



HAL
open science

Logic of gauge

Alexander Afriat

► **To cite this version:**

| Alexander Afriat. Logic of gauge. 2013. hal-00769688v1

HAL Id: hal-00769688

<https://hal.science/hal-00769688v1>

Preprint submitted on 3 Jan 2013 (v1), last revised 3 Jan 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Logic of gauge

Alexander Afriat

January 3, 2013

Abstract

The logic of gauge theory is considered by tracing its development from general relativity to Yang-Mills theory, through Weyl's two gauge theories.

1 Weyl's first gauge theory

1.1 Geometrical justice

First, there was general relativity.¹ Levi-Civita (1917) noticed that the connection determined by Einstein's covariant derivative transported the *direction* of a vector anholonomically, but not its *length*, which was left unchanged. This was unfair, protested Weyl—length deserved the same treatment as direction.² To remedy he proposed a more general theory that propagated length just as anholonomically as direction. *Congruent* transport would also be governed by a connection, which Weyl defined³ as a bilinear mapping between neighbouring points: linear in the object propagated and in the direction of propagation. A connection transporting the (squared) length l from $a = \gamma(a)$ to its neighbour⁴ $b = \gamma(b)$ along $\gamma : [a, b] \rightarrow M$ (the universe) would therefore be a real-valued⁵ one-form $A = A_\mu dx^\mu$ applied— $\langle A, \dot{\gamma} \rangle = A_\mu \dot{\gamma}^\mu$ —to the direction $\dot{\gamma} = \dot{\gamma}^\mu \partial_\mu \in T_a M$ and multiplied by the initial length l_a , yielding the small difference

$$\delta l = l_b - l_a = l_a \langle A, \dot{\gamma} \rangle$$

subtracted from l_a . The final length l_b is $l_a(1 - \langle A, \dot{\gamma} \rangle)$ —unless a and b are too far apart for γ to remain straight in the interval, in which case l_b is

$$l_a \exp \int_\gamma A.$$

¹Einstein (1916)

²See Afriat (2009). Ryckman (2005) provides an alternative analysis.

³More on connections in §1.4.

⁴Which is so close to a it practically belongs to the tangent space $T_a M$; see Weyl (1927) p. 28, Weyl (1931a) p. 52.

⁵Here the structure group is the multiplicative group $\langle \mathbb{R}_{\neq 0}, \times \rangle$ of dilations, generated by the Lie algebra $\langle \mathbb{R}, +, [\cdot, \cdot] \rangle$ or rather $\langle \mathbb{R}, + \rangle$; the Lie product $[\cdot, \cdot]$ vanishes since real numbers commute.

To deal with the geometrical injustice that A was introduced to remedy, the curvature

$$(1) \quad F = \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu = dA = \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu)dx^\mu \wedge dx^\nu$$

cannot vanish—unlike the three-form

$$(2) \quad dF = d^2A = \frac{1}{6}(\partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} + \partial_\sigma F_{\mu\nu})dx^\mu \wedge dx^\nu \wedge dx^\sigma,$$

which does. Seeing all this, Weyl could not help thinking of the electromagnetic four-potential A , the Faraday two-form $F = dA$ and Maxwell's two homogeneous equations⁶ $dF = 0$: he had unified gravity and electromagnetism, by mistake!⁷ And indeed Einstein would soon point out the mistake: the anholonomy on which Weyl had founded his theory was not observed in nature, as we shall see in §1.3.

1.2 Gauge

Weyl seeks to rectify general relativity using the curvature (1), which ensures geometrical justice. Differentiation is destructive (or rather irreversible); what d destroys is the freedom

$$(3) \quad A \mapsto A' = A - d\lambda$$

invisible to $F = dA = dA'$, in the sense that the inverse image of F under d is the whole equivalence class $[A]$ given by the equivalence relation $A \sim (A + d\lambda)$. If A only served to produce the curvature F , (3) would be meaningless; but A appears elsewhere too, notably in the law of propagation⁸

$$(4) \quad \nabla g = A \otimes g,$$

which is not indifferent to (3). To make (4) invariant, (3) therefore has to be balanced by

$$(5) \quad g \mapsto g' = e^\lambda g.$$

Such compensation is typical⁹ of a gauge theory: an invariant expression (here (4)) is sensitive to a first transformation, and to a second as well—but indifferent to the two together, if their variations are appropriately constrained, and balance one another.

So far we have two logical ingredients

1. GR: *general relativity*

⁶ $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0$

⁷See Ryckman (2005).

⁸This generalises the length-preserving condition $\nabla g = 0$ satisfied by the Levi-Civita connection.

⁹Typical but mysterious, even for Weyl (1931a) p. 54: “insbesondere konnte ich nichts a priori Einleuchtendes vorbringen zugunsten der Koppelung des willkürlichen additiven Gliedes $\partial\lambda/\partial x_p$, das nach der Erfahrung in den Komponenten des elektromagnetischen Potentials steckt, mit dem von der klassischen Geometrie geforderten Eichfaktor e^λ .”

2. GJ: geometrical justice

which together yield Weyl's theory of electricity and gravity; I'll write

$$\text{GR \& GJ} \rightarrow \text{W18.}$$

The next logical ingredients will be MW: *matter wave* and EO: *avoid Einstein's objection* ...

1.3 Einstein's objection

The tangent of a worldline's *image* $\bar{\gamma} \subset M$ only has a direction; the length $l = \|\dot{\gamma}\|^2 = g(\dot{\gamma}, \dot{\gamma})$ of the tangent $\dot{\gamma} = d\gamma/dt$ is given by the parameter rate $\partial\gamma/\partial t$. If the values of the parameter are identified with the readings of a clock describing γ , the length l giving the proper ticking rate should remain constant—the hands of a good clock don't accelerate. But far from remaining constant, lengths in Weyl's theory are not even integrable:

$$l_b(\gamma) = l_a \exp \int_{\gamma} A$$

depends on γ —whereas an exact connection $A = d\mu$ would of course give

$$l_b = l_a \exp \int_a^b d\mu = l_a \exp \Delta\mu$$

along any path, $\Delta\mu$ being the difference $\mu(b) - \mu(a)$ between the final and initial values of μ . In addition to the *first* clock effect (Langevin's twins) already present in Einstein's theory, Weyl's theory therefore has a *second* clock effect expressed in the anholonomy of ticking rates.

Einstein objected that *nature provides integrable clocks*.¹⁰ Two clocks trace out a loop $\bar{\gamma} = \partial\omega$ enclosing a region ω (without holes): starting from the same point a they describe worldlines γ_1, γ_2 that meet at b . They will tick at the same rate if A is exact, for then

$$\oint_{\partial\omega} d\mu = \iint_{\omega} d^2\mu$$

¹⁰Letter to Weyl dated 15 April 1918: "So schön Ihre Gedanke ist, muss ich doch offen sagen, dass es nach meiner Ansicht ausgeschlossen ist, dass die Theorie die Natur entspricht. Das ds selbst hat nämlich reale Bedeutung. Denken Sie sich zwei Uhren, die relativ zueinander ruhend neben einander gleich rasch gehen. Werden sie voneinander getrennt, in beliebiger Weise bewegt und dann wieder zusammen gebracht, so werden sie wieder gleich (rasch) gehen, d. h. ihr relativer Gang hängt nicht von der Vorgeschichte ab. Denke ich mir zwei Punkte P_1 & P_2 die durch eine Zeitartige Linie verbunden werden können. Die an P_1 & P_2 anliegenden zeitartigen Elemente ds_1 und ds_2 können dann durch mehrere zeitartigen Linien verbunden werden, auf denen sie liegen. Auf diesen laufende Uhren werden ein Verhältnis $ds_1 : ds_2$ liefern, welches von der Wahl der verbindenden Kurven unabhängig ist.—Lässt man den Zusammenhang des ds mit Massstab- und Uhr-Messungen fallen, so verliert die Rel. Theorie überhaupt ihre empirische Basis." Another letter to Weyl, four days later: "wenn die Länge eines Einheitsmassstabes (bezw. die Gang-Geschwindigkeit einer Einheitsuhr) von der Vorgeschichte abhängen. Wäre dies in der Natur wirklich so, dann könnte es nicht chemische Elemente mit Spektrallinien von bestimmter Frequenz geben, sondern es müsste die relative Frequenz zweier (räumlich benachbarter) Atome der gleichen Art im Allgemeinen verschieden sein. Da dies nicht der Fall ist, scheint mir die Grundhypothese der Theorie leider nicht annehmbar, deren Tiefe und Kühnheit aber jeden Leser mit Bewunderung erfüllen muss."

vanishes—in fact (without holes) it is enough for A to be closed,

$$\oint_{\partial\omega} A = \iint_{\omega} dA$$

vanishes too if dA does. But if the loop encloses an electromagnetic field $F = dA$, one of the clocks will tick faster than the other once they're compared at b . In any case the theory didn't work: from the beginning it rested on an anholonomy not seen in nature.

1.4 Connections

A few more words about connections, which provide a notion of constancy or free fall or absence of force, even in the presence of influences that can complicate it. The fundamental operation is linear and infinitesimal, and can be extended by integration. It serves to displace an object, say a vector V , from a to a neighbouring point b . One could first try to produce the image $V_b \in \mathbb{V}_b$ *directly*, by multiplying the initial vector V_a ; say $V_b = \Xi V_a$, where the linear operator $\Xi : \mathbb{V}_a \rightarrow \mathbb{V}_b$ belongs to a representation of an appropriate group. If one stopped there, at b , such an operation would be natural; but indeed one wants to continue, by integration—which is a limit of sums, not of products. To the initial vector one therefore has to add the corrections accumulated along the way. One proceeds more indirectly, by producing the small difference

$$\delta V_b = V_b - V'_b = \mathbf{T}V_a \in \mathbb{V}_b$$

between the 'perturbed' and 'unperturbed' vectors V'_b and V_b . The perturbed components $V'^{\mu} = \langle \mathbf{e}^{\mu}, V' \rangle$ can be characterised as sharing the perturbation of the tetrad:

$$V'_b{}^{\mu} = \langle \mathbf{e}'^{\mu}, V'_b \rangle = V_a{}^{\mu} = \langle \mathbf{e}_a^{\mu}, V_a \rangle.$$

The infinitesimal generator $\mathbf{T} : \mathbb{V}_a \rightarrow \mathbb{V}_b$ acts on the components $V_a{}^{\mu}$ through a representation $T_{\mu}^{\nu} = \langle \mathbf{e}^{\nu}, \mathbf{T}\mathbf{e}_{\mu} \rangle$:

$$\delta V_b{}^{\nu} = V_b{}^{\nu} - V'_b{}^{\nu} = T_{\mu}^{\nu} V_a{}^{\mu}.$$

If the tetrad is in unperturbed 'free fall,' V_a is in the kernel of \mathbf{T} and the difference δV_b vanishes.

We have seen that the congruent transport of Weyl's first gauge theory is produced by a real-valued connection A , which, applied to a direction $\dot{\gamma} \in T_a M$ of propagation, yields a dilation generator $\langle A, \dot{\gamma} \rangle$ belonging to the Lie algebra \mathbb{R} of the dilation group $\langle \mathbb{R}_{\neq 0}, \times \rangle$. In the second gauge theory¹¹ the structure group becomes $\mathbb{U}(1) = e^{i\mathbb{R}}$, whose Lie algebra is $i\mathbb{R} = \text{Lie}\mathbb{U}(1)$.

2 Weyl's second gauge theory

The setback of 1918 and Einstein's objection (his *preaching!*¹²) had their benefits, they taught Weyl the experimental character of physics, an *empirical discipline founded*

¹¹Weyl (1929a,b,c). See Straumann (1987), O'Raifeartaigh (1997), O'Raifeartaigh & Straumann (2000), Brading (2002), Scholz (2005) for more recent accounts.

¹²Letter to Seelig—quoted in Seelig (1960) p. 274—in which Weyl quotes Einstein: "So – das heisst auf so spekulative Weise, ohne ein leitendes, anschauliches physikalisches Prinzip – macht man keine Physik!"

directly on experience and not a geometrical fantasy deduced from aesthetic hunches: “All these geometrical leaps-in-the-air [W18] were premature, we return [W29] to the solid ground of physical facts.”¹³ In 1929 he’s understood, matured, and will have his “revenge”;¹⁴ he’ll even claim that his new theory came straight out of experience,¹⁵ directly derived from spectrographic data¹⁶ . . .

For his new theory takes account of the electron’s spin—which in fact got there through the Dirac equation; and in Dirac’s argument spin does not come (straight) out of experience¹⁷ but out of a mathematical, aesthetic, *a priori* principle, in much the same spirit as the geometrical justice that produced Weyl’s first gauge theory.

2.1 The new undulatory ontology

But let us go back a few years. Louis de Broglie (1924), Schrödinger (1926) *et al.* had meanwhile produced an *undulatory* world. Weyl had no reason to get rid of electricity or gravitation; to those existing ingredients he therefore had to add a matter wave, to update his ontology. As long as the theory included only gravity and electricity, the gauge relation (3)-(5) could only hold between *them*; but now, with a third element, as many compensations were in principle possible, of which only two were plausible: the old relation between gravity and electricity, and a new one between electricity and the matter wave. With (3)-(5) the theory would have remained subject to Einstein’s objection—which the presence of the electron’s wavelength¹⁸ h/mc in the Dirac equation made even more convincing,¹⁹ by providing an absolute standard of length allowing the distant comparisons Weyl wanted to prevent in 1918.²⁰ The other possibility

¹³Weyl (1931a) p. 56: “Alle diese geometrischen Luftsprünge waren verfrüht, wir kehren zurück auf den festen Boden der physikalischen Tatsachen.”

¹⁴“Rache”; Pauli (1979) p. 518: “Als Sie früher die Theorie mit $g'_{ik} = \lambda g_{ik}$ machten, war dies reine Mathematik und unphysikalisch. Einstein konnte mit Recht kritisieren und schimpfen. Nun ist die Stunde der Rache für Sie gekommen; jetzt hat Einstein den Bock des Fernparallelismus geschossen, der auch nur reine Mathematik ist und nichts mit Physik zu tun hat, und Sie können schimpfen!”

¹⁵Weyl (1931a) p. 57: “Das neue Prinzip ist aus der *Erfahrung* erwachsen und resümiert einen gewaltigen, aus der Spektroskopie entsprungene Erfahrungsschatz.”

¹⁶Weyl (1931a) p. 57: “Dieses Transformationsgesetz der ψ ist zuerst von PAULI aufgestellt worden und folgt mit unfehlbarer Sicherheit aus den spektroskopischen Tatsachen, genauer aus den Termdoublets der Alkalispektren und der Tatsache, daß die Doublettenkomponenten nach Ausweis ihres Zeemaneffekts *halbganze* innere Quantenzahlen besitzen.”

¹⁷On the logical priority of relativity (or spin) *cf.* Weyl (1931b) p. 193: “Da die Möglichkeit einer solchen relativitätsinvarianten Gleichung für ein skalares ψ nicht vorhanden ist, erscheint *der spin als ein durch die Relativitätstheorie notwendig gefordertes Phänomen.*”

¹⁸But here Planck’s constant h , the speed of light c —and even charge—are set equal to one.

¹⁹Weyl (1929c) p. 284: “By this new situation, which introduces an atomic radius into the field equations themselves—but not until this step—my principle of *gauge-invariance*, with which I had hoped to relate gravitation and electricity, is robbed of its support.” Weyl (1931a) p. 55: “Die Atomistik gibt uns ja absolute Einheiten für alle Maßgrößen an die Hand. [...] So geht in die DIRACsche Feldgesetze des Elektrons die „Wellenlänge des Elektrons“, die Zahl h/mc , als eine absolute Konstante ein. Damit fällt das Grundprinzip meiner Theorie, das Prinzip von der Relativität der Längenmessung, dem Atomismus zum Opfer und verliert seine Überzeugungskraft.”

²⁰See also Weyl (1929c) p. 290.

remained: (3) with a quantum version of (5),²¹ of which the best was manifestly²²

$$(6) \quad \psi \mapsto \psi' = e^{i\lambda}\psi,$$

where $\mathbb{U}(1)$ replaced the multiplicative group \mathbb{R} of (5).²³ As ψ was now part of a four-dimensional spacetime theory, it could no longer obey the Schrödinger equation, which violates relativity by treating space and time very differently.²⁴ Weyl adopted what amounted to a Dirac equation, but cut in half: deprived of mass and the associated crisscrossing of component pairs ...

We now have four logical ingredients:

1. GR: *general relativity*
2. GJ: *geometrical justice*
3. MW: *matter wave*
4. EO: *avoid Einstein's objection;*

W29 ← W18 & MW & EO & ?

A final ingredient, EL: *twice too many energy levels*, will almost be enough to produce the second gauge theory.

2.2 Dirac-Weyl theory

We can take $H = p_x^2$ as the simplified Hamiltonian of a particle whose mass is one-half. Momentum p in quantum mechanics is represented by differentiation, in the sense

²¹Weyl (1929c) p. 284: "this principle has an equivalent in the quantum-theoretical field equations which is exactly like it in formal respects; the laws are invariant under the simultaneous replacement of ψ by $e^{i\lambda}\psi$, φ_α by $\varphi_\alpha - \partial\lambda/\partial x_\alpha$ where λ is an arbitrary real function of position and time."

²²The conservation requirement $\|\psi'\|^2 = \|\psi\|^2$ being very natural. And transformation (6) is not even observable (with respect to position at any rate); cf. Weyl (1931b) p. 195: "Aus der Natur, dem Transformationsgesetz der Größe ψ ergibt sich, daß die vier Komponenten ψ_ρ relativ zum lokalen Achsenkreuz nur bis auf einen gemeinsamen Proportionalitätsfaktor $e^{i\lambda}$ durch den physikalischen Zustand bestimmt sind, dessen Exponent λ willkürlich vom Orte in Raum und Zeit abhängt, und daß infolgedessen zur eindeutigen Festlegung des kovarianten Differentials von ψ eine Linearform $\sum_\alpha f_\alpha dx_\alpha$ erforderlich ist, die so mit dem Eichfaktor in ψ gekoppelt ist, wie es das Prinzip der Eichinvarianz verlangt."

²³Weyl (1931a) p. 55: "In dem theoretischen Weltbild bedeutet die Verwandlung von f_p in $-f_p$ eine objektive Änderung des metrischen Feldes; denn es ist etwas anderes, ob sich eine Strecke bei kongruenter Verpflanzung längs einer geschlossenen Bahn vergrößert oder verkleinert. Nach dem angenommenen Wirkungsgesetz aber ist die Entscheidung über das Vorzeichen der f_p auf Grund der beobachteten Erscheinungen unmöglich. Hier enthält darum, in Widerstreit mit einem oben ausgesprochenen erkenntnistheoretischen Grundsatz, das theoretische Weltbild eine Verschiedenheit, welche sich auf keine Weise für die Wahrnehmung aufbrechen läßt." P. 57: "Die an der alten Theorie gerügte Unsicherheit des Vorzeichens $\pm f_p$ löst sich dadurch in das unbestimmte Vorzeichen der $\sqrt{-1}$ auf. Schon damals, als ich die alte Theorie aufstellte, hatte ich das Gefühl, daß der Eichfaktor die Form $e^{i\lambda}$ haben sollte; nur konnte ich dafür natürlich keine geometrische Deutung finden. Arbeiten von SCHRÖDINGER und F. LONDON stützten die Forderung durch die allmählich sich immer deutlicher abzeichnende Beziehung zur Quantentheorie." See also Weyl (1931b) p. 89.

²⁴Weyl (1931b) pp. 187-8: "Es ist klar, daß man zu einer befriedigenden Theorie des Elektrons nur kommen wird, wenn es gelingt, das Grundgesetz seiner Bewegung in der von der Relativitätstheorie geforderten, gegenüber Lorentz-Transformationen invarianten Form zu fassen."

that²⁵

$$(7) \quad p \mapsto id,$$

in components $p_\mu \mapsto i\partial_\mu$. Our quantum Hamiltonian will therefore be $-\partial_x^2$, which means that Schrödinger's equation $i\partial_t\psi = \partial_x^2\psi$ differentiates space twice as much as time. But by what should it be replaced? The d'Alembertian $\square = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2$ and Klein-Gordon equation $(\square - m^2)\psi = 0$ treat space the same way as time, they have the right transformation properties; but \square is 'squared' and there are reasons to prefer a wave operator and especially a time derivative²⁶ that aren't. In seeking a square root $\sqrt{\square}$ Dirac found $\not{\partial} = \gamma^\mu\partial_\mu$, where the γ^μ 's have the algebraic properties needed to get rid of the cross terms that appear when squaring. He therefore proposed the *Dirac equation*

$$(8) \quad (m - i\not{\partial})\psi = 0$$

which not only treats the three spatial derivatives $\gamma^k\partial_k$ the same way as the time derivative $\gamma^0\partial_0$, but differentiates with respect to time only once.²⁷ The γ^μ 's, which do not commute, cannot be numbers; they admit for instance the canonical representations

$$(9) \quad \gamma^0 \leftrightarrow \begin{pmatrix} 0 & \sigma^0 \\ -\sigma^0 & 0 \end{pmatrix} \quad \gamma^k \leftrightarrow \begin{pmatrix} 0 & \sigma^k \\ \sigma^k & 0 \end{pmatrix},$$

where all four quaternions $\sigma^\mu : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ are hermitian and unitary; σ^0 is the identity, and the three traceless operators σ^k satisfy $2i\sigma^j = \varepsilon_{jkl}[\sigma^k, \sigma^l]$.

The wave ψ on which the γ^μ 's act therefore has four (complex) components—*embarras de richesses* which Weyl found most troubling: “doppelt zu viel Energieiveaus”! The anti-diagonality of the γ^μ 's governs the embarrassing excess by swapping the two two-spinors making up ψ . As the embarrassment is due to the *sign* that distinguishes between the different interweavings²⁸ produced by the γ^μ 's, Weyl deals with it by choosing the only mass—zero—that doesn't distinguish between plus and minus.²⁹

²⁵See Weyl (1931b) p. 89.

²⁶Weyl (1931b) p. 188: “Sie ist nicht im Einklang mit dem allgemeinen Schema der Quantenmechanik, welches verlangt, daß die zeitliche Ableitung nur in der ersten Ordnung auftritt.” P. 193: “Legt man die de Brogliesche Wellengleichung für das skalare ψ zugrunde, in welche die elektromagnetischen potentiale $[A_\mu]$ durch die Regel [(10)] eingeführt sind, so ergibt sich aber für die elektrische Dichte ein Ausdruck, der außer ψ die zeitliche Ableitung $\partial\psi/\partial t$ enthält und nichts mit der Ortswahrscheinlichkeit zu tun hat; sein Integral ist überhaupt keine Einzelform. Dies ist nach *Dirac* das entscheidendste Argument dafür, daß die Differentialgleichungen des in einem elektromagnetischen Feld sich bewegenden Elektrons von 1. Ordnung in bezug auf die zeitliche Ableitung sein müssen.”

²⁷Weyl (1931b) p. 190: “Nach dem allgemeinen Schema der Quantenmechanik sollte, wie schon erwähnt, die Differentialgleichung für ψ von 1. Ordnung hinsichtlich der zeitlichen Ableitung von ψ sein. Gemäß dem Relativitätsprinzip kann sie aber dann auch nur die 1. Ableitungen nach den räumlichen Koordinaten enthalten.”

²⁸Symplectic for time but simply ‘NOT’ for space. The interweaving produced by a purely NOT γ^0 would be gratuitous; the symplecticity given by the sign difference is essential—with respect to the three γ^k 's with merely NOT anti-diagonality.

²⁹Weyl (1929c) p. 292: “The [mass] term (5) of the Dirac theory is, however, more doubtful. It must be admitted that if we retain it we can obtain all details of the line spectrum of the hydrogen atom—of one electron moving in the electrostatic field of a nucleus—in accord with what is know from experiment. But

Without mass and half the components, (8) becomes $\sigma^\mu \partial_\mu \zeta = 0$. The reduced wave ζ has two complex components ζ_1, ζ_2 but four ('lightlike',³⁰) real ones: the squared length $x_0 = \|\zeta\|^2 = \bar{\zeta} \sigma^0 \zeta$ and the three hermitian quadratic forms $x_k = \bar{\zeta} \sigma^k \zeta$.

We now have five logical ingredients:

1. GR: *general relativity*
2. GJ: *geometrical justice*
3. MW: *matter wave*
4. EO: *avoid Einstein's objection*
5. EL: *twice too many energy levels*;

W18 & MW & EO & EL \rightarrow W29.³¹ The foundations are in place, the rest will follow.

2.3 Tetrads and spinors

The need for tetrads follows from the presence of spinors (introduced, as we have seen, to satisfy relativistic transformation requirements).

As long as there were only tensors

$$B = B_{\sigma \dots}^{\mu \nu \dots} \partial_\mu \otimes \partial_\nu \otimes \dots \otimes dx^\sigma \otimes \dots,$$

their components $B_{\sigma \dots}^{\mu \nu \dots}$ and the (holonomic) bases ∂_μ and dx^σ with respect to which they were represented were subject to (appropriate representations of) $\text{GL}(4, \mathbb{R})$. But this is too comprehensive a group to preserve scalar products—whose preservation was, however, required by the spinors representing the new arrival, the matter wave.³² The two complex components of a normalised spinor ϕ are subject to a 2×2 complex representation of $\text{SU}(2)$; the three real components $\bar{\phi} \sigma^k \phi$ to a 3×3 real representation of $\text{SO}(3)$; the four real components $\bar{\zeta} \sigma^\mu \zeta$ of the arbitrarily long spinor ζ to a 4×4 real

we obtain twice too much; if we replace the electron by a particle of the same mass and positive charge $+e$ (which admittedly does not exist in nature) the Dirac theory gives, contrary to all reason and experience, the same energy terms as for a negative electron, except for a change in sign. Obviously an essential change is here necessary." P. 294: "Be bold enough to leave the term involving mass entirely out of the field equations."

³⁰The condition $x_0^2 = x_1^2 + x_2^2 + x_3^2$ applies.

³¹W18 & MW & EO give something like Dirac-Maxwell theory in curved spacetime.

³²Weyl (1929c) p. 285: "The tensor calculus is not the proper mathematical instrument to use in translating the quantum-theoretic equations of the electron over into the *general theory of relativity*. Vectors and terms are so constituted that the law which defines the transformation of their components from one Cartesian set of axes to another can be extended to the most general linear transformation, to an affine set of axes. That is not the case for the quantity ψ , however; this kind of quantity belongs to a representation of the rotation group which cannot be extended to the affine group. Consequently we cannot introduce components of ψ relative to an arbitrary coordinate system in general relativity as we can for the electromagnetic potential and field strengths. We must rather describe the metric at a point P by local Cartesian axes $e(\alpha)$ instead of by the g_{pq} . The wave field has definite components $\psi_1^+, \psi_2^+; \psi_1^-, \psi_2^-$ [full Dirac theory] relative to such axes, and we know how they transform on transition to any other Cartesian axes in P . The laws shall naturally be invariant under arbitrary rotation of the axes in P , and the axes at different points can be rotated independently of each other; they are in no way bound together."

representation of $\mathbb{S}\mathbb{O}^+(1, 3)$. Weyl acts on the two complex components of ζ with a 2×2 complex representation of a group³³ one can call

$$\mathbb{W}(2, \mathbb{C}) = \{g \in \mathbb{GL}(2, \mathbb{C}) : |\det g| = 1\},$$

which is slightly larger than $\mathbb{SL}(2, \mathbb{C})$.

The Lorentz group is not enough for the ‘global’ needs of the holonomic bases ∂_μ and dx^σ , which require all of $\mathbb{GL}(4, \mathbb{R})$. To adapt to the spinorial requirements which impose $\mathbb{S}\mathbb{O}^+(1, 3)$, Weyl introduced the (anholonomic) tetrads e^μ —his *Achsenkreuze*—whose orthogonality is preserved by the Lorentz group.³⁴ The *Achsenkreuze* and their orthogonality are more primitive than the metric

$$g = \eta_{\mu\nu} e^\mu \otimes e^\nu = \eta_{\mu\nu} (\Lambda_\sigma^\mu e^\sigma) \otimes (\Lambda_\tau^\nu e^\tau),$$

which derives from them by construction; the (second) equality holds for all transformations $\Lambda \in \mathbb{S}\mathbb{O}^+(1, 3)$.

2.4 The displacement of spinors

Spinors have directions, whose differentiation and parallel propagation have to take account of the spacetime curvature of the region crossed. We have seen that in 1929 Weyl abandons the congruent transport of his first gauge theory, to return to the old Levi-Civita transport of Einstein’s theory, generated by the Lie algebra $\mathfrak{o}(1, 3) = \text{Lie } \mathbb{S}\mathbb{O}^+(1, 3)$.

The spin connection ω is a one-form which, applied to a direction $\dot{\gamma} \in T_a M$, yields a generator $\omega_{\dot{\gamma}} = \langle \omega, \dot{\gamma} \rangle : \mathbb{C}_a^2 \rightarrow \mathbb{C}_b^2$ of transport belonging to the Lie algebra $\mathfrak{w}(2, \mathbb{C}) = \text{Lie } \mathbb{W}(2, \mathbb{C})$. Subtracting the difference $\delta|\psi_b\rangle = \omega_{\dot{\gamma}}|\psi_a\rangle$ from the initial spinor $|\psi_a\rangle$ we obtain the transported spinor $|\psi_b\rangle = (\mathbb{1}_2 - \omega_{\dot{\gamma}})|\psi_a\rangle$, where $\mathbb{1}_2$ is the unit matrix in \mathbb{C}^2 . But numbers are easier to compare than spinors at different points, so it is best to take components $\psi^k = \langle \varphi^k | \psi \rangle$:

$$\psi_b^k = \psi_a^k - \omega_{\mu j}^k \dot{\gamma}^\mu \psi_a^j = \psi_a^k - (\omega_{\dot{\gamma}})^k_j \psi_a^j,$$

$\langle \varphi^k |$ being an orthonormal basis, $k = 1, 2$.

Even if the gauge group \mathcal{G} given by (6) changes the direction of the whole wavefunction in Hilbert space, the structure group $\mathbb{U}(1)$ only changes the phase of the spinor, not its direction. But generalisation to a non-Abelian³⁵ structure group that also changes the *directions* of spinors is natural, and was accomplished by Yang & Mills (1954), as we shall see in §3.

³³Weyl (1929b) p. 333: “man beschränke sich auf solche lineare Transformationen U von ψ_1, ψ_2 , deren Determinante den absoluten Betrag 1 hat.”

³⁴Weyl (1931b) p. 195: “Ferner bedarf man in der allgemeinen Relativitätstheorie an jeder Weltstelle P eines aus vier Grundvektoren in P bestehenden normalen Achsenkreuzes, um die Metrik in P festzulegen und relativ dazu die Wellengröße ψ durch ihre vier [full Dirac theory again] Komponenten ψ_ρ beschreiben zu können; die gleichberechtigten normalen Achsenkreuze in einem Punkte gehen durch die Lorentztransformationen auseinander hervor.”

³⁵Even if the structure group $\mathbb{U}(1)$ makes W29 an Abelian gauge theory, we have seen that parallel transport against its curved background requires $\mathbb{W}(2, \mathbb{C})$ as well; as a gauge theory it is Abelian, but it is also a non-Abelian theory of spinors on curved spacetime.

2.5 Three gauge arguments

2.5.1 The inherited inexact connection

To reach, from Weyl's first gauge theory, the compensation of (6) by (3), with³⁶

$$(10) \quad d \mapsto D = d + iA$$

(or $\partial_\mu \rightarrow D_\mu = \partial_\mu + iA_\mu$) we only need a handful of principles. In §2.1 we saw that Einstein's objection, reinforced by the absolute length h/mc , favours (6) over (5). The addition of the electromagnetic potential to momentum (and hence to the derivative) comes from analytical mechanics, where³⁷

$$(11) \quad p \mapsto p + A$$

(or $p_\mu \mapsto p_\mu + A_\mu$). Together (7) and (11) give (10), the one-form A being the one that figures, by a natural identification, in (3). The compensation of (6) by (3) can be seen in the Lagrangian

$$\mathcal{L} = \bar{\psi} \sigma^\mu D_\mu \psi = \bar{\psi}' \sigma^\mu (D_\mu - i\partial_\mu \lambda) \psi'.$$

The inexact connection A and its nonvanishing curvature (1) were there long before it even made sense to apply (6).³⁸

I'd say there's something of a 'gauge argument' here already. But Weyl has another gauge argument,³⁹ which extracts electromagnetism from the $\mathbb{U}(1)$ freedom left by the $h : \mathbb{W}(2, \mathbb{C}) \rightarrow \mathbb{SO}^+(1, 3)$ homomorphism and expressed by

$$(12) \quad h(e^{i\lambda} g) = h(g) \in \mathbb{SO}^+(1, 3),$$

$g \in \mathbb{W}(2, \mathbb{C})$.⁴⁰

2.5.2 How Weyl extracts an inexact connection from $\mathbb{U}(1)$

$\mathbb{SO}^+(1, 3) = G$ and $\mathbb{W}(2, \mathbb{C}) = G'$ are just 'structure' groups, acting at a generic spacetime point. What about the corresponding gauge groups \mathcal{G} , \mathcal{G}' acting on all of spacetime M ? In special relativity "there's just a single tetrad"; so there's just one $\mathbb{SO}^+(1, 3) = G = \mathcal{G}$, one $\mathbb{W}(2, \mathbb{C}) = G' = \mathcal{G}'$, and above all one $e^{i\lambda}$.⁴¹ But with

³⁶See Weyl (1931b) p. 89.

³⁷See Weyl (1931b) p. 88.

³⁸Weyl (1929c) p. 283 simply *provides* the inexact electromagnetic potential φ_α (of the *nonvanishing* field $F_{\alpha\beta}$).

³⁹Weyl (1929b) p. 348, Weyl (1929c) p. 291

⁴⁰Weyl (1929c) p. 291: "It is my firm conviction that we must seek the origin of the electromagnetic field in another direction. We have already mentioned that it is impossible to connect the transformations of the ψ in a unique manner with the rotations of the axis system; however we may attempt to accomplish this by means of invariants which can be used as constituents of an action quantity we always find that there remains an arbitrary "gauge factor" $e^{i\lambda}$. Hence the local axis-system does not determine the components of ψ uniquely, but only within such a factor of absolute magnitude 1." See also the text quoted in footnote 22.

⁴¹Weyl (1929b) p. 348: "In der speziellen Relativitätstheorie muß man diesen Eichfaktor als eine Konstante ansehen, weil wir hier ein einziges, nicht an einen Punkt gebundenes Achsenkreuz haben." Weyl (1929c) p. 291: "In the special theory of relativity, in which the axis system is not tied up to any particular point, this factor is a constant."

spacetime curvature the tetrad varies,⁴² and so does λ . This could mean the following:⁴³ Only a *flat* $\mathfrak{o}(1, 3)$ -valued connection \mathcal{A} allows the assignment of the *same* tetrad to distant points—only with flatness can there be *global* constancy or ‘sameness.’ With curvature it becomes meaningless to say that tetrads at distant points are the same. Where tetrads cannot remain constant, one has to suppose they *vary*. A flat real-valued phase connection A alongside a curved \mathcal{A} can of course be countenanced, but it is in the spirit of Weyl’s argument for both to be flat or both curved. So if the tetrad varies, λ might as well too.

The group homomorphism h determines the Lie algebra homomorphism

$$\mathfrak{h} : \mathfrak{w}(2, \mathbb{C}) \rightarrow \mathfrak{o}(1, 3),$$

where the Lie algebra $\mathfrak{w}(2, \mathbb{C}) = \text{Lie } \mathbb{W}(2, \mathbb{C})$ is the direct sum $\mathfrak{sl}(2, \mathbb{C}) \oplus i\mathbb{R}\mathbb{1}_2$, and $i\mathbb{R} = \text{Lie } \mathbb{U}(1)$. Doing away with the additive freedom λ (or rather $i\lambda\mathbb{1}_2$) we’re left with the isomorphism between $\mathfrak{w}(2, \mathbb{C})/i\mathbb{R}\mathbb{1}_2 = \mathfrak{sl}(2, \mathbb{C})$ and $\mathfrak{o}(1, 3)$. Instead of the phase $e^{i\lambda} \in \mathbb{U}(1)$ we have $i\lambda\mathbb{1}_2 \in i\mathbb{R}\mathbb{1}_2$; instead of $\mathbb{U}(1)$ we have the Lie algebra $i\mathbb{R}\mathbb{1}_2$; and instead of (12),

$$\mathfrak{h}(\gamma \oplus i\lambda\mathbb{1}_2) = \mathfrak{h}(\gamma) \in \mathfrak{o}(1, 3),$$

$\gamma \in \mathfrak{w}(2, \mathbb{C})$.⁴⁴

The additive freedom $i\lambda\mathbb{1}_2$ is in the Lie algebra $\mathfrak{w}(2, \mathbb{C})$ where the spin connection \mathfrak{A} has its values; and connections are there to generate parallel transport—in a *direction*.⁴⁵ A direction $\dot{\gamma} \in T_a M$ will therefore characterise the propagation of λ , whose infinitesimal variation $\delta\lambda$ has to be linear in λ and in $\dot{\gamma}$. The object needed is a one-form; applied to the direction $\dot{\gamma}$ it yields the infinitesimal generator $\langle A, \dot{\gamma} \rangle \in \mathbb{R}$, which then multiplies λ to produce the increment $\delta\lambda = \lambda \langle A, \dot{\gamma} \rangle$. So there’s a connection for

⁴²The gauge groups become infinite-dimensional. Weyl (1929b) p. 348: “Anders in der allgemeinen Relativitätstheorie: jeder Punkt hat sein eigenes Achsenkreuz und darum auch seinen eigenen willkürlichen Eichfaktor; dadurch, daß man die starre Bindung der Achsenkreuze in verschiedenen Punkten aufhebt, wird der Eichfaktor notwendig zu einer willkürlichen Ortsfunktion.” Weyl (1929c) p. 291: “But it is otherwise in the general theory of relativity when we remove the restriction binding the local axis-systems to each other; we cannot avoid allowing the gauge factor to depend arbitrarily on position.”

⁴³Here I am indebted to Johannes Huisman.

⁴⁴Weyl (1929b) p. 348: “Dann ist aber auch die infinitesimale lineare Transformation dE der ψ , welche der infinitesimalen Drehung $d\gamma$ entspricht, nicht vollständig festgelegt, sondern dE kann um ein beliebiges rein imaginäres Multiplum $i \cdot df$ der Einheitsmatrix vermehrt werden.” Weyl (1929c) p. 291: “Then there remains in the infinitesimal linear transformation dE of ψ , which corresponds to the given infinitesimal rotation of the axis-system, an arbitrary additive term $+id\varphi \cdot 1$.”

⁴⁵Weyl (1929b) p. 348: “Zur eindeutigen Festlegung des kovarianten Differentials $\delta\psi$ von ψ hat man also außer der Metrik in der Umgebung des Punktes P auch ein solches df für jedes von P ausgehende Linienelement $\overrightarrow{PP'}$ ($= dx$) nötig. Damit $\delta\psi$ nach wie vor linear von dx abhängt, muß

$$df = f_p(dx)^p$$

eine Linearform in den Komponenten des Linienelements sein. Ersetzt man ψ durch $e^{i\lambda}$, so muß man sogleich, wie aus der Formel für das kovariante Differential hervorgeht, df ersetzen durch $df - d\lambda$.” Weyl (1929c, p. 291): “The complete determination of the covariant differential $\delta\psi$ of ψ requires that such a $d\varphi$ be given. But it must depend linearly on the displacement PP' : $d\varphi = \varphi_p(dx)^p$, if $\delta\psi$ shall depend linearly on the displacement. On altering ψ by multiplying it by the gauge factor $e^{i\lambda}$ we must at the same time replace $d\varphi$ by $d\varphi - d\lambda$ as is immediately seen from this formula of the covariant differential.” Weyl’s notation is confusing: whereas the one-form $d\lambda$ (which is a differential) is necessarily exact, df and $d\varphi$ (my A) aren’t.

tetrads, another for spinors, and a third one—*A*—for the residual $\mathbb{U}(1)$ freedom caught ‘in between’ tetrads and spinors.

The whole point of allowing the propagation of λ to depend on direction is to admit anholonomies. So the curvature (1) of *A* will not necessarily vanish. In (1), *A* and (2) Weyl again⁴⁶ saw⁴⁷ the electromagnetic field, its potential and Maxwell’s two homogeneous equations.

2.5.3 How the standard gauge argument extracts an exact connection from $\mathbb{U}(1)$

An alternative logic,⁴⁸ which is claimed to produce electromagnetism from the indifference of \mathcal{L} to (6), is popular: The local phase transformation gives rise to a new Lagrangian

$$\mathcal{L}' = \bar{\psi}' \sigma^\mu \partial_\mu \psi' = \bar{\psi} e^{-i\lambda} \sigma^\mu (e^{i\lambda} \partial_\mu + e^{i\lambda} i \partial_\mu \lambda) \psi = \bar{\psi} \sigma^\mu (\partial_\mu + i \partial_\mu \lambda) \psi.$$

As the components $i \partial_\mu \lambda$ of $i d\lambda$ account for the difference, invariance will be restored if the same term is subtracted, thus producing the covariant derivative $D' = d - i d\lambda$ and the invariant Lagrangian

$$\hat{\mathcal{L}} = \bar{\psi}' \sigma^\mu D'_\mu \psi'.$$

It is then argued that an interaction $F = d^2 \zeta$ is thereby deduced, whose potential *A* is $d\zeta$. But since d^2 vanishes the interaction does too, as has often been pointed out.⁴⁹

3 Yang-Mills theory

Here the structure group $\mathbb{S}\mathbb{U}(N)$ replaces $\mathbb{U}(1)$. Instead of the connection $A = A_\mu dx^\mu$ with values in the Lie algebra $i\mathbb{R}$ we have the Yang-Mills connection

$$\mathcal{A} = \sum_{\mu=0}^3 A_\mu \otimes dx^\mu = \sum_{\mu=0}^3 \sum_{k=1}^N A_\mu^k \mathbf{T}_k \otimes dx^\mu,$$

with values in the Lie algebra $\mathfrak{su}(N)$ spanned by $\mathbf{T}_1, \dots, \mathbf{T}_N$. Applied to a transport direction $\dot{\gamma} \in T_a M$, the connection \mathcal{A} gives the infinitesimal generator

$$\langle \mathcal{A}, \dot{\gamma} \rangle = A_\mu \dot{\gamma}^\mu = A_\mu^k \dot{\gamma}^\mu \mathbf{T}_k = \mathbf{T} : \mathbb{C}_a^N \rightarrow \mathbb{C}_b^N$$

which turns the initial spinor $|\Psi_a\rangle \in \mathbb{C}_a^N$ into the increment $\delta|\Psi_b\rangle = \mathbf{T}|\Psi_a\rangle \in \mathbb{C}_b^N$. In components $\Psi^k = \langle \Phi^k | \Psi \rangle$ we can write $\Psi_b^k = \Psi_a^k - T_l^k \Psi_a^l$. The curvature

$$\mathcal{F} = d\mathcal{A} = \mathbf{T}_k \otimes (dA_\mu^k) \wedge dx^\mu = \frac{1}{2} \mathcal{F}_{\mu\nu}^k \mathbf{T}_k \otimes dx^\mu \wedge dx^\nu = \frac{1}{2} \mathcal{F}_{\mu\nu} \otimes dx^\mu \wedge dx^\nu$$

is a two-form with values in $\mathfrak{su}(N)$.

⁴⁶See §1.1.

⁴⁷Weyl (1929b) p. 349, Weyl (1929c) p. 291-2

⁴⁸See for instance Yang & Mills (1954) p. 192, Sakurai (1967) p. 16, Aitchison & Hey (1982) p. 176, Mandl & Shaw (1984) p. 263, Göckeler & Schücker (1987) p. 48, Ramond (1990) pp. 183-91, Ryder (1996) p. 93, O’Raifeartaigh (1997) p. 118.

⁴⁹Auyang (1995) p. 58, Brown (1999) pp. 50-3, Teller (2000) pp. S468-9, Lyre (2001, 2004a,b), Healey (2001) p. 438, Martin (2002) p. S229, Martin (2003) p. 45, Catren (2008) pp. 512, 520

4 Final remarks

Summing up, we obtain Yang-Mills theory (YM) by putting together the following ingredients:

1. GR: *general relativity*
2. GJ: *geometrical justice*
3. MW: *matter wave*
4. EO: *avoid Einstein's objection*
5. NA: *non-Abelian structure group*;

W18 & MW & EO & NA \rightarrow YM.

I thank Ermenegildo Caccese, Johannes Huisman, Marc Lachièze-Rey, Thierry Levasseur and Jean-Philippe Nicolas for many valuable conversations and clarifications.

References

- Afriat, A. (2009) "How Weyl stumbled across electricity while pursuing mathematical justice" *Studies in History and Philosophy of Modern Physics* **40**, 20-5
- Aitchison, I. J. R. and A. J. G. Hey (1982) *Gauge theories in particle physics*, Hilger, Bristol
- Auyang, S. Y. (1995) *How is quantum field theory possible?*, Oxford University Press
- Brading, K. (2002) "Which symmetry? Noether, Weyl, and the conservation of electric charge" *Studies in History and Philosophy of Modern Physics* **33**, 3-22
- Broglie, L. de (1924) *Recherches sur la théorie des quanta*, Thèse, Paris
- Brown, H. (1999) "Aspects of objectivity in quantum mechanics" pp. 45-70 in J. Butterfield and C. Pagonis (editors) *From physics to philosophy*, Cambridge University Press
- Catren, G. (2008) "Geometric foundations of classical Yang-Mills theory" *Studies in History and Philosophy of Modern Physics* **39**, 511-31
- Einstein, A. (1916) "Grundlage der allgemeinen Relativitätstheorie" *Annalen der Physik* **49**, 769-822
- Göckeler, M. and T. Schücker (1987) *Differential geometry, gauge theories, and gravity*, Cambridge University Press
- Healey, R. (2001) "On the reality of gauge potentials" *Philosophy of Science* **68**, 432-55

- Levi-Civita, T. (1917) “Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana” *Rendiconti del Circolo matematico di Palermo* **42**, 173-205
- Lyre, H. (2001) “The principles of gauging” *Philosophy of Science* **68**, S371-81
- Lyre, H. (2004a) *Lokale Symmetrien und Wirklichkeit: eine Naturphilosophische Studie über Eichtheorien und Strukturenrealismus*, Mentis, Paderborn
- Lyre, H. (2004b) “Holism and structuralism in $U(1)$ gauge theory” *Studies in History and Philosophy of Modern Physics* **35**, 643-70
- Mandl, F. and G. Shaw (1984) *Quantum field theory*, Wiley, Chichester
- Martin, C. (2002) “Gauge principles, gauge arguments and the logic of nature” *Philosophy of Science* **69**, S221-34
- Martin, C. (2003) “On continuous symmetries and the foundations of modern physics” pp. 29-60 in K. Brading and E. Castellani (editors) *Symmetries in physics*, Cambridge University Press
- O’Raifeartaigh, L. (1997) *The dawning of gauge theory*, Princeton University Press
- O’Raifeartaigh, L. and N. Straumann (2000) “Gauge theory: historical origins and some modern developments” *Reviews of Modern Physics* **72**, 1-23
- Pauli, W. (1979) *Wissenschaftlicher Briefwechsel, Band I: 1919-1929*, Springer, Berlin
- Ramond, P. (1990) *Field theory: a modern primer*, Westview Press, Boulder
- Ryckman, T. (2005) *The reign of relativity: philosophy in physics 1915-1925*, Oxford University Press
- Ryder, L. (1996) *Quantum field theory*, Cambridge University Press
- Sakurai, J. J. (1967) *Advanced quantum mechanics*, Addison-Wesley, Reading
- Scholz, E. (2005) “Local spinor structures in V. Fock’s and H. Weyl’s work on the Dirac equation (1929)” pp. 284-301 in D. Flament *et al.* (editors) *Géométrie au vingtième siècle, 1930-2000*, Hermann, Paris
- Schrödinger, E. (1926) “Quantisierung als Eigenwertproblem (erste Mitteilung)” *Annalen der Physik* **79**, 361-76
- Seelig, K. (1960) *Albert Einstein*, Europa Verlag, Zurich
- Straumann, N. (1987) “Zum Ursprung der Eichtheorien bei Hermann Weyl” *Physikalische Blätter* **43**, 414-21
- Teller, P. (2000) “The gauge argument” *Philosophy of Science* **67**, S466-81
- Weyl, H. (1918) “Gravitation und Elektrizität” pp. 147-59 in *Das Relativitätsprinzip*, Teubner, Stuttgart, 1990

- Weyl, H. (1927) *Philosophie der Mathematik und Naturwissenschaft*, Oldenbourg, Munich
- Weyl, H. (1929a) "Gravitation and the electron" *Proceedings of the National academy of sciences, USA* **15**, 323-34
- Weyl, H. (1929b) "Elektron und Gravitation" *Zeitschrift für Physik* **56**, 330-52
- Weyl, H. (1929c) "Gravitation and the electron" *The Rice Institute Pamphlet* **16**, 280-95
- Weyl, H. (1931a) "Geometrie und Physik" *Die Naturwissenschaften* **19**, 49-58
- Weyl, H. (1931b) *Gruppentheorie und Quantenmechanik*, Hirzel, Leipzig
- Yang, C. N. and R. Mills (1954) "Conservation of isotopic spin and isotopic gauge invariance" *Physical Review* **96**, 191-5