WEYL'S THEORY OF THE COMBINED GRAVITATIONAL-ELECTROMAGNETIC FIELD

© William O. Straub, PhD
Pasadena, California

1. Early Aftermath of Einstein's Theory of Relativity

Imagine that it is early 1918, a little more than two years since Einstein's announcement of the theory of general relativity. Karl Schwarzschild, a German physicist (and, tragically, a fatality on the front line of the Great War), has already shown that Einstein's equations predict the previously unexplained perihelion shift of Mercury exactly. In less than a year, the noted British scientist Arthur Eddington will make use of a total solar eclipse to measure the deflection of starlight by the sun, which will again confirm the predictions of general relativity. Ten years earlier, Einstein's 1905 theory of special relativity had become accepted scientific thought, and now even the lay public is excited about Einstein's strange ideas of warped space and relative time. There is no "force" of gravity, no action at a distance pulling one gravitating body to another. Planets move in ellipses around the sun because warped space tells them how to move, and the sun's gravity tells space how to warp. As far as a planet is concerned, it's moving in a straight line, doing the only thing it knows how to do. Similarly, time isn't what we think it is. The ticking rate of a wristwatch depends not only on the wearer's relative velocity to some observer, it is also affected by gravity. Maybe time travel really is possible! Einstein is on the cusp of becoming the world's first scientific superstar.

For physicists everywhere, these are especially exciting times because the mathematics of general relativity promise new discoveries in other branches of physics. There are only two forces of nature known – gravitation and electromagnetism – and now the first force has been explained. Is there any doubt that the electrodynamic equations of Maxwell will also be explained as a geometric consequence of relativity's strange new mathematics?

The renowned German mathematician Hermann Klaus Hugo Weyl is already working along these lines. He has written a paper that he believes demonstrates that electrodynamics, like gravitation, is a geometric consequence of general relativity. Before the paper is submitted for publication, Weyl shows it to his friend and colleague Einstein, who is initially ecstatic, perhaps in part because he senses yet another success for his own theory. The paper is published in a respected scientific journal, *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, along with a brief addendum by Einstein. However, while he praises Weyl's theory, Einstein expresses his belief that, while elegant and beautiful, it cannot possibly describe reality. Weyl desperately tries to repair things, but it is no good. Within a few years, the theory is dead.

And yet, in just another ten years, Weyl will resurrect his theory and be completely vindicated, albeit in a way he never dreamed earlier. Not only that, his theory will become what can easily be called one of the seminal concepts in 20th-century mathematical physics, helping to bring about a revolution in a field of physics that Einstein never allowed himself to fully accept – quantum mechanics. And many years later, Weyl's ideas will provide future physicists a powerful tool in the successful development of theories that describe two other fundamental forces in nature – the strong and weak interactions.

What exactly was Weyl's theory? Why did it die off in spite of the earnest praise of Einstein and other noted physicists of the time? How was it reborn, and what significance does it have for modern physics? In the following, we'll look at the theory and its implications, point out its problems, and leave it to the reader to decide its significance.

2. Notation

Elsewhere on this website you will find a description of the principles of general relativity, its mathematics and the notation that it presupposes. As a reminder, note that partial derivatives are denoted by a single subscripted bar while covariant derivatives are indicated by a double subscripted bar, so we have quantities like $A_{\mu|\nu} = \partial A_\mu / dx^\nu$ and $F^\alpha_{\mu||\nu} = \partial F^\alpha_\mu / dx^\nu + F^\alpha_\lambda \Gamma^\lambda_{\mu\nu} - F^\lambda_\mu \Gamma^\alpha_{\lambda\nu}$, where $\Gamma^\lambda_{\mu\nu}$ are coefficients of connection, or just *connections*. Please note that in Riemannian space, the connections reduce to the Christoffel symbols: $\Gamma^\lambda_{\mu\nu} = -\left\{ {\lambda \atop \mu\nu} \right\}$.

3. The Idea of Gauge Invariance

It was observed long ago that Maxwell's equations were invariant to a certain change in the electromagnetic 4-potential A_μ; that is, if we define a new potential given by $\hat{A}_\mu = A_\mu + f_{|\mu}$, where $f_{|\mu}$ is the gradient of some scalar field $f(x)$, then the electromagnetic field tensor, defined as $F_{\mu\nu} = A_{\mu|\nu} - A_{\nu|\mu}$, is unchanged. Thus, there is an arbitrary aspect to the potential that can often be exploited to simplify problems in electrodynamics.

In 1918, Hermann Weyl attempted to formulate a new kind of gauge theory involving the metric tensor $g_{\mu\nu}$ and the tensor formalism of general relativity and differential geometry. In fact, Weyl himself coined the term "gauge transformation," likening it to a change in the scale of railroad tracks. Today, so-called "gauge theories" have nothing to do with geometrical objects like $g_{\mu\nu}$; instead, they involve *local phase changes* in quantum fields [i.e., $\hat{\psi} \longrightarrow \exp i\pi(x)\,\psi$], which are fundamental in the description of the weak and strong interactions. We can also thank Weyl for this development, which sprung from his first attempt to apply such transformations to geometry (and which we summarize in the following).

4. Vector Magnitude in Riemannian Space

You may recall from some class in elementary math or physics that the square of the magnitude or length of a vector $\vec{\zeta}$ is defined by the dot product $\vec{\zeta} \cdot \vec{\zeta} = |\zeta|^2$. In tensor notation, we write this as

$$l^2 = g_{\mu\nu}\,\zeta^\mu\,\zeta^\nu \tag{4.1}$$

where $g_{\mu\nu}$ is the symmetric metric tensor and $\zeta^\mu(x)$ is an arbitrary vector. If we take the total derivative of this expression we get

$$2l\,dl = g_{\mu\nu|\alpha}\,\zeta^\mu\,\zeta^\nu dx^\alpha + g_{\mu\nu}\,d\zeta^\mu\,\zeta^\nu + g_{\mu\nu}\,\zeta^\mu\,d\zeta^\nu$$

Using the identity $d\zeta^\mu = \Gamma^\mu_{\alpha\beta}\,\zeta^\alpha\,dx^\beta$, and adjusting the indices somewhat, we have

$$2l\,dl = g_{\mu\nu|\alpha}\,\zeta^\mu\,\zeta^\nu + g_{\mu\lambda}\,\Gamma^\lambda_{\nu\alpha}\,\zeta^\mu\,\zeta^\nu\,dx^\alpha + g_{\lambda\nu}\Gamma^\lambda_{\mu\alpha}\,\zeta^\mu\,\zeta^\nu\,dx^\alpha = g_{\mu\nu||\alpha}\,\zeta^\mu\,\zeta^\nu\,dx^\alpha \tag{4.2}$$

where $g_{\mu\nu||\alpha}$ is the covariant derivative of the metric tensor. In a Riemannian space this quantity vanishes identically, so the change in vector magnitude dl is zero. From this expression we see immediately that the quantity $g_{\mu\nu||\alpha}$ alone determines whether vector magnitude is constant or variable.

Weyl wondered if Riemannian space might be altered in some way that would allow for a non-zero dl. While thinking this over, he discovered that if the metric tensor $g_{\mu\nu}$ was "re-gauged" to $\lambda g_{\mu\nu}$ then vector magnitude would no longer be a constant. To see this, assume that the metric tensor undergoes the infinitesimal change of scale $g_{\mu\nu} \longrightarrow (1+\varepsilon\pi)g_{\mu\nu}$, where $\pi(x)$ is some function of the coordinates; that is, we define a new metric in Riemannian space in which $\hat{g}_{\mu\nu} = (1+\varepsilon\pi)g_{\mu\nu}$. The length of a vector in this new gauge is then

$$\hat{l}^2 = \hat{g}_{\mu\nu}\,\zeta^\mu\,\zeta^\nu = (1+\varepsilon\pi)\,l^2$$

Differentiation now gives

$$\begin{aligned} 2\,\hat{l}\,d\hat{l} &= \varepsilon\pi_{|\alpha}\,g_{\mu\nu}\,\zeta^\mu\,\zeta^\nu dx^\alpha + (1+\varepsilon\pi)g_{\mu\nu||\alpha}\,\zeta^\mu\zeta^\nu\,dx^\alpha \\ &= \varepsilon\pi_{|\alpha}\,g_{\mu\nu}\,\zeta^\mu\,\zeta^\nu dx^\alpha \end{aligned}$$

(Remember we're still in Riemannian space, which is why $g_{\mu\nu||\alpha}$ vanishes here.) Weyl now multiplied both sides of this expression by $\hat{l}$, giving

$$2\,(1+\varepsilon\pi)g_{\mu\nu}\,\zeta^\mu\,\zeta^\nu\,d\hat{l} = \varepsilon\pi_{|\alpha}\,g_{\mu\nu}\,\zeta^\mu\,\zeta^\nu dx^\alpha\,\hat{l}$$

Because $g_{\mu\nu}\,\zeta^\mu\,\zeta^\nu$ is arbitrary, we can divide this term out, giving (to first order in ε)

$$\begin{aligned} 2\,(1+\varepsilon\pi)\,d\hat{l} &= \varepsilon\pi_{|\alpha}\,dx^\alpha\,\hat{l}\,, \quad \text{or} \\ d\hat{l} &= \frac{1}{2}\varepsilon\pi_{|\alpha}\,dx^\alpha\hat{l} \end{aligned} \tag{4.3}$$

Weyl saw that a regauging of the metric in Riemannian space resulted in a non-zero $d\hat{l}$. This was no big deal in itself, but Weyl thought that an adjustment in Riemannian geometry might allow vector magnitude to vary from point to point without any regauging of the metric tensor. At the same time, Weyl knew that the electromagnetic 4-potential A_μ of electrodynamics also allowed a kind of regauging (also called a gauge transformation) that had no effect on Maxwell's equations (Maxwell's equations are therefore *gauge invariant*). The apparent similarity between a regauging of the metric and a gauge transformation of the 4-potential seemed to Weyl to represent a potential means of unifying general relativity with electromagnetism.

Another consideration that Weyl made involves the integration of (4.3). Assume that the vector is transported around a closed loop; integration then gives

$$\begin{aligned} \hat{l}^2 &= \hat{l}_0^2 \exp[\varepsilon \oint \pi_{|\alpha}\, dx^\alpha] \\ &= \hat{l}_0^2 \exp[\varepsilon \oint_S \text{curl}\, \pi\, dS\,] \end{aligned}$$

where we have used Stoke's theorem. However, the curl of a gradient field is zero, so the length of a "regauged" vector transported around a closed loop in Riemannian space is unchanged. Weyl believed that a suitable generalization of Riemannian geometry might allow the gradient $\pi_{|\alpha}$ to be replaced by a vector field, leading to a much more interesting situation.

5. Vector Magnitude in a Weyl Space

To start, Weyl looked at the definition $d\zeta^\mu = \Gamma^\mu_{\alpha\beta}\, \zeta^\alpha\, dx^\beta$ and assumed that a change in vector magnitude must obey a similar expression. He therefore assumed that

$$dl = \phi_\alpha\, dx^\alpha\, l \tag{5.1}$$

where $\phi_\mu(x)$ is a vector quantity of unknown origin that keeps dl from vanishing. Plugging this definition for dl into (4.2), Weyl obtained

$$\begin{aligned} 2l^2 \phi_\alpha\, dx^\alpha &= g_{\mu\nu||\alpha}\, \zeta^\mu\, \zeta^\nu dx^\alpha \text{ or} \\ 2g_{\mu\nu}\, \phi_\alpha &= g_{\mu\nu||\alpha} \end{aligned} \tag{5.2}$$

Therefore, the metric covariant derivative can be expressed in terms of the metric tensor and the field ϕ_α. Since the metric covariant derivative is no longer zero, Weyl had the non-Riemannian space he needed. He then considered the identity

$$g_{\mu\nu||\alpha} = g_{\mu\nu|\alpha} + g_{\mu\lambda}\, \Gamma^\lambda_{\nu\alpha} + g_{\lambda\nu} \Gamma^\lambda_{\mu\alpha}$$

and then wrote down the three expressions that result from a cyclic permutation of the indices μ, ν and α in $2g_{\mu\nu}\, \phi_\alpha$ and $g_{\mu\nu||\alpha}$:

$$\begin{aligned} 2g_{\mu\nu}\, \phi_\alpha &= g_{\mu\nu|\alpha} + g_{\mu\lambda}\, \Gamma^\lambda_{\nu\alpha} + g_{\lambda\nu} \Gamma^\lambda_{\mu\alpha} \\ 2g_{\alpha\mu}\, \phi_\nu &= g_{\alpha\mu|\nu} + g_{\alpha\lambda}\, \Gamma^\lambda_{\mu\nu} + g_{\lambda\mu} \Gamma^\lambda_{\nu\alpha} \\ 2g_{\nu\alpha}\, \phi_\mu &= g_{\nu\alpha|\mu} + g_{\nu\lambda}\, \Gamma^\lambda_{\mu\alpha} + g_{\lambda\alpha} \Gamma^\lambda_{\mu\nu} \end{aligned}$$

He then subtracted the first equation from the sum of the second and third, obtaining

$$g_{\lambda\alpha} \Gamma^\lambda_{\mu\nu} = -[\mu\nu, \alpha] + g_{\nu\alpha}\, \phi_\mu + g_{\alpha\mu}\, \phi_\nu - g_{\mu\nu}\, \phi_\alpha$$

where the term in brackets is the Christoffel symbol of the first kind. Contraction with $g^{\alpha\beta}$ then left him with

$$\Gamma^\alpha_{\mu\nu} = -\begin{Bmatrix} \alpha \\ \mu\nu \end{Bmatrix} + \delta^\alpha_\mu\, \phi_\nu + \delta^\alpha_\nu\, \phi_\mu - g_{\mu\nu}\, g^{\alpha\beta}\, \phi_\beta \tag{5.3}$$

where the term in curly brackets is the Christoffel symbol of the second kind:

$$\begin{Bmatrix} \alpha \\ \mu\nu \end{Bmatrix} = \frac{1}{2} g^{\alpha\beta} [g_{\mu\beta|\nu} + g_{\beta\nu|\mu} - g_{\mu\nu|\beta}]$$

Weyl therefore succeeded in deriving a non-Riemannian connection term that conveniently reduced to its Riemannian counterpart when $\phi_\mu = 0$.

Next, Weyl investigated what would happen to the connection $\Gamma^\alpha_{\mu\nu}$ when the metric underwent a gauge transformation. For simplicity, he considered the infinitesimal transformation $\hat{g}_{\mu\nu} = \exp(\varepsilon\pi)\, g_{\mu\nu}$, where ε is some small constant and $\pi(x)$ is the gauge parameter. Thus, the metric tensor transforms like $\hat{g}_{\mu\nu} = (1+\varepsilon\pi)\, g_{\mu\nu}$. To see how ϕ_μ transforms, he recalled (4.3) for Riemannian space; he therefore determined that, for $\lambda = \exp(\varepsilon\pi)$, $\hat{\phi}_\mu = \phi_\mu + \frac{1}{2}\varepsilon\pi_{|\mu}$, or $\delta\phi_\mu = \frac{1}{2}\varepsilon\pi_{|\mu}$. It was now a simple matter for Weyl to show that

$$\hat{\Gamma}^\alpha_{\mu\nu} = \Gamma^\alpha_{\mu\nu}$$

Therefore, the Weyl connection $\Gamma^\lambda_{\mu\nu}$ is invariant with respect to arbitrary changes in the gauge. At this point, Weyl must have felt that he was onto something. He still needed to investigate the consequences of this new space (which we now call a *Weyl space*) with respect to general relativity and Maxwell's equations, but this was a great start.

6. Weyl's Variational Principle

When Einstein initially developed his general theory of relativity, he was apparently unaware of what Pais refers to as the "royal road" to the field equations. Recall that in free space Einstein's field equations are

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0 \tag{6.1}$$

and not $R_{\mu\nu} = 0$, as he first surmised (here, $R_{\mu\nu}$ is the Ricci tensor and R is the Ricci scalar, defined as $R = g^{\mu\nu}R_{\mu\nu}$). Had Einstein known about the variational principle from the beginning, he would have reached the above equation immediately. The principle I am referring to involves the fact that the quantity I defined by

$$I = \int \sqrt{-g}\, R\, d^4x$$

has its minimum value if and only if (6.1) holds. Recall the expression for $R_{\mu\nu}$ in a Riemannian manifold:

$$R_{\mu\nu} = \begin{Bmatrix}\alpha\\ \mu\alpha\end{Bmatrix}_{|\nu} - \begin{Bmatrix}\alpha\\ \mu\nu\end{Bmatrix}_{|\alpha} + \begin{Bmatrix}\alpha\\ \mu\beta\end{Bmatrix}\begin{Bmatrix}\beta\\ \mu\alpha\end{Bmatrix} - \begin{Bmatrix}\alpha\\ \mu\nu\end{Bmatrix}\begin{Bmatrix}\beta\\ \alpha\beta\end{Bmatrix}$$

If we vary I with respect to some arbitrary change in $g^{\mu\nu}$, we easily find that

$$\delta I = \int \sqrt{-g}\, \left[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right] \delta g^{\mu\nu}\, d^4x \tag{6.2}$$

Setting this equal to zero then gives Einstein's free-space field equations, as asserted (if you're curious as to how this calculation is performed, see any introductory text on general relativity). Similarly, there exists a variational principle for Maxwell's equations in charge-free space, which starts with the quantity

$$I = \int \sqrt{-g}\, F_{\mu\nu}F^{\mu\nu}\, d^4x$$

where $F_{\mu\nu} = A_{\mu|\nu} - A_{\nu|\mu}$ is the antisymmetric electromagnetic field tensor, which in Cartesian matrix form looks like

$$F_{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & -B_x & 0 \end{bmatrix}$$

In this case, we have two quantities we can vary under the integral. The first is $g^{\mu\nu}$(which is contained in $\sqrt{-g}$ and in the contravariant form of $F_{\mu\nu}$), and the other is A_μ. If we vary $\sqrt{-g}\, F_{\mu\nu}F^{\mu\nu}$ with respect to $g^{\mu\nu}$, we get the energy-momentum tensor

$$T_{\mu\nu} = g^{\alpha\beta}\, F_{\mu\alpha}F_{\nu\beta} - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} \tag{6.3}$$

which is traceless ($g^{\mu\nu}T_{\mu\nu} = 0$), whereas variation with respect to A_μ leads to

$$[\sqrt{-g}\, F^{\mu\nu}]_{|\nu} = 0$$

which expresses the fact that the electromagnetic source density is zero for charge-free space. The latter equation results from what we refer to as a gauge variation δA_μ, while the former result comes about from taking the metric variation $\delta g^{\mu\nu}$. Noether's theorem (which is a fantastic discovery in its own right) states that the variational principle associates symmetries (invariances with respect to $\delta g^{\mu\nu}$, δA_μ and the like) with conservation laws (I'll have to write up something someday about Emmy Noether, whose 1918 theorem showed once and for all that female mathematicians could work with the best of them).

Anyway, Weyl must have had a brainstorm at this point. He would combine the Einstein and Maxwell scalar densities into a single integral formula, do the variations with respect to $\delta g^{\mu\nu}$and δA_μ, and see if the resulting equations corresponded to reality. Hopefully, this would show that electodynamics, like gravitation, was a geometrical construct. But he ran into a problem before he could set pencil to paper. The problem has to do with what we now call gauge weights of tensor densities, which I'll explain now.

You may recall two important aspects of the integrands we employ in variational principles: one, they must be scalars, and thus not dependent on any particular coordinate system (that's the whole idea of relativity!); and two, they must be scalar *densities*, which is why the $\sqrt{-g}$ appears (the combination $\sqrt{-g}\, d^4x$ is itself invariant with respect to coordinate change). This should all be very familiar material to you, but the concept of gauge weight may not be. Gauge weight is a number that is associated with the number of times the quantity $g_{\mu\nu}$ appears in the integrand, either explicitly or implicitly. The most important example is $\sqrt{-g}$ itself (the minus sign is unimportant, another fact you should already be familiar with). Recall the definition of g, which is the determinant of the metric tensor $g_{\mu\nu}$ (or metric matrix, as you may want to think of it). In four-dimensional Cartesian space we have $g = -g_{00}\, g_{11}\, g_{22}\, g_{33}$; g therefore has a gauge weight of 4, so $\sqrt{-g}$ has a weight of exactly 2. In three dimensions, $\sqrt{-g}$ would have a gauge weight of 3/2, etc. Similarly, each occurrence of $g^{\mu\nu}$ is counted as a weight of -1. If we rescale the metric tensor with the infinitesimal gauge factor $\varepsilon\pi$, then $\varepsilon\pi$ becomes the "counter." Thus, $\delta g_{\mu\nu} = \varepsilon\pi$, while $\delta g^{\mu\nu} = -\varepsilon\pi$. Some quantities have no gauge weight; for example, ϕ_μ has no gauge weight because $\delta\phi_\mu = 1/2\, \varepsilon\pi_{|\mu}$ is not a multiple of the gauge parameter ϕ_μ. All of this is due to Weyl.

Another example is the quantity $\sqrt{-g}\, F_{\mu\nu}F^{\mu\nu} = \sqrt{-g}\, F_{\mu\nu}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}$, which has a gauge weight of zero in four-dimensional space. In any other dimension, it is non-zero. Weyl believed that because the electromagnetic scalar density is of zero weight, then all physically-relevant tensor quantities should also have zero gauge weights. This is an appealing idea, and it also gives a special emphasis to four-dimensional space. In 1918, it is doubtful that anyone had even considered the possibility of spacetimes having more than four dimensions, so Weyl's rationale was certainly valid at the time (the five-dimensional theories of Kaluza and Klein didn't appear until the mid-1920s). Consequently, Weyl required that his variational principle be based upon a weight-zero scalar density. In Weyl's geometry, the Ricci scalar is $R = \widetilde{R} + 6g^{\mu\nu}\phi_\mu\phi_\nu - 6g^{\mu\nu}\phi_{\mu||\nu}$, where $\widetilde{R}$ is the scalar's form in ordinary Riemannian space. But wait! The scalar density that leads to Einstein's equations is $\sqrt{-g}\, R$, which is of gauge weight 1 (remember that the weight of $R_{\mu\nu}$ in Weyl's geometry is zero). Consequently, it is unsuitable because it is not gauge invariant. After some thought, Weyl decided that he would use the square of the Ricci scalar R to balance things out (there are other quantities, including $R_{\mu\nu\alpha\beta}\, R^{\mu\nu\alpha\beta}$, that he could have used, but R^2 is certainly the simplest). Also, we may reason that, since $F_{\mu\nu}F^{\mu\nu}$ is the "square" of $F_{\mu\nu}$, we should use a quadratic form for the Ricci scalar as well. So Weyl went with the quantity

$$I = \int \sqrt{-g}\, [R^2 + AF_{\mu\nu}F^{\mu\nu}\,]\, d^4x \tag{6.4}$$

in his variational principle, where A is some constant (in what follows, it turns out that $A = 6$ is a convenient choice, and it's the one I have used).

Before we launch into Weyl's integral, we'll need to know a few identities in Weyl space. Using the fact that $g_{\mu\nu||\alpha} = 2g_{\mu\nu}\phi_\alpha$, it is a simple matter to show that the contravariant covariant derivative is $g^{\mu\nu}_{\ \ ||\alpha} = -2g^{\mu\nu}\phi_\alpha$. The metric determinant also has a non-zero covariant derivative, which is $(\sqrt{-g})_{||\alpha} = 4\sqrt{-g}\,\phi_\alpha$. Furthermore, partial derivatives of vector densities are numerically the same as their covariant counterparts, so $(\sqrt{-g}\,\zeta^\mu)_{||\mu} = (\sqrt{-g}\,\zeta^\mu)_{|\mu}$, a fact that will be of use when we do integration by parts under the integral sign (actually, this second identity holds in Riemannian space as well).

I'm going to tell you right now that passing the variational operator δ for $g^{\mu\nu}$ and ϕ_μ through the Weyl integral is a major pain in the neck. The process is straightforward, but it's tedious and prone to error if you don't get things right all the way through (Mathematica has a tensor package that might make short work of this problem, but I haven't tried it). The main headache is the fact that you've got to integrate by parts many times, and the potential for making simple algebraic errors is always present. After about half an hour my calculations are several pages long but, fortunately (and very typical of tensor problems), terms start cancelling out all over the place, leaving a relatively simple answer. (Historical note: Weyl used a kind of round-about approach in which he initially set R equal to a constant, then corrected for it at the end of his calculations. It's rather tough to follow, but the results are identical to what follows below, so I suppose he knew what he was doing.) For brevity's sake (and my own, as I don't like typesetting all this stuff), I will just write down the results. We end up with

$$\delta I = \int \sqrt{-g}\,[W_{\mu\nu}\,\delta g^{\mu\nu} + S^\mu \delta\phi_\mu]\,d^4x$$

where $W_{\mu\nu}$ and S^μ are coefficients of the indicated variations. Setting each of these to zero, we find that

$$\sqrt{-g}\,S^\mu = [\sqrt{-g}\,F^{\mu\nu}]_{|\nu} = \sqrt{-g}g^{\mu\nu}[R\phi_\nu + \frac{1}{2}R_{|\nu}] \tag{6.5}$$

and

$$g^{\mu\nu}W_{\mu\nu} = 2\,g^{\mu\nu}[R\,\phi_\mu\phi_\nu + R_{|\mu}\phi_\nu + \frac{1}{2}R\,\phi_{\mu||\nu} + \frac{1}{4}R_{|\mu||\nu}] = 0 \tag{6.6}$$

(the latter calculation also generates the energy-momentum tensor $T_{\mu\nu}$, but this vanished when I contracted with $g^{\mu\nu}$).

As you are probably aware, the quantity $[\sqrt{-g}\,F^{\mu\nu}]_{|\nu}$ appears in Maxwell's equations, where it is equal to$\sqrt{-g}\,S^\mu$, the electromagnetic source density. Equation (6.5) implies that the electromagnetic source is derivable from purely geometric quantities. You are also no doubt aware that the divergence of this quantity is zero, reflecting the fact that electromagnetic charge is conserved. Does Weyl's theory provide for this? Amazingly, it does. To see this, let's take the divergence of (6.5) with respect to the remaining variable μ:

$$\begin{aligned} [\sqrt{-g}\,F^{\mu\nu}]_{|\nu|\mu} &= \{\sqrt{-g}g^{\mu\nu}[R\phi_\nu + \frac{1}{2}R_{|\nu}]\}_{||\mu} \\ &= 2\sqrt{-g}g^{\mu\nu}[R\,\phi_\mu\phi_\nu + R_{|\mu}\phi_\nu + \frac{1}{2}R\,\phi_{\mu||\nu} + \frac{1}{4}R_{|\mu||\nu}] \end{aligned}$$

The first term is zero because $F^{\mu\nu}$ is completely antisymmetric in its indices, so it gets wiped out by the symmetric double partial derivatives. The term on the right side is just (6.6), which also vanishes. Consequently, Weyl's expressions are fully consistent with Maxwell's equations. To me, this is a truly beautiful aspect of the Weyl theory.

Another way of obtaining this result is to let $\delta g^{\mu\nu}$ and $\delta\phi_\mu$ represent variations of the gauge, where $\delta g^{\mu\nu} = -\varepsilon\pi g^{\mu\nu}$ and $\delta\phi_\mu = 1/2\,\varepsilon\pi_{|\mu}$. We then have

$$\begin{aligned} \delta I &= \int \sqrt{-g}\,[-W_{\mu\nu}\,\pi g^{\mu\nu} + \frac{1}{2}S^\mu\pi_{|\mu}]\,\varepsilon\,d^4x \\ &= \int [-\sqrt{-g}\,g^{\mu\nu}W_{\mu\nu} - \frac{1}{2}(\sqrt{-g}S^\mu)_{|\mu}]\,\varepsilon\pi\,d^4x \end{aligned}$$

where we have integrated by parts over the last term in the integrand. Setting this to zero, we get the same results as before.

But there's more. In the variation of (6.4) with respect to $g^{\mu\nu}$, the full expression that obtains is

$$\begin{aligned} &\sqrt{-g}\,[-\frac{1}{2}g_{\mu\nu}R^2 + 2\,R\,R_{\mu\nu} - 8\,g_{\mu\nu}g^{\alpha\beta}R\phi_\alpha\phi_\beta + 8\,R\phi_\mu\phi_\nu + 8\,R_{|\mu}\phi_\nu - 8\,g_{\mu\nu}g^{\alpha\beta}R_{|\alpha}\phi_\beta \\ &-2g_{\mu\nu}g^{\alpha\beta}R_{|\alpha||\beta} - 4\,g_{\mu\nu}g^{\alpha\beta}R\phi_{\alpha||\beta} + 2\,R_{|\mu||\nu} + 4\,R\phi_{\mu||\nu}]\,\delta g^{\mu\nu} = 0 \end{aligned}$$

Dividing out $\delta g^{\mu\nu}$and contracting with $g^{\mu\nu}$ gives (6.6). Now let us assume that space is Riemannian after all, so that $\phi_\mu = 0$. We then get

$$-\frac{1}{2}g_{\mu\nu}R^2 + 2R\,R_{\mu\nu} - 2g_{\mu\nu}g^{\alpha\beta}R_{|\alpha||\beta} + 2R_{|\mu||\nu} = 0$$

Contraction with $g^{\mu\nu}$shows that $g^{\mu\nu}R_{|\mu||\nu} = 0$, leaving us with

$$R\,[R_{\mu\nu} - \frac{1}{4}g_{\mu\nu}R]\; + R_{|\mu||\nu} = 0 \tag{6.7}$$

Compare this (Weyl's equation with $\phi_\mu = 0$) with the Einstein free-space equation:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0$$

The similarity is there, but there there's also something else. You might know that Einstein searched for an equation that, like the energy-density equation for the electromagnetic field (6.3), is traceless. Weyl's equation is traceless; multiply (6.7) by $g^{\mu\nu}$ and see for yourself. However, Einstein's equation is divergenceless, however, whereas Weyl's is not. Nevertheless, it is instructive to note that Weyl's equation (6.7) reproduces all of the predictions that general relativity makes via Einstein's equation, and it may indicate even more (the next part is where I really stick my neck out).

If you're familiar with Schwarzschild's solution of Einstein's free-space field equations, you should have no trouble reproducing them using (6.7). First, assume that the invariant line element in spherical coordinates can be written as

$$ds^2 = e^{\nu(r)}(dx^0)^2 - e^{\lambda(r)}\,dr^2 - r^2 d\theta^2 - r^2\sin^2\theta\,d\phi^2$$

I will just give you the answer, because I'm too lazy to write out all the expressions for $R_{\mu\nu}$, etc. It is

$$\begin{aligned} e^{\nu} &= 1 - \frac{2m}{r} + \frac{R}{12}r^2 \\ e^{\lambda} &= (1 - \frac{2m}{r} + \frac{R}{12}r^2)^{-1} \\ R &= \text{constant} \end{aligned} \tag{6.8}$$

This is just the Schwarzschild solution with an additional term proportional to R (which in the Einstein case is zero). If we set $R = 0$ in Weyl's equation, all of the predictions that general relativity makes are reproduced. If we retain R, on the other hand, we get a repulsive acceleration factor due to the $r^2R/12$ term, which tends to counteract the attractive force of gravity. Many years ago Cartan showed that the most general form of Einstein's equations for free space is

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 0$$

where Λ is the cosmological constant, known to be a very small number (perhaps zero). The solution to Cartan's equations is just (6.8), where $R = -4\Lambda$. Therefore, the acceleration term in Weyl's theory must be very small. Could it have anything to do with the observed repulsive force that seems to permeate all space? (I told you I was sticking my neck out).

7. Problems with Weyl's Theory

Now for the bad part. Weyl's theory was struck down, by Einstein no less, and it never really took hold, in spite of the fact that everyone thought it was an elegant and beautiful idea. What happened?

Einstein's objection to the theory can be traced to (5.1). If we integrate this, we get

$$l = l_0 \exp \int \phi_\mu\,dx^\mu \tag{7.1}$$

where l_0 is the length a vector would have in the absence of the ϕ_μ field. Einstein noted that vector length could be made proportional to the ticking of a clock (from the dx^0 part of the integral). If ϕ_μ varies from point to point in space, the clock's setting would change more and more with time. The spacing of atomic spectral lines, for example, would depend on their history and be subject to change, with unpredictable (and probably disastrous) results. Since this is not the case, Einstein declared Weyl's theory to be unphysical.

There's another problem, one that I feel is more critical. Recall Weyl's definition of the metric covariant derivative,

$$g_{\mu\nu||\alpha} = 2g_{\mu\nu}\phi_\alpha$$

In view of this definition, consider the fact that there are vectors whose magnitudes are absolute (the Compton wavelength of the electron, for example), so any theory that changes their magnitude would be nonsensical. Unfortunately, Weyl's theory cannot distinguish absolute vectors from variable-length vectors. To make this quantitative, consider the line element as an example of an absolute vector:

$$\begin{aligned} ds^2 &= g_{\mu\nu}dx^\mu dx^\nu \text{ , or} \\ 1 &= g_{\mu\nu}\lambda^\mu\lambda^\nu \end{aligned}$$

where λ^μ is the unit vector dx^μ/ds (we assume that $d\lambda^\mu/ds = \Gamma^\mu_{\alpha\beta}\lambda^\alpha\lambda^\beta$ is non-zero; after all, this represents the geodesic equations!). Taking the total derivative of this like we did earlier, we get

$$\begin{aligned} 0 &= g_{\mu\nu||\alpha}\lambda^\mu\lambda^\nu dx^\alpha \\ &= g_{\mu\nu||\alpha}\lambda^\mu\lambda^\nu\lambda^\alpha \end{aligned}$$

Therefore, either the metric covariant derivative $g_{\mu\nu||\alpha}$ is identically zero, or it satisfies the peculiar cyclic symmetry condition

$$g_{\mu\nu||\alpha} + g_{\alpha\mu||\nu} + g_{\nu\alpha||\mu} = 0 \tag{7.2}$$

Clearly, Weyl's definition $g_{\mu\nu||\alpha} = 2g_{\mu\nu}\phi_\alpha$ does not satisfy this condition. Yet (7.2) is obviously valid, indicating that the Weyl theory is not even consistent with elementary tensor calculus. So where do we go from here?

First of all, let's try to clear up Einstein's objection. It's no big deal; if the term $\exp\int\phi_\mu\, dx^\mu$ must be a constant to keep vector magnitude from changing, then let's just make it a constant. To be precise, let's assume (like we did earlier) that the vector is moved around a closed path. Then we can set

$$\oint \phi_\mu dx^\mu = 2\pi i n \tag{7.3}$$

Remembering that $\exp 2\pi i n = 1$, this makes l invariant, albeit at the expense of making ϕ_μ an imaginary quantity (can you now see why Weyl's theory might have application in quantum mechanics?). In fact, we can take this argument even further by appealing to the Bohr model of the hydrogen atom. Following Adler et al. (the original idea is due to the German physicist F. London), we consider an electron in the static spherical field of a proton which consists of the non-zero Weyl potential ϕ_0 ($\phi_i = 0$). A single planetary electron circles the proton, making a complete circuit in the time t. The electron's velocity is therefore $v = 2\pi r/t$, where r is the Bohr radius. Equating electrostatic and centrifugal forces, we have

$$\frac{mv^2}{r} = \frac{e^2}{4\pi\epsilon_0\, r}$$

where the proton's classical electric potential, $A_0 = e/4\pi\epsilon_0 r$, has been used. Using (7.1), we have

$$\oint \phi_0(r)\, dx^0 = \phi_0 \oint c\, dt = 2\pi i n$$

so that, for one electron orbit, $t = 2\pi i n/c\phi_0$. Putting all this together, we can derive the formula for the Bohr radius

$$r = \frac{\epsilon_0\, n^2 h^2}{\pi m e^2}$$

if we make the identification

$$\phi_0 = \frac{ie}{\hbar c}A_0$$

This simple argument demonstrates that there is reason to associate Weyl's ϕ_μ field with the electromagnetic 4-potential after all. Note also that Stoke's theorem allows us to write

$$\oint \phi_\mu \, dx^\mu = \int_S \text{curl } \boldsymbol{\phi} \, dS$$

so this integral can vanish only if the curl of ϕ_μ (which is $F_{\mu\nu}$) is zero, again indicating a connection with electrodynamics. However, the appearance of an imaginary potential $ie/\hbar c \, A_\mu$ makes us wonder if we're not barking up the wrong tree. After all, Planck's constant $\hbar$ shows up in quantum theory, not geometry. London felt the same way, and in 1927 suggested that Weyl's gauge theory might find a happier home in quantum mechanics. Weyl immediately recognized that this was indeed the case, and in 1929 he published a paper in *Zeitschrift für Physik* that established the connection once and for all (Weyl's 1929 paper will be the subject of a future discussion on this site).

So much for Einstein's remark. Now let's have a look at (7.2), which Weyl's manifold does not obey. Is there any way it can be fixed? The offending equation is Weyl's prescription for the metric covariant derivative, $g_{\mu\nu||\alpha} = 2g_{\mu\nu}\phi_\alpha$. Let's consider this expression for a moment.

Since $g_{\mu\nu}$ is a symmetric tensor, in n-space it has a total of $n(n+1)/2$ distinct terms. The field ϕ_α is just a vector and has just n terms. Therefore, in Weyl's geometry the metric covariant derivative has a total of $n^2(n+1)/2$ terms. Now consider the cyclic expression (7.2). The metric tensor $g_{\mu\nu}$ is still symmetric, but now the quantities $g_{\mu\mu||\mu}$ (no sum) are zero. Careful consideration of the symmetry properties in (7.2) shows that $g_{\mu\nu||\alpha}$ has only $n(n^2-1)/3$ distinct terms. Thus, the quantity $g_{\mu\nu||\alpha}$ must be a much different creature than the one Weyl came up with.

Pursuing this further, let's now do something that Weyl might have considered, but didn't. Let us contract (5.2) with $g^{\mu\nu}$; this gives us the new identity

$$\phi_\alpha = \frac{1}{2n} g^{\mu\nu} g_{\mu\nu||\alpha} \tag{7.4}$$

Notice that if this quantity had been given from the start, basic tensor analysis would forbid its "uncontraction" back to (5.2), although ϕ_α (whatever it might be) can still be used to define a non-vanishing metric covariant derivative. However, while Weyl's ϕ_α field no longer violates (7.2), we seem to have lost our simple definition for $g_{\mu\nu||\alpha}$. To get it back, notice that, in view of (7.2), we can also express (7.4) as

$$\begin{aligned} \phi_\alpha &= -\frac{1}{n} g^{\mu\nu} g_{\mu\alpha||\nu} \\ &= \frac{1}{n} g^{\mu\nu}{}_{||\nu} g_{\mu\alpha} \end{aligned}$$

Contracting this expression through with $g^{\alpha\lambda}$ leads us to

$$g^{\alpha\lambda}{}_{||\alpha} = n \, g^{\alpha\lambda} \phi_\alpha$$

This equation implies that the metric covariant derivative can be expressed in terms of the metric tensor and ϕ_μ (somewhat like Weyl's original recipe). We therefore venture to guess that the necessary definition for $g_{\mu\nu||\alpha}$ goes like

$$g_{\mu\nu||\alpha} = A \, g_{\mu\nu}\phi_\alpha + B \, g_{\alpha\mu}\phi_\nu + C \, g_{\nu\alpha}\phi_\mu$$

where A, B and C are constants. Using the fact that $g_{\mu\nu}$ is symmetric, we see immediately that $B = C$. Consideration of the cyclic symmetry in (7.2), and contracting with $g^{\mu\nu}$ shows that the unique solution is

$$g_{\mu\nu||\alpha} = \frac{2n}{n-1} \, g_{\mu\nu}\phi_\alpha - \frac{n}{n-1} \, g_{\alpha\mu}\phi_\nu - \frac{n}{n-1} \, g_{\nu\alpha}\phi_\mu \tag{7.5}$$

where n is the dimension of spacetime. Similarly, we have

$$g^{\mu\nu}{}_{||\alpha} = \frac{n}{n-1}\delta^\mu_\alpha \, g^{\beta\nu}\phi_\beta + \frac{n}{n-1}\delta^\nu_\alpha \, g^{\beta\mu} \, \phi_\beta - \frac{2n}{n-1} g^{\mu\nu}\phi_\alpha \tag{7.6}$$

Earlier we showed how Weyl derived his connection $\Gamma^{\alpha}_{\mu\nu}$ by adding and subtracting cyclic permutations for the expanded form of $g_{\mu\nu||\alpha}$. The exact same procedure for the revised $g_{\mu\nu||\alpha}$ definition shows that we must now write

$$\Gamma^{\alpha}_{\mu\nu} = -\left\{ {\alpha \atop \mu\nu} \right\} - g_{\mu\nu||\beta} g^{\alpha\beta} \tag{7.7}$$

$$= -\left\{ {\alpha \atop \mu\nu} \right\} + \frac{n}{n-1}\delta^{\alpha}_{\mu}\,\phi_{\nu} + \frac{n}{n-1}\delta^{\alpha}_{\nu}\,\phi_{\mu} - \frac{2n}{n-1} g_{\mu\nu}\, g^{\alpha\beta}\,\phi_{\beta} \tag{7.8}$$

Well, this is troublesome. Not only is it distinctly different from Weyl's definition, but a straightforward gauge transformation of this quantity using $\delta\phi_{\mu} = 1/2\,\varepsilon\pi_{|\mu}$ reveals that it is not gauge invariant for any choice of n (although the contracted form $\Gamma^{\alpha}_{\mu\alpha}$ is gauge invariant). Be this as it may, Schrödinger has shown that the most general symmetric connection that is at all possible has the form

$$\Gamma^{\alpha}_{\mu\nu} = -\left\{ {\alpha \atop \mu\nu} \right\} - B_{\mu\nu\beta}\, g^{\alpha\beta}$$

where $B_{\mu\nu\beta}$ is a special rank-three tensor that has precisely the same symmetry properties as $g_{\mu\nu||\beta}$. Consequently, the revised connection is at least partially justified. It would be interesting to know how Weyl would have responded to this.

At any rate, this is as far as we can take things without really going off the deep end. While it is still possible to find a tensor quantity T in the revised Weyl geometry such that the combination $\sqrt{-g}T^{\,2}$ is gauge invariant, it doesn't provide any more insight than the original Weyl theory. In fact, this same T quantity has almost exactly the same form as (6.5), and also has zero divergence. But of course, this is only a consequence of Noether's theorem, and nothing more (the theorem guarantees that symmetries will produce conservation laws, but doesn't guarantee that they're physically meaningful!). Perhaps there is an infinite number of such quantities, which are initially appealing but really signify nothing.

And that may be the biggest problem with Weyl's theory – in spite of the neat mathematics, it just doesn't predict anything that can be checked experimentally. Like many unified field theories, it may only be empty formalism. Be that as it may, it's still a beautiful mathematical edifice, and for that reason alone it's worth knowing about.

8. End of the Story

I think that's enough for now. I hope you have found this interesting, if only as a glimpse into what was probably the very first unified field theory. I encourage you to look at how Weyl carried the gauge concept over to quantum mechanics, where it made an enormous and lasting contribution to theoretical physics. It also set the stage for a revolution in the physics of the strong and weak interactions, where Weyl's idea evolved into what is called the principle of *local gauge invariance.* This gauge principle is now believed to lie at the heart of all of physics. Weyl reflected on his gauge theory not long before his death in 1955, when he wrote:

> The principle of general relativity had resulted above all in a theory of the gravitational field. While it was not difficult to adapt also Maxwell's equations of the electromagnetic field to this principle, it proved insufficient to reach the goal at which classical field physics is aiming: a unified field theory deriving all forces of nature from one common structure of the world and one uniquely determined law of action. [I attempted] to attain this goal by a new principle which I called gauge invariance (*Eichinvarianz*). This attempt has failed. There holds, as we now know, a principle of gauge invariance in nature; but it does not connect the electromagnetic potentials ϕ_{μ}, as I had assumed, with Einstein's gravitational potentials $g_{\mu\nu}$, but ties them to the four components of the wave field ψ by which Schrödinger and Dirac taught us to represent the electron. Of course, one could not have guessed this before the "electron field" ψ was discovered by quantum mechanics!

Praise God for creating such a wonderful universe, and for giving us thoughtful geniuses like Hermann Weyl!

References

Adler, R., Bazin, M. and Schiffer, L. *Introduction to General Relativity.* McGraw-Hill (2nd ed.),1975.

Cartan, E. *On Einstein's Gravitational Equations.* J. Math. Pures Appl., 1 (1922).

London, F. *Quantum Mechanical Interpretation of Weyl's Theory.* Zeit. f. Phys. 42 (1927).

Pais, A. *Subtle is the Lord ... The Science and the Life of Albert Einstein.* Oxford University Press (1982).

Schrödinger, E. *Space-Time Structure.* Cambridge Univ. Press (1950).

Schrödinger, E. *The Final Affine Field Laws.* Proc. Royal Irish Acad., 51 (1947-48).

Weyl, H. *Gravitation and Electricity.* Sitz. Berichte d. Preuss. Akad. d. Wissenschaften, 465 (1918).