = S(\partial_\mu - i\epsilon B_\mu)\psi \end{aligned} \tag{49}$$

其中 $\psi' = S\psi$ 並用 (40) 式將 (49) 式展開得

$$\begin{aligned} &(\partial_\mu S)\psi + S(\partial_\mu \psi) - i\epsilon B'_\mu S\psi = S(\partial_\mu \psi) - i\epsilon SB_\mu \psi \\ \Rightarrow &(\partial_\mu S) - i\epsilon B'_\mu S = -i\epsilon SB_\mu \\ \Rightarrow &i\epsilon B'_\mu S = (\partial_\mu S) + i\epsilon SB_\mu \\ \Rightarrow &B'_\mu S = \frac{-i}{\epsilon}(\partial_\mu S) + SB_\mu \\ \Rightarrow &B'_\mu SS^{-1} = \frac{-i}{\epsilon}(\partial_\mu S)S^{-1} + SB_\mu S^{-1} \\ \Rightarrow &B'_\mu = SB_\mu S^{-1} - \frac{i}{\epsilon}(\partial_\mu S)S^{-1} \end{aligned} \tag{50}$$

(50) 式中第二項與論文中的 (3) 式差了一個負號，不過這只是因為我們這裡使用的 S 變換方向與論文中的方向是相反的，但這並不重要，因為在同位旋空間的任意轉動可以用順時針旋轉或逆時針旋轉得到，即在考慮微小變換 $\epsilon\lambda$ 下 ($\lambda$ 是一個向量且是時空的函數)S 可以展開並保留到 $\lambda$ 一階項，寫成

$$S \approx 1 \pm 2iT \cdot \epsilon\lambda(x^\mu) \tag{51}$$

(50) 式中的 S 是使用 (51) 式中取正號，論文中的 S 則是使用負號，我將繼續使用取正號的 S，也就是

$$S \approx 1 + 2iT \cdot \epsilon\lambda(x^\mu) \quad , \quad S^{-1} \approx 1 - 2iT \cdot \epsilon\lambda(x^\mu) \tag{52}$$

在 (50) 式中展現了非阿貝爾群性質所帶來的複雜性，因為 $B_\mu$ 和 S 並不滿足交換率，所以 $B_\mu$ 的轉換方式比電磁學中的電磁勢轉換方式 ($A'_\mu = A_\mu + \partial_\mu \lambda$) 複雜很多。為了簡化 $B_\mu$ 的轉換方式我們使用 S 在微小變換下的近似

$$\begin{aligned} B'_\mu \approx &(1 + 2iT \cdot \epsilon\lambda)B_\mu(1 - 2iT \cdot \epsilon\lambda) \\ &- i(2iT \cdot \partial_\mu \lambda)(1 - 2iT \cdot \epsilon\lambda) \end{aligned} \tag{53}$$

將 (53) 式乘開並保留到 $\lambda$ 一階項得

$$B'_\mu \approx B_\mu - i2\epsilon[B_\mu, T \cdot \lambda] + 2T \cdot \partial_\mu \lambda \tag{54}$$

(54) 式中的 [ , ] 為對易算符。因為「同位旋角動量」T 的本質是包立矩陣，而 $B_\mu$ 根據 (45) 式的定義是兩倍的向量 $b_\mu$ 內積 T，如同前面提到的使用 $B_\mu$ 來計算會比用 $b_\mu$ 來的困難，因此這裡我將把尋找 $B_\mu$ 轉換方式改成是找 $b_\mu$ 的轉換方式，$B_\mu$ 和 $b_\mu$ 之間的關係如 (45) 式所示，如此一來便可使用包立矩陣的以下性質 (使用愛因斯坦求和約定)

$$\begin{aligned} &(\sigma \cdot a)(\sigma \cdot b) \\ &= \sigma_i a_i \sigma_j b_j = a_i b_j (\sigma_i \sigma_j) = a_i b_j (\delta_{ij} + i\epsilon_{ijk}\sigma_k) \\ &= a_i b_j \delta_{ij} + i\epsilon_{ijk} a_i b_j \sigma_k = a \cdot b + i\sigma \cdot (a \times b) \end{aligned} \tag{55}$$

其中 a 和 b 為任意實數向量，運用該關係式來計算 (54) 式中的對易關係即可得

$$\begin{aligned} [B_\mu, T \cdot \lambda] &= \left[\sigma \cdot b_\mu, \sigma \cdot \frac{\lambda}{2}\right] \\ &= (\sigma \cdot b_\mu)\left(\sigma \cdot \frac{\lambda}{2}\right) - \left(\sigma \cdot \frac{\lambda}{2}\right)(\sigma \cdot b_\mu) \\ &= i\sigma \cdot (b_\mu \times \lambda) \end{aligned} \tag{56}$$

因此 (54) 式可寫成

$$\begin{aligned} \sigma \cdot b'_\mu &\approx \sigma \cdot b_\mu - i2\epsilon(i\sigma \cdot (b_\mu \times \lambda)) + \sigma \cdot \partial_\mu \lambda \\ &= \sigma \cdot b_\mu + 2\epsilon\sigma \cdot (b_\mu \times \lambda) + \sigma \cdot \partial_\mu \lambda \end{aligned} \tag{57}$$

所以 $b_\mu$ 在考慮無窮小偶域規範變換 $\lambda \ll 1$ 時的轉換方式為

$$b'_\mu \approx b_\mu + 2\epsilon(b_\mu \times \lambda) + \partial_\mu \lambda \tag{58}$$

若比照論文內容，其中 $\epsilon\lambda$ 就等於論文中的 $\delta\omega$。所以在好不容易找到 $b_\mu$ 的轉換關係後，我們回去看原本的拉氏量在我們為了符合 SU(2) 偶域規範不變性而引入向量場 $b_\mu$ 後變成

$$\begin{aligned}\mathscr{L} &= i\overline{\psi}\gamma^\mu D_\mu\psi - m\overline{\psi}\psi \\ &= i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi + (\epsilon\overline{\psi}\gamma^\mu\sigma\psi)\cdot b_\mu \end{aligned} \tag{59}$$

這邊一樣仿照電磁學中的流程，(59) 式中我們得到旋量場 $\psi$ 本身和與向量場 $b_\mu$ 的交互作用關係，那麼理應也要有 $b_\mu$ 場自身的交互作用項，由於 $b_\mu$ 實際上長 $b_\mu = [b_{1\mu}, b_{2\mu}, b_{3\mu}]$，因此要在拉氏量中加入而外的三個項，將表示 $b_\mu$ 場自身交互作用的拉氏量記作 $\mathscr{L}_b$，其形式仿照電磁學為

$$\begin{aligned}\mathscr{L}_b &= \frac{-1}{16\pi}\left(f_1^{\mu\nu}f_{\mu\nu 1} + f_2^{\mu\nu}f_{\mu\nu 2} + f_3^{\mu\nu}f_{\mu\nu 3}\right) \\ &= \frac{-1}{16\pi}f^{\mu\nu}\cdot f_{\mu\nu}\end{aligned} \tag{60}$$

同樣因為侷域規範不變的要求所以排除了 $b^\mu \cdot b_\mu$ 項存在的可能性 (這裡的 $b^\mu\cdot b_\mu$ 一樣代表三項 $b_i^\mu b_{\mu i}$ 縮併的和)，然而根據 (26) 式的 Proca 作用量可知這一項與傳遞 $b_\mu$ 場作用力的玻色子質量有關，但在要求侷域規範不變性下此項一定為零，也就是傳遞 $b_\mu$ 場作用力的玻色子質量肯定是零，從這影發的質量問題將會在後面討論。現在的問題是 $b_\mu$ 本身的變換關係就較為複雜了，能夠滿足侷域規範不變的 $f^{\mu\nu}$ 肯定也比電磁學中複雜許多，因此為了降低複雜度，我們同樣考慮無窮小侷域規範變換的近似。但儘管說是降低了複雜度，但要給出論文中 (9) 式的變換規則可真的是十分複雜，這篇論問許多輕描淡寫帶過的公式背後的數學複雜度乍看之下是難以體會的，實際操作過後才真的能體悟楊振寧先生的聰明才智。話說回來，由於 $f^{\mu\nu}$ 應該等於 $\partial^\mu b^\nu - \partial^\nu b^\mu$ 再加上修正項，於是現在的策略是先考慮在 (58) 式中 $b_\mu$ 的變換規則下 $f^{\mu\nu} = \partial^\mu b^\nu - \partial^\nu b^\mu$ 會多出什麼使得 $f^{\mu\nu}\cdot f_{\mu\nu}$ 無法在無窮小侷域規範變換下前後不一致，最後再加入修正項抵銷多出的部分，所以在假設 $f^{\mu\nu} = \partial^\mu b^\nu - \partial^\nu b^\mu$ 的狀況下，其無窮小侷域規範變換為

$$\begin{aligned}f'^{\mu\nu} =&\partial^\mu(b^\nu + 2\epsilon(b^\nu\times\lambda) + \partial^\nu\lambda) \\ &- \partial^\nu(b^\mu + 2\epsilon(b^\mu\times\lambda) + \partial^\mu\lambda) \\ =&\partial^\mu b^\nu - \partial^\nu b^\mu + \partial^\mu\partial^\nu\lambda - \partial^\nu\partial^\mu\lambda \\ &+ 2\epsilon[\partial^\mu(b^\nu\times\lambda) - \partial^\nu(b^\mu\times\lambda)] \\ =&f^{\mu\nu} + 2\epsilon[(\partial^\mu b^\nu)\times\lambda + b^\nu\times(\partial^\mu\lambda) \\ &- (\partial^\nu b^\mu)\times\lambda - b^\mu\times(\partial^\nu\lambda)] \\ =&f^{\mu\nu} + 2\epsilon[f^{\mu\nu}\times\lambda + b^\nu\times(\partial^\mu\lambda) - b^\mu\times(\partial^\nu\lambda)]\end{aligned} \tag{61}$$

下一步計算 $f'^{\mu\nu}\cdot f'_{\mu\nu}$，省略 $\lambda^2$ 以上的高次並，觀察多出的項

$$\begin{aligned}f'^{\mu\nu}\cdot f'_{\mu\nu} =&\{f^{\mu\nu} + 2\epsilon[f^{\mu\nu}\times\lambda + b^\nu\times(\partial^\mu\lambda) - b^\mu\times(\partial^\nu\lambda)]\} \\ &\cdot\{f_{\mu\nu} + 2\epsilon[f_{\mu\nu}\times\lambda + b_\nu\times(\partial_\mu\lambda) - b_\mu\times(\partial_\nu\lambda)]\} \\ =&f^{\mu\nu}\cdot f_{\mu\nu} + 4\epsilon\{f^{\mu\nu}\cdot[b_\nu\times(\partial_\mu\lambda) - b_\mu\times(\partial_\nu\lambda)]\} \\ =&f^{\mu\nu}\cdot f_{\mu\nu} - 8\epsilon\{f^{\mu\nu}\cdot[b_\mu\times(\partial_\nu\lambda)]\} \\ =&f^{\mu\nu}\cdot f_{\mu\nu} + 8\epsilon\{\partial_\nu\lambda\cdot[b_\mu\times f^{\mu\nu}]\}\end{aligned} \tag{62}$$

所以現在要在 $f^{\mu\nu} = \partial^\mu b^\nu - \partial^\nu b^\mu$ 中加入一個修正項來抵銷計算 $f'^{\mu\nu}\cdot f'_{\mu\nu}$ 產生的 $8\epsilon\{\partial_\nu\lambda\cdot[b_\mu\times f^{\mu\nu}]\}$。這一步就有點要靠通靈了，總之聰明絕頂的楊振寧先生給出正確 $f^{\mu\nu}$ 的關係為 (在一次演講中楊振寧先生曾提到這一步困擾了他很久，而正是和米爾斯合作後才突破了這個關口)

$$f^{\mu\nu} = \partial^\mu b^\nu - \partial^\nu b^\mu - 2\epsilon(b^\nu\times b^\mu) \tag{63}$$

$f^{\mu\nu}$ 在無窮小規範變換下的轉換關係為

$$f'^{\mu\nu} = f^{\mu\nu} + 2\epsilon(f^{\mu\nu}\times\lambda) \tag{64}$$

所以在此定義下，(60) 式中的 $\mathscr{L}_b$ 在有限侷域規範變換下是不變的，此時完整的楊-米爾斯拉氏量用 $f^{\mu\nu}$ 可寫成

$$\mathscr{L} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi - \frac{1}{16\pi}f^{\mu\nu}\cdot f_{\mu\nu} + (\epsilon\overline{\psi}\gamma^\mu\sigma\psi)\cdot b_\mu \tag{65}$$

若是使用 (46) 式中定義的協變求導 $D_\mu$ 則可寫成較為簡潔的形式

$$\mathscr{L} = i\overline{\psi}\gamma^\mu D_\mu\psi - m\overline{\psi}\psi - \frac{1}{16\pi}f^{\mu\nu}\cdot f_{\mu\nu} \tag{66}$$

這裡因為我使用的是高斯單位制，而原始論文中使用的是 SI 單位制，所以在 $f^{\mu\nu}\cdot f_{\mu\nu}$ 係數有 $1/16\pi$ 和 $1/4$ 的差別 (本文與論文皆令 $\hbar = 1, c = 1$)，還有其他如狄拉克場拉氏量前面是否有個 i 所造成的一些小差異，不過以上這些不是特別重要，我這裡使用的是現在比較常用的傳統，且本文前後的表達設定都是自洽的。總之該拉氏量描述了兩個質量相等的狄拉克場與三個質量為零的向量規範場 $b_\mu$ 的相互作用。所有的結果都來自於要求 (36) 式中兩個質量相等的狄拉克場的拉氏量滿足 SU(2) 的全域與侷域的規範不變性。若仿照電磁學中的方法，由兩個質量相等的狄拉克場產生了三個流

$$J^\mu = (-\epsilon\overline{\psi}\gamma^\mu\sigma\psi) \tag{67}$$

$J^\mu$ 代表了規範場 $b_\mu$ 中三個分量的三個源，此時若把規範場 $b_\mu$ 本身的拉氏量寫出來

$$\mathscr{L}_b = -\frac{1}{16\pi}f^{\mu\nu}\cdot f_{\mu\nu} - J^\mu\cdot b_\mu \tag{68}$$

可以發現這與電磁學中的拉氏量十分相像，但背後的複雜程度卻不可同日而語，原因就在於楊-米爾斯理論是非阿貝爾的規範理論。

我覺得本文至此整個楊-米爾斯理論的核心概念「一個整體不變性 (對稱性) 應該也要在侷域成立，並藉此決定交互作用的規則」已經表達的十分清楚了，所以關於論文的解讀我想在這邊告一個段落。論文在後面還討論了規範場 $b_\mu$ 的量子化，但楊振寧先生的論文字字珠璣，雖然看上去十分簡短，但由於本身知識的嚴重不足，為了解讀前面內容背後的數學並將前因後果理清已經讓我筋疲力盡了，所以請容許我略過規範場 $b_\mu$ 的量子化部分的內容。

## III. 對物理發展的影響

### A. 質量問題

前面曾經提到為了滿足侷域的規範不變性，規範場 $b_\mu$ 的質量必定為零，也就是傳遞交互作用力的規範玻色子質量必須為零，從這衍生的問題使的楊-米爾斯理論剛被提出時並不被世人重視，其根本原因在於規範玻色子是否有質量決定了該交互作用是長程力或是短程力。無質量的規範玻色子例如光子，其傳遞的電磁交互作用力就是長程力，而強力和弱力因為只在原子核中產生作用，所以顯然強力和弱力是短程力，這也就代表傳遞強力和弱力的規範玻色子應該要有質量才對，由於在 1950 年代世人對強力和弱力的了解十分有限，因此楊-米爾斯理論在應用層面上出現了致命的缺點，當然後來我們明白傳遞強力的膠子確實是無質量的，造成強力是短程力的原因與夸克漸近自由的性質有關。至於傳遞弱力的 W 與 Z 玻色子確實是有質量的，其原因後來由希格斯機制給出了規範玻色子獲得質量的原因。

### B. 標準模型

楊-米爾斯理論雖然一開始從強力中的同位旋守恆出發，但由於理論中是使用兩個「質量完全相等」的狄拉克場的拉氏量，畢竟與實際上質子與中子的狀況不符，因此透過同位旋守恆使用 SU(2) 群作為描述強力的對稱群自然不會得到正確描述強力交互作用的理論。但楊-米爾斯理論對於之後標準模型的建立有著至關重要的地位，作為一套數學框架，只要將它用在合適的地方便能得到豐厚的成果，以下列出楊-米爾斯理論應用的理論與其使用的對稱群

- 量子電動力學　U(1)
- 電弱交互作用　U(1)×SU(2)
- 量子色動力學　SU(3)
- 標準模型　U(1)×SU(2)×SU(3)

當然以上這些理論之所以能成功除了依靠楊-米爾斯理論外，也仰賴了其他許多偉大物理學家提出的理論與貢獻，例如夸克的概念、非阿貝爾規範場的重整化和希格斯機制等等。

[1] C. N. Yang and R. L. Mills, Conservation of isotopic spin and isotopic gauge invariance, Phys. Rev. **96**, 191 (1954).