Seminar on Gauge theory of gravity and spacetime

Willem van der Feltz

November 14, 2019

Abstract

The gauge principle seems to lie at the heart of many important and successfull theories like Yang-Mills and Quantum Electrodynamics. The 'gauge idea' that the physical system does not depend on the coordinates you use to describe it resonates very much with Einstein's idea of relativity, and we are motivated to try to derive gravity from a gauge theoretical point of view. In this talk, the gauge approach to gravity and spacetime is sketched and closely compared to the more well-known gauge approach of QED. In turns out that from the gauge approach, we arrive at the so-called Einstein-Cartan spacetime which is a generalisation of the usual Riemann spacetime of GR. The Einstein-Cartan spacetime possesses, besides curvature, also a mysterious property called torsion, linked to the intrinsic spin of matter. Without spin, the theory reduces to the Einstein field equations.

1 Introduction

This field of research tries to find out whether gauge theory is actually an underlying framework of GR: whether GR and, in general, other spacetime theories can be derived from a gauge theoretical point of view.

Friedrich Hehl, a professor here at Cologne, had been very active in this field and I have based a large part of this seminar on his paper 'Gauge theory of gravity and spacetime'. [1]

Gauge theories have in the past been very successful: the interactions within the standard model are based on the Yang-Mills gauge theory, Quantum Electrodynamics (QED) can emerge from a gauge theory and in fact gauge principles already play a central role in GR as we learned it, although we might not have noticed. All of these are motivations to study and develop gauge theories of gravity and spacetime.

Before proceeding, I think it is important to discuss and clarify the notion of gauge. The word gauge is used widely and frequently throughout physics, but its precise meaning, implications and the philosophy behind it are controversial. Some speak of gauge symmetries and some of gauge redundancies. So what are they exactly?

I would like to give an illustration of what gauge means based on a blog by Terence Tao, a mathematician and field-medal winner. [2]

2 Gauge by T. Tao

The ancient Greeks roughly separated mathematics into two fields: geometry and number theory, where geometry is the study of abstract forms. Those two fields can be combined by the concept of a coordinate system. Using a coordinate system, geometrical objects can be described by numerical ones and vice-versa. For example, one can describe a sphere as a numerical relation $x^2+y^2 = 1$ within a two dimensional Cartesian coordinate system.

Although not always necessary, it has been proven to be very useful to use coordinate systems: numerical tools such as Fourier transformation, differentiation and integration can only be used when coordinates have been chosen. A specific choice of coordinates to describe a system is called a gauge. One should always keep in mind that the geometric properties of an object should not depend on the coordinate system or gauge that was chosen to describe it. In other words, certain properties should not change when changing the coordinate system. Switching from one coordinate system to another is called a gauge transformation.

Consider a two dimensional sphere; the surface of the earth, to first approximation. Let's say we want to describe in some model the wind direction at every location on this surface. This can be done by assigning to every point on the sphere (the base space) a unit vector, or equivalently some point on a circle S^1, similar to a tanget space. If one wants now to concretely, or numerically, describe a wind-direction configuration, a gauge or coordinate system has to be fixed. In our case, we need, for every unit circle belonging to a point on the sphere, define a reference point and a preferred direction. The vector at a point can then be described numerically compared to this reference. For example, we could choose true north as the reference point on every circle, and the wind direction at every point of the sphere could be expressed with a unit imaginary number $\mathrm{e}^{i\phi}$ where ϕ is the angle away from true north. This would give us a concrete, numerical function to play with: $u(x) = \mathrm{e}^{i\phi(x)}$.

However, a priori there is not a clear or unique choice for this gauge. We might also define magnetic north as a reference. We might even change the reference points by angles $\theta(x)$ differently at every point on the base space. The true, 'physical' situation should not depend on this, however; the true wind directions do not care about how we choose to describe them. To compensate for this change of reference, our function needs to be adjusted accordingly: $u(x) \to \mathrm{e}^{-i\theta(x)}u(x)$. We recognise here already in this simple example terms that remind strongly of gauge transformations in Electrodynamics.

Depending on the physical problem or configuration, choosing a certain gauge might greatly simplify or complexify calculations you want to do with it. Choosing, in practise, a suitable gauge is far from trivial.

Often in physics, gauges are called symmetries; symmetries we impose in the hope of getting some new interesting physical behavior. However, in the light of this more mathematical discussion, they should be viewed more correctly simply as choices of coordinate systems. Since gauges and gauge transformations merely reflect different numerical descriptions of the same system, I would say that *philosophically* they should not have an effect on it. A sensible theory *has* to be gauge invariant.

3 Quantum electrodynamics

Let's now, before discussing gauges further, move on to a physical relevant situation we all know: quantum electrodynamics. [3] Although everyone already knows this example, I think it is useful to discuss it from a truly gauge-theoretical point of view. Furthermore, we can make a clear analogy with the gauge theory for spacetime later.

We start from the bare-bones action that corresponds to the Dirac equation of relativistic quantum mechanics:

$$S = \int \bar{\psi}(i\hbar c\gamma^\mu \partial_{x^\mu} - mc^2)\psi \, \mathrm{d}^4 x \, .$$

We see immediately that this action is invariant under a global phase shift $\psi \to \psi \mathrm{e}^{i\phi}$. This *global* symmetry is now, due to Noether's theorem, linked to a conserved current:

$$J^\mu(x) = \frac{e}{\hbar}\bar{\psi}(x)\gamma^\mu\psi(x) \tag{1}$$

Gauge theory now enters the stage. We apply what is called the *gauge principle*: we demand that this global symmetry also holds locally. In other words, we demand the action stays invariant under a phase shift dependent on position in spacetime: $\psi \to \psi e^{i\phi(x)}$. This is the principle of gauge theories: a global symmetry is recognized and then made *local*, without giving up the invariance of the Lagrangian. In light of the previous discussion we could look at it this way: the dirac spinors live in 4 dimensional spacetime (the base space) and consist of imaginary components in $\mathbb{C}$. It should not matter what reference orientation we choose for the complex plane in which those components live.

However, we see that the action as defined above is not invariant under such a local transformation: the derivative will also act on this local phase change:

$$\partial_\mu(e^{i\phi(x)}\psi(x)) = e^{i\phi(x)}\partial_\mu\psi(x) + i(\partial_\mu\phi(x))\psi(x)$$

What is going wrong here is that we are using the normal partial derivative. The partial derivative essentially compares in this case two spinor vectors at infinitesimal distance. What we should keep in mind here is that we use differently oriented complex coordinates to describe the spinors at each point in spacetime. We should come up with a proper way to compare two of them at different positions. The object that does this is called the *connection*.

The connection is an essential object in differential geometry and we already encountered it in GR as the Christoffel symbols. In this example I will simply introduce it as an additional term we use to construct an invariant derivative. We construct the so called *covariant derivative*:

$$D_\mu := \partial_\mu - i\frac{e}{\hbar}A_\mu, \tag{2}$$

and demand that it does not depend on the gauge:

$$D'_\mu\psi' = D'_\mu(e^{i\phi(x)}\psi) = e^{i\phi(x)}(D_\mu\psi).$$

One can show this is satisfied when our introduced connection A_μ transforms as follows:

$$A_\mu \to A'_\mu = A_\mu - \frac{\hbar}{e}\partial_\mu\phi(x),$$

and we can see that this extra term will cancel with the unwanted term from the partial derivative above. Besides being called the connection, this A_μ is also called the *gauge field* or *gauge potential* in this context. Physically, we can identify it with the electromagnetic 4-potential.

With this covariant derivative we arrive at a new, now gauge invariant Lagrangian:

$$S = \int \bar{\psi}(i\hbar c\gamma^\mu D_\mu - mc^2)\psi \, d^4x = \int \left[\bar{\psi}(i\hbar c\gamma^\mu\partial_\mu - mc^2)\psi + ce\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu\right] d^4x \,. \tag{3}$$

We see that by applying the gauge principle, the interaction between the electron field and the electromagnetic field pops up naturally.

To quote Einstein, it is unnatural to have something that acts on the system without being acted upon. In this case, we would like to add a term to the Lagrangian that describes the dynamics of the gauge field. We can construct a simple, gauge invariant term starting from the commutator of the covariant derivatives

$$[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu =: F_{\mu\nu}$$

equivalently, in terms of forms (A = diff. 1-form):

$$F := dA$$

which we recognize as the gauge invariant electromagnetic tensor. From its definition we could identify the electric and magnetic field as well as the homogenious Maxwell equations. We can now construct the scalar

$$F_{\mu\nu}F^{\mu\nu}$$

Adding this scalar to the Lagrangian we arrive at the starting point of QED:

$$S_{\mathrm{QED}} = \int \left(\bar{\psi}(i\hbar c\gamma^{\mu} D_{\mu} - mc^2)\psi + c\hbar J^{\mu} A_{\mu} - \frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu}\right) \mathrm{d}^4 x\,. \tag{4}$$

Note that this addition to the Lagrangian is not unique, we could also add other dynamical terms for A_{μ}. It turns out, however, to be the most simple addition and therefore a logical starting point. Varying this Lagrangian with respect to the gauge field A_{μ} will lead to conditions which we could identify with the inhomogenious Maxwell equations:

$$\frac{\delta S_{\mathrm{QED}}}{\delta A_{\mu}} = 0 \qquad \Rightarrow \qquad \partial_{\alpha} F^{\alpha\beta} = \mu_0 J^{\beta} \tag{5}$$

4 Principle of relativity and gauge principle

Let's now move to GR. As we will see, the gauge idea lies really at the heart of it, although with an essential twist from the above examples.

Einstein came up with general relativity when he thought about the equivalence of gravitational mass and inertial mass. Because they seem to be identical up to very high order, an observer with mass inside an enclosed capsule that experiences some normal force could not tell whether the capsule is accelerating, dragging him along, or if it is at rest but in the influence of some gravitational field.

Einstein made the radical conclusion that not only are the gravitational and inertial masses equal, the physical effects are in fact identical.

This statement is equivalent to saying that the difference between gravitational and inertial effects is only a difference in the mathematical description of the *same* physical system. It's merely a difference in the choice of coordinate system.

This reminds us immediately of the gauge principle: local changes in coordinate systems have no physical meaning. In the case of GR, we demand that local, general coordinate transformations (gauge transformations) have no effect on the physical reality.

We learned in GR-I about the covariant derivative. It was introduced in the context of differential geometry as a derivative that is invariant under general coordinate transformations. It is equal to a partial derivative plus a term we called the connection:

$$D_{\alpha} V^{\mu} := \partial_{\alpha} V^{\mu} + \Gamma^{\mu}_{\alpha\nu} V^{\nu}.$$

We can now see that this all fits completely with the gauge viewpoint on GR! D_{μ} is the correct derivative that is invariant under gauge transformations. The connection, or Christoffel symbols, is a similar object to the gauge field from QED.

There is one essential difference with the other gauge theories, however. In the context of QED, the gauge transformation was dependent on the position in spacetime but had an effect on the abstract space of phases of wave functions. In the case of GR, a gauge transformation not only depends on the point in spacetime, but also affects that very same spacetime! The once rigid space time has changed its role in the theater of physics from stage to actor.

Einstein did not derive GR from a gauge theoretical point of view. The deep connection between GR and the gauge principle motivates us, however, to try to do this. It leads us, in fact, to a more general theory of spacetime.

5 Geometry

5.1 Coframe and connection

Let's now briefly discuss geometry of spacetime in the Cartan-formalism, used by Hehl. It is not necessary to understand this deeply to get the idea later.

The first ingredient we need is a differentiable manifold where at each point we can span the local cotangent frame by four linearly independent cotangent vectors, the *coframe* $\theta^\alpha = e_i{}^\alpha dx^i$, where dx^i are the coordinate covectors. From the coframe, we could define the *dual*: the usual tangent vector space. In this discussion, the coframe is most useful, however. The cotangent frame might be anholonomic, defined by $d\theta^\alpha \neq 0$, which would translate to the fact that the $e_i{}^\alpha$ are not constants.

The second essential ingredient is the connection, $\Gamma_\alpha{}^\beta = \Gamma_{i\alpha}{}^\beta dx^i$ which allows us to compare tangent covectors at different points on the manifold. It allows us to define parallel transport and covariant differentiation.

Note that at this point, we haven't defined a metric yet. Whereas in GR, we mostly reason from the metric and also derive the connection from it, it can be seen separately as a more fundamental object. In fact, also Einstein recognized the connection as the directly relevant conceptual element for GR: it can break the rigidity of space allowing for curvature. The Riemannian metric seems in some sense to be of secondary importance.

The coframe and the connection are a good starting point for a description of spacetime. Also, we could see they correspond to the gauge potentials (equation 15). From it, we can define the curvature and torsion as follows:

$$T^\alpha := D\theta^\alpha = d\theta^\alpha + \Gamma_\beta{}^\alpha \wedge \theta^\alpha = \frac{1}{2} T_{ij}{}^\alpha dx^i \wedge dx^j, \tag{6}$$

$$R_\alpha{}^\beta := D\Gamma_\alpha{}^\beta = d\Gamma_\alpha{}^\beta + \Gamma_\alpha{}^\gamma \Gamma_\gamma{}^\beta = \frac{1}{2} R_{ij\alpha}{}^\beta dx^i \wedge dx^j \tag{7}$$

which we could now also call the gauge field strengths. We already encountered the curvature in GR I. There it was the Riemann tensor. It can also be written as the commutator of the covariant derivative

$$R^\beta{}_{\alpha\mu\nu} V_\beta := [D_\mu, D_\nu] V_\alpha$$

and it could be interpreted as the change in direction a vector gets when it is parallel-transported along a closed curve. Similarly, the torsion tells about the change of spin a vector gets.

5.2 Metric

If we want now to measure angles and time and space intervals, we need to introduce a metric tensor

$$g = g_{\alpha\beta} \theta^\alpha \otimes \theta^\beta = e_i{}^\alpha e_j{}^\beta g_{\alpha\beta} dx^i dx^j. \tag{8}$$

We can choose a coframe such that the metric becomes orthonormal.

From it we can define the nonmetricity:

$$Q_{\alpha\beta} := -Dg_{\alpha\beta} = \qquad\qquad . \tag{9}$$

The nonmetricity tells about how the metric would change around a closed loop. We require it to be zero if we want the space and angles to be integrable.

Now that we have a metric, we could compare our connection to the more specific Levi-Cevita connection. They are equal when the torsion and nonmetricity vanish.

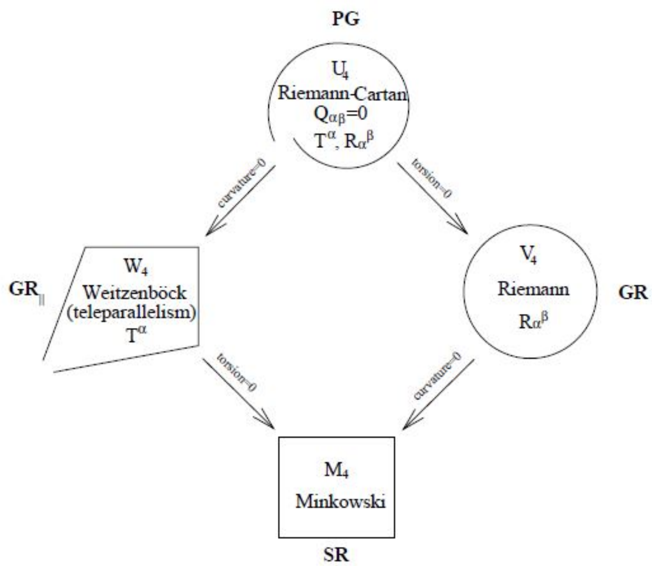

Figure 1: Different types of spaces, where Riemann-Cartan is the most general one with both nonzero curvature and nonzero torsion.

6 Gauge theory of gravity and spacetime

6.1 Dirac lagrangian and conserved currents

Let's now return to the gauge theory. We start, like in QED, with the Dirac Lagrangian in Minkowski spacetime. As is stressed by Hehl, in a gauge theory it is important to start with a conserved current connected to some global symmetry of the system. We set $c = 1$:

$$\mathfrak{L} = \frac{i}{2}\bar{\Psi}(\gamma^i\partial_i + m)\Psi. \tag{10}$$

We first consider the invariance of the Lagrangian under global (or rigid) Poincaré transformations. From the translational invariance we get, from Noether's theorem, the energy momentum tensor density:

$$\mathfrak{T}^i_\alpha = \delta^k_i\mathfrak{L} - \frac{\partial\mathfrak{L}}{\partial\partial_k\Psi}\partial_i\Psi \tag{11}$$

Similarly, we get from the Lorentz-invariance the canonical total angular momentum tensor density, consisting of intrinsic and orbital part:

$$\mathfrak{J}_{ij}{}^k = \mathfrak{S}_{ij}{}^k + x_i\mathfrak{T}_j{}^k - x_j\mathfrak{T}_i{}^k = -\mathfrak{J}_{ji}{}^k \tag{12}$$

where the intrinsic, canonical spin angular momentum density is given by

$$\mathfrak{S}_{ij}{}^k = \frac{\partial\mathfrak{L}}{\partial\partial_k\Psi}l_{ij}\Psi = -\mathfrak{S}_{ji}{}^k \tag{13}$$

Both give us conservation laws:

$$\partial_i\mathfrak{T}_\alpha{}^i = 0, \qquad \partial_k\mathfrak{J}_{ij}{}^k = 0. \tag{14}$$

6.2 Gauging the Poincaré group

We have learned in GR that the Poincaré group is a Lie group with 10 generators: 4 translations in spacetime and 6 Lorentz transformations consisting of 3 spatial rotations and 3 boosts. It is a semidirect product of the translational and the Lorentz group:

$$P(1,3) = T(4) \rtimes SO(1,3).$$

In gauge theory, a new field (gauge field) is naturally introduced for every generator of the considered Lie group. These fields are then used to construct a gauge-invariant Lagrangian.

The case for the Poincaré group in Minkovski spacetime is far from trivial, especially since the group is non-abelian (meaning it does not commute, roughly) and since the translations and Lorentz transformations are interrelated in a non-trivial way. It was first successfully analyzed by Tom Kibble in 1960. [4]

He derived that from the 4 spacetime translations, we get 4 gauge potentials: θ^α. From the Lorenz group, we get 6 potentials: $\Gamma^{\alpha\beta} = -\Gamma^{\beta\alpha}$ whose symmetry is inherited from the symmetry of the Lorentz generators. Because of the four dimensional spacetime we get in total 40 new field parameters.

$$\theta^\alpha = e_i{}^\alpha dx^i, \qquad \Gamma^{\alpha\beta} = \Gamma_i{}^{\alpha\beta} dx^i. \tag{15}$$

These gauge potentials correspond exactly to the coframe and connection, and together they give a new structure to the spacetime! A gauge invariant Lagrangian was constructed by Kibble:

$$\mathfrak{L} = \frac{ie}{2} e^i{}_\alpha \bar{\Psi} \gamma^\alpha (\partial_i + \frac{i}{4} \sigma_{\beta\gamma} \Gamma_i{}^{\beta\gamma}) \Psi + \text{herm. conj.} + m\bar{\Psi}\Psi \tag{16}$$

where $\sigma_{\beta\gamma} := \frac{i}{2}[\gamma_\beta, \gamma_\gamma]$. It can now be shown now that:

$$\frac{\delta \mathfrak{L}}{\delta e_i{}^\alpha} = \mathfrak{T}_\alpha{}^i, \qquad \frac{\delta \mathfrak{L}}{\delta \Gamma_i{}^{\alpha\beta}} = \mathfrak{S}_{\alpha\beta}{}^i; \tag{17}$$

so varying the action w.r.t. the gauge field gives us back the very same conserved currents we derived from Noether's theorem!

6.3 Sciama Kibble field equations

We now want to add to the Lagrangian a term that describes the dynamics of the gauge potentials, like in the case for QED. Again, there is no unique way of doing this. The simplest way is to construct a scalar from the curvature. We create first the Ricci tensor $\mathrm{Ric}_i{}^\alpha = e^j{}_\beta R_{ij}{}^{\alpha\beta}$ and then the scalar density $ee^i{}_\alpha \mathrm{Ric}_i{}^\alpha$. We can now add this to our Lagrangian and get the total action:

$$\mathrm{W}_{\text{tot}} = \int \mathrm{d}^4x [\frac{1}{2\kappa} e(e^i{}_\alpha \mathrm{Ric}_i{}^\alpha - 2\Lambda) + \mathfrak{L}(e_k{}^\gamma, \Psi, D\Psi)], \tag{18}$$

where also the cosmological constant Λ was added and $e := \det\left(e^i{}_\alpha\right)$. Varying this Lagrangian now with respect to our gauge fields $e_i{}^\alpha$ and $\Gamma_i{}^{\alpha\beta}$ leads finally to the following equations:

$$\mathrm{Ric}_\alpha{}^i - \frac{1}{2} e_i{}^\alpha \mathrm{Ric}_\gamma{}^\gamma + \Lambda e_i{}^\alpha = \frac{\kappa}{e} \mathfrak{T}_\alpha{}^i \tag{19}$$

$$\mathrm{Tor}_{\alpha\beta}{}^i - e^i{}_\alpha \mathrm{Tor}_{\beta\gamma}{}^\gamma + e^i{}_\beta \mathrm{Tor}_{\alpha\gamma}{}^\gamma = \frac{\kappa}{e} \mathfrak{S}_{\alpha\beta}{}^i. \tag{20}$$

These equations are known as the Einstein-Cartan field equation and they are the final result of this gauge-theoretical approach to gravity and spacetime. They are in fact a more general version of the Einstein field equations; we already see a big similarity in equation 19. When the intrinsic spin ${\mathfrak{S}_{\alpha\beta}}^{i}$ is zero, we get vanishing torsion and the equations reduce completely to the Einstein field equations.

With spin, we expect, according to the Einstein-Cartan equations, that torsion will become relevant. In 'daily' life, its effects are predicted to be very small and they haven't been measured yet. However, a critical density is known

$$\rho_{\mathrm{EC}} \approx \frac{m}{\lambda_{\mathrm{Compton}} l_{\mathrm{Plank}}^{2}},$$

at which the torsion *will* become relevant. It is expected that densities were high enough during the early stages of our universe to actually require Einstein-Cartan formalism for a correct discription!

References

[1] F. W. Hehl, "Gauge theory of gravity and spacetime," 2012.

[2] "What is a gauge?." Retrieved 11-08-2019 from `https://terrytao.wordpress.com/2008/09/27/what-is-a-gauge/`.

[3] "Gauge theory." Retrieved 11-09-2019 from `https://en.wikipedia.org/wiki/Gauge_theory#An_example:_Electrodynamics`.

[4] T. W. Kibble, "Lorentz invariance and the gravitational field," *Journal of mathematical physics*, vol. 2, no. 2, pp. 212–221, 1961.