1. Introduction

It is almost always claimed\(^1\) that Weyl deliberately unified gravitation and electricity in the rectification of general relativity he attempted in 1918. In fact the unification, as Bergia (1993) and Ryckman (2005) have pointed out and a couple of passages\(^2\) show, was the unintended outcome of a priori prejudice. But what prejudice?

The evidence suggests that the theory came straight out of Weyl’s sense of mathematical ‘justice,’ which led him to put the direction and length of a vector on an equal footing. Levi-Civita (1917) had discovered that the parallel transport determined by Einstein’s covariant derivative was not integrable—while length, far from depending on the path taken, remained unaltered. For Weyl this was unfair: both features deserved the same treatment.\(^3\) He remedied with a connection that made congruent transport (of length) just as path-dependent as parallel transport. This ‘total’ connection restored justice through a length connection it included, an inexact one-form Weyl could not help identifying with the electromagnetic potential \(A\), whose 4-curl \(F = dA\) being closed (for \(dF = d^2A\) vanishes identically), provides Maxwell’s two homogeneous equations, not to mention source-free versions of the other two (up to Hodge duality at any rate). Electromagnetism away from sources thus came, quite unexpectedly, out of Weyl’s surprising sense of mathematical justice.

Admittedly there were also intimations,\(^4\) from the beginning, announcing an ‘infinitesimal’ agenda of sorts; but it was largely unmotivated back then, and too vague to produce the theory on its own—in fact it may even have been suggested by the theory. The agenda would take shape over the next years, acquiring justification and grounding; Ryckman has found roots in Husserl, and connected, or even identified it with a ‘teenthistic’ opposed to distant comparisons. But his compelling reconstruction of the infinitesimal programme and its philosophical background rests largely on a text (footnote 13) from a subsequent ‘context of justification.’

‘Mathematical justice’ has, in the context of discovery, a more conspicuous (though sometimes, as we shall see in §3.1, thinly disguised) presence than the infinitesimal agenda. It is also logically stronger, being enough—together with a couple of simple and natural operations—to yield all of source-free electromagnetism. So I contend that what was really at work in the spring of 1918, what effectively gave rise to Weyl’s theory of gravitation and electricity, was the equal rights of direction and length.

\(^3\) See Weyl (1918a) p.148 “und es ist dann […] als integriibel herausgestellt hat.”
\(^4\) Weyl (1918a) p.148 “In der oben […] Element erhalten,” for instance; or (same page): “Eine Wahrhafte […] benachbarten kennen.” Further adumbrations—such as the title: Reine Infinitesimalgeometrie—can be found in Weyl (1918b), which came out about half a year after the communication of Weyl (1918a).
2. Background: Einstein, Levi-Civita

We can begin with aspects of Einstein’s theory of gravitation, since Weyl’s theory grew out of it. What interests us above all is affine structure, given by the Christoffel symbols $\Gamma^a_{bc}$. Through the geodesic equation

$$\frac{d^2x^a}{ds^2} + \sum_{b,c=0}^3 \Gamma^a_{bc} \frac{dx^b}{ds} \frac{dx^c}{ds} = 0$$

and the wordlines satisfying it, the Christoffel symbols provide a notion of (parametrised) straightness, of inertial, unaccelerated motion, of free fall.

The left-hand side of (1) gives the components $\langle dx^a, \nabla_a \sigma \rangle$ of the covariant derivative $\nabla_a \sigma$ of the vector $\sigma$ with components $dx^a / ds = \langle dx^a, \sigma \rangle$, in the direction $\sigma$ tangent to the worldline $\sigma : I \rightarrow M : s \mapsto \sigma(s)$ with coordinates $\sigma^a(s) = x^a(\sigma(s))$, where $I$ is an appropriate interval and $M$ the differential manifold representing the universe; $a = 0,\ldots,3$. The Christoffel symbols are related to $\nabla$ by $\Gamma^a_{bc} = \langle dx^a, \nabla_b \partial_c \rangle$, where the basis vectors $\partial_a = \partial_a(x) = \partial / \partial x^a$ are tangent to the coordinate lines of the system $x^a$. Einstein only appears to have explored the infinitesimal behaviour of the parallel transport determined by his covariant derivative. It was Levi-Civita (1917) who first understood that if $\nabla$ vanishes, as in Einstein’s theory, the direction of the vector $V_s \in T_{\alpha(s)} M$ transported according to $\nabla_\sigma V_s = 0$ depends on the path $\sigma$ taken—whereas the squared length $\langle \sigma, V \rangle$ remains constant along $\sigma$, for $\nabla g = 0 = \nabla_{\sigma} V_s$ means that $dl_s / ds = \nabla_{\sigma} l_s$ vanishes.

3. The emergence of Weyl’s theory

3.1 The equal rights of direction and length

Weyl felt that as parallel transport depended on the path taken, congruent transport ought to as well. In fact his generalisation of Einstein’s theory appears to have been almost entirely determined by the intention of putting direction and length on an equal footing. The following table\(^5\)—parts of which may for the time being be more intelligible than others—outlines Weyl’s programme.

<table>
<thead>
<tr>
<th>DIRECTION</th>
<th>LENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>coordinates (up to gauge)</td>
<td>gauge</td>
</tr>
<tr>
<td>parallel transport</td>
<td>congruent transport</td>
</tr>
<tr>
<td>gravitation</td>
<td>electricity</td>
</tr>
<tr>
<td>Levi-Civita connection $\Gamma^a_{bc}$</td>
<td>length connection $A$</td>
</tr>
<tr>
<td>$\delta V^a = -\Gamma^a_{bc} X^b V^c$</td>
<td>$\delta l = -\langle \alpha, X \rangle = -A X^b l$</td>
</tr>
<tr>
<td>Riemann curvature $R^a_{bcd}$ (of $\Gamma^a_{bc}$)</td>
<td>length curvature $F = dA$</td>
</tr>
<tr>
<td>geodesic coordinates $y^a$ (at $P$): $\Gamma^a_{bc} = 0$</td>
<td>geod. gauge (at $P$): $A' = A + d\lambda = 0$</td>
</tr>
<tr>
<td>equivalence principle: $\bar{x}^a = \Gamma^a_{bc} \bar{x}^b \bar{x}^c \mapsto \bar{y}^a$</td>
<td>equival. principle: $\alpha = -lA \mapsto \alpha' = 0$</td>
</tr>
</tbody>
</table>

\(^5\) Parts of it were inspired by Coleman and Korté (2001) pp.204-5, 211-12.
A few words about “coordinates (up to gauge).” The parallel between coordinates and gauge, which Weyl draws over and over, can be seen as a parallel between direction and length. For surely Weyl does not mean “coordinates including gauge—as opposed to gauge,” since that would be redundant. And up to gauge, coordinates provide no more than direction: The coordinates \( x^a \) assign to each event \( P \in M \) a basis \( \partial_a \in T_p M \), and a dual basis \( dx^a = g^b(\partial_a) = g(\partial_a, \cdot) \in T^* M \) providing the components \( V^a = \langle dx^a, V \rangle \) of any vector \( V \in T_p M \); \( a = 0, \ldots, 3 \). The recalibration\(^6\) \( g \mapsto e^{\lambda g} \) induces a transformation \( V \mapsto e^V \), or \( V^a \mapsto e^V a^a \), through \( e^{\lambda g}(V, V) = g(\gamma \partial_a, \gamma \partial_a)V^a V^b = g(\partial_a, \partial_a)e^V a^a e^V b \). Direction, given by the ratios \( e^V a^a : e^V b^b : e^V c^c = V^a : V^b : V^c \), remains unaffected.

Weyl clearly distinguishes between a ‘stretch’ (like a stretch of road) and its numerical length, determined by the gauge chosen. Just as a direction (all )\( e^V \lambda \) is ‘expanded’ with respect to a coordinate system, which provides its numerical representation (the ratios \( V^a : \cdots : V^c \)), a stretch gets ‘expanded’ in a gauge, which likewise gives a numerical representation, the (squared) length \( l = e^{\lambda g}(V, V) \). The rest of the table should in due course become clearer. Let us now see how the inexact one-form \( A \) from which so much of electromagnetism can be derived, emerges from the equal rights of direction and length.

### 3.2 Electromagnetism from equal rights

Weyl calls a manifold \( M \) affinely connected if the tangent space \( T_p M \) at every point \( P \in M \) is connected to all the neighbouring tangent spaces \( T_p M \) by a mapping \( \Xi : T_p M \to T_p M : V_p \mapsto V'_p = \Xi V_p \), linear both in the ‘main’ argument \( V_p \in T_p M \) and in the (short) directional argument \( X = P' - P \), where \( P' \) (being near \( P \)) and hence \( X \) are viewed as lying in \( T_p M \). Being linear, \( \Xi \) will be represented by a matrix, in fact by \( \Xi = \langle dx^a, V \rangle = \langle dx^a, \Xi \partial_a \rangle = \langle dx^a, \Xi \partial_a \partial_a \rangle \). Weyl specifically refers to the components \( \delta V^a = \langle dx^a, V'_p \rangle - \langle dx^a, V_p \rangle \), requiring them to be linear in the components \( X^b \) and \( V'_p = \langle dx_p, V_p \rangle \). The bilinear function \( \Gamma^a(\{X^b\}, \{V^c\}) = \delta V^a \) will be a matrix, represented by \( \Gamma_{bc}^a \); the small difference \( \delta V^a \) is therefore \( -\Gamma_{bc}^a X^b V^c \).

With respect to the geodesic coordinates \( y^a \) which make \( \Gamma_{bc}^a = \Gamma_{bc}^a X^b = \langle dy^a, V \partial_a \rangle \) and \( \delta V^a \) vanish, leaving the components \( V^a \) unchanged, \( \Xi^a \) becomes the identity matrix \( \delta a = \langle dy^a, \Xi \partial_a \rangle \leftrightarrow \text{diag}(1,1,1,1) \). Physically this has to do with the equivalence principle, according to which a gravitational field \( \Gamma_{bc}^a \) can always be eliminated or generated at \( P \) by an appropriate choice of coordinates.

With equal rights in mind Weyl turns to length, using the very same scheme. To clarify his procedure we can take just a single component of the difference \( \{\delta V^0, \ldots, \delta V^3\} \), calling it \( \delta l \) (this will be the ‘squared-length-difference scalar’).\(^9\) The upper index of \( \Gamma_{bc}^a \) accordingly disappears, leaving \( \delta l = -\Gamma_{bc}^a X^b V^c \) (one can perhaps think of the hybrid, intermediate

---


\(^{7}\) Cf. Weyl (1918) p.149: “Wird die Mannigfaltigkeit […] Verhältnis nach festgelegt.”

\(^{8}\) The convenient ‘exponential’ recalibration is not used by Weyl.

\(^{9}\) Weyl appears to use \( d \) and \( \delta \) interchangeably, and \( d \) in a way—see the next footnote—that is not only unusual today, but was even then. He does not distinguish between the scalar representing the difference in squared length, and the corresponding one-form (as we would call it). The distinction nonetheless seems useful.
connection $\Gamma_{bc}$ as being something like $\langle A, \nabla \partial, \partial \rangle$). If we now take a single component of the main argument $\{V^0, \ldots, V^3\}$, calling it $l$ (this will be the squared length), the second index of $\Gamma_{bc}$ disappears as well, and we are left with $\partial l = -\Gamma_b^c X^b l$, where $\Gamma_b^c = \langle A, \partial_b \partial_c \rangle$ are the components of a one-form, denoted $A$ with electricity in mind.

But this is not really Weyl’s argument, which is better conveyed as follows. The object $A$ generating the squared-length-difference scalar $\partial l$ has to be linear in the squared length $l$ and the direction $X$. A linear function $A(l, X) = \partial l$ of a scalar $l$ and vector $X$ yielding a scalar $\partial l$ will be a one-form.\footnote{Weyl in fact writes $dl = -d\varphi$, whereas I write $\alpha = -\lambda A$. The misleading $d$’s cannot be understood globally—or even locally, in the theory of gravitation and electricity, in which $F = d^2 \varphi$ will be the Faraday two-form: where $d\varphi$ is closed, in other words the differential (even only locally) of a function $\varphi$, there would be no electromagnetism.} $\partial l = -\langle \alpha, X \rangle = -\alpha X^b = -\langle A, X \rangle l = -A_{X^b l}$, where $\alpha$ is the squared-length-difference one-form. An exact one-form $A = d\mu$ would make congruent transfer integrable, removing the dependence of the recalibration $\exp(\int A) = \exp(\int d\mu) = e^{i\gamma}$ on the path $\gamma : [0,1] \to M$, where $\Delta \mu = \mu_1 - \mu_0$ is the difference between the values $\mu_1 = \mu(P_i)$ and $\mu_0 = \mu(P_o)$ of $\mu$ at $P_i = \gamma(1)$ and $P_o = \gamma(0)$. Mathematical justice therefore demands that $A$ be inexact; so the curl $F = dA$ cannot vanish identically.

Confirmation that $A$ has to be one-form, possibly inexact, is provided by Weyl’s requirement that the squared-length-difference one-form $\alpha = -\lambda A$ be eliminable\footnote{See Weyl (1923) p.122; and Lyre (2004), who speaks of a generalised equivalence principle.} at any point $P$ by recalibration. As $l$ is given (and does not vanish), this amounts to $A + d\lambda = 0$ at $P$, where the gauge $\lambda$ is geodesic. Since $d\lambda$ is a one-form, $A$ must be one too. Though $d\lambda$ is exact, Weyl only asks that it cancel $A$ at $P$—so $A$ needn’t even be closed, or locally exact.

With $F = dA$ and its consequence $dF = 0$ before him Weyl could not help seeing the electromagnetic four-potential $A$, the Faraday two-form $F = \frac{1}{2} F_{ab} dx^a \wedge dx^b = dA$ (which vanishes wherever $A$ is closed) and Maxwell’s two homogeneous equations, expressed by $dF = 0$—not to mention an electromagnetic ‘equivalence principle’ according to which the squared-length-difference scalar $\partial l$ and one-form $\alpha$, as well as the electromagnetic potential $A$, can be eliminated or generated at a point by the differential $d\lambda$ of an appropriate gauge function $\lambda$. In coordinates $F_{ab} = F(\partial_a \partial_b, \partial_c) = \partial_a^c F_{bc} - \partial_b^c F_{ac} + \partial_c^d F_{ab}$; the vanishing three-form $dF = \frac{1}{2} dF_{bc} \wedge dx^b \wedge dx^c$ has components $dF(\partial_a \partial_b \partial_c) = \partial_a^c F_{bc} + \partial_b^c F_{ac} + \partial_c^d F_{ab}$. Maxwell’s other two equations are obtained, in ‘source-free’ form, by setting $d^2 F$ equal to zero, where $^* F$ is the Hodge dual of the Faraday two-form. Electromagnetism thus emerged, altogether unexpectedly, from the equal rights of direction and length.

3.3 The illegitimacy of distant comparisons

Weyl has another \textit{a priori} prejudice, rooted, as Ryckman (2005) has cogently argued, in Husserl’s transcendental phenomenology. It is expressed in two similar passages,\footnote{Weyl (1927) p.98: Erkennt man neben dem physischen […] zwischen Ebene und Fläche.} \footnote{Weyl (1931) p.52: Die Philosophen mögen recht haben, daß unser […] zwischen Fläche und Ebene.} which roughly say: As the curvature $R(P)$ is subtle and hard to perceive directly, a “cognizing ego” at the “ego center” $P \in M$ takes itself to be immersed in the ‘psychologically privileged’
tangent space $T_pM$. The universe $M$ resembles $T_pM$ in the immediate vicinity $U$ of $P$, where they practically coincide, and ‘cover’ one another. Beyond $U$ the relation between $M$ and the ‘intuitive’ space $T_pM$ grows looser, as the universe goes its own way, bending as the energy-momentum tensor $T$ varies. Ryckman writes (p.148) that

Weyl restricted the homogeneous space of phenomenological intuition, the locus of phenomenological Evidenz, to what is given at, or neighboring, the cognizing ego [...]. But in any case, by delimiting what Husserl termed “the sharply illuminated circle of perfect givenness,” the domain of “eidetic vision,” to the infinitely small homogeneous space of intuition surrounding the “ego-center” [...].

This restriction or delimitation can be understood in two ways: directly, in terms of the limitations of our senses, and of an accordingly circumscribed domain of sensory access or “eidetic vision”; or more mathematically, as follows: The cognizing ego attaches a kind of intuitive ‘certainty’ to all of $T_pM$, which, being flat and homogeneous (curvature—which vanishes identically—and the metric are constant) can be captured or ‘understood’ in its entirety once any little piece is. The universe shares that certainty as long as it resembles $T_pM$, and hence only in $U$, outside of which it is subject to all sorts of unforeseeable variations. Integrable congruent propagation had to be rejected as allowing the certain comparison of lengths well beyond $U$, indeed at any distance, without the welcome ambiguities related to the path followed. Returning to Ryckman (p.149):

Guided by the phenomenological methods of “eidetic insight” and “eidetic analysis”, the epistemologically privileged purely infinitesimal comparison relations of parallel transport of a vector, and the congruent displacement of vector magnitude, will be the foundation stones of Weyl’s reconstruction. The task of comprehending “the sense and justification” of the mathematical structures of classical field theory is accordingly to be addressed through a construction or constitution of the latter within a world geometry entirely built up from these basic geometrical relations immediately evident within a purely infinitesimal space of intuition. A wholly epistemological project, it nonetheless coincides with the explicitly metaphysical aspirations of Leibniz and Riemann to “understand the world from its behaviour in the infinitesimally small.

Removed from the context of Weyl’s theory, the two prejudices are entirely distinct. While one is markedly infinitesimal, the other—‘mathematical justice’—has nothing (necessarily) infinitesimal about it: in a spirit of equal rights one could require, for instance, both the directions and lengths of the vectors in some set to have the same kind of distribution—uniform, say, or Gaussian—around a given vector. Nothing infinitesimal about that.

An abundant insistence in the early going on the equal rights of direction and length, together with the absence, back then, of any developed, articulated expression of the telesccepticism just outlined, suggests the following account. First there was mathematical justice, which, far from being at odds with Weyl’s nascent infinitesimal agenda, supported and perhaps even suggested it. In due course Weyl’s ‘purely infinitesimal geometry’ acquired more explicit philosophical grounding—expressed in footnotes 12, 13 and rooted in Husserl’s transcendental phenomenology—which can in hindsight be viewed as justifying and motivating the surprising, apparently gratuitous early insistence on equal rights.

4. Final remarks

Out of a sense of mathematical justice, then, Weyl made congruent displacement just as path-dependent as parallel transport. But experience, objected Einstein, is unfair, showing congruent displacement to be integrable. In a letter to Weyl dated April 15th (1918) he argued that clocks running at the same rate at one point will continue to run at the same rate at
another, however they get there—whatever the demands of mathematical justice. Four days later he reformulated the objection in terms of the ‘proper frequencies’ of atoms (rather than genuine macroscopic clocks) “of the same sort”: if such frequencies depended on the path followed, and hence on the different (electromagnetic) vicissitudes of the atoms, the chemical elements they would make up if brought together would not have the clean spectral lines one sees.

But even if experience shows congruent displacement to be integrable, it would be wrong to conclude that the equal rights of direction and length led nowhere; for the structure that came out of Weyl’s ostensibly groundless sense of mathematical justice would survive in our standard gauge theories, whose accuracy is less questionable.

There are various levels of ‘experience,’ ranging from the most concrete to the most abstract: from the most obvious experimental level, having to do with the results of particular experiments, to principles, perhaps even instincts, distilled from a lifetime of experience. One such principle could be Einstein’s “I am convinced that God does not play dice,” to which, having—we may conjecture—noticed that the causal regularities behind apparent randomness eventually tend to emerge, he may ultimately have been led by experience: by his own direct experience, together with his general knowledge of science and the world. One would nonetheless hesitate to view so general and abstract a principle as being a posteriori, empirical. It is clearly not a posteriori with respect to any particular experiment; only, if at all, with respect to a very loose, general and subjective kind of ongoing experience, capable of being interpreted in very different ways.

An unexpected empirical fertility of apparently a priori and unempirical prejudice can sometimes be accounted for in terms of a derivation, however indirect, from experience: by attributing remote empirical roots to considerations which at first seem to have nothing at all to do with experience. Fair enough, the world can be experienced in very different ways, some much less obvious and straightforward than others. But here we have a prejudice which—however subtle and developed one’s faculties for interpreting experience—seems to be completely unempirical. Perhaps the empirical shortcomings of the theory are best blamed, then, on the totally unempirical nature of the prejudice from which it was derived.

Or is it so completely unempirical? As mathematical justice is at issue, the principle of sufficient reason can come to mind: if there is an imbalance, an unexpected difference, there had better be a reason for it—failing which, balance, or rather justice should prevail. Even Einstein’s dice may come to mind: If a situation of apparent balance, such as \( |\psi\rangle = (|\alpha\rangle + |\beta\rangle)^{-1/2} \), gives rise to an imbalance (as it must, if a measurement is made), such as the eigenvalue +1 of the operator \( A = |\alpha\rangle\langle\alpha| - |\beta\rangle\langle\beta| \), there had better be a reason—a circumstance unrepresented in \( |\psi\rangle \) which favours \( |\alpha\rangle \). For God does not play dice: symmetry-breaking is never entirely spontaneous. But the ‘balance’ before the disruption is not always so easily seen; what tells us in general which objects or entities are to be put on an equal footing, for imbalances to be visible? Judgment, surely; a judgment somehow founded in experience, which assesses the relevant peculiarities of the context and determines accordingly. And here Weyl’s judgment and sense of balance determine that, in the context of parallel transport, the direction and length of a vector should have the same status and treatment.
Can the success of modern gauge theories be really attributed to Weyl’s sense of mathematical justice? Is the connection between those gauge theories and the equal rights of direction and length too tenuous to be worth speaking of? The lineage is unmistakable, and can be traced through Yang and Mills (1954) and Weyl (1929), back to 1918; the scheme of compensation already present in Weyl’s theory of gravitation and electricity survives in today’s gauge theories, and is central to their success … but any attempt to answer these questions would take us too far from our subject.

Whatever the relationship between mathematical justice and experience, we have a surprising example of how directly an elaborate theory can emerge from simple a priori prejudice. The prejudice seems gratuitous in the context of discovery, and only acquires justification and grounding years later, by association with an articulated ‘telescepticism,’ which provides epistemology and motivation.

I thank Mario Alai, Ermenegildo Caccese, Nancy Cartwright, Dennis Dieks, John Earman, Rossella Lupacchini, Antonio Masiello, George Sparling for many fruitful conversations; the Center for philosophy of science, University of Pittsburgh, where as Visiting Fellow I began work on Weyl’s theory; and above all Thomas Ryckman, for inspiration and discussion.

References
Weyl, H. (1918b) “Reine Infinitesimalgeometrie” Mathematische Zeitschrift 2, 384-411
Weyl, H. (1918c) letter from Hermann Weyl to Albert Einstein dated 10 December 1918
Weyl, H. (1927) Philosophie der Mathematik und Naturwissenschaft, Oldenbourg, Munich