Transforming Tensor Transformation Trauma's

Deep down into the (rabbit) hole

Roel Andringa-Boxum

January 14, 2021

1 Introduction

I have a confession to make. Even though I worked a lot with the theory of General Relativity and even have a PhD in the field of Quantum Gravity, I always found the precise meaning of general covariance and diffeomorphism invariance, the discussion between active and passive coordinate transformations, how to interpret a Lie derivative etcetera very confusing. Especially the passive versus active discussion was (and still can be!) frustrating; whenever I thought I had figured everything out from one point of view, I found a discrepancy in my understanding of the other one. It seemed like conservation of confusion. My confusion was like the components of the connection and my transformation from the active to the passive point of view like a coordinate transformation: whenever the confusion disappeared in one frame, it popped up again in the other, and vice versa. What made this confusion even more frustrating, was that a lot of my collegues simply didn't seem to care so much about these subtleties. We applied (infinitesimal) general coordinate transformations all the time, but somehow I felt I (and they?) didn't *truly* understand what was going on. I experienced it as a 'shut up and calculate' kind of attitude. From the philosophical side there are quite some papers about the meaning of 'diffeomorphism invariance', and especially Einstein's 'hole argument',[1] which kept Einstein away from his field equations for nearly two years. Even Einstein was confused! But many authors of these papers apparently really like the abstract mathematical approach to this issue, leaving aside concrete examples and coordinate representations. It was like they thought coordinate representations of tensors are for the dummie-physicists. In combination with different usages of terminology and notation, and differing opinions about something which I considered to be the very heart of General Relativity, my frustration grew. So with these highly personal notes I hope to transform away mine (and your?) confusion about some of these issues once and for all. But be prepared: exposing this stuff in all of its gory glory details can look very ugly notationally.

Some knowledge about relativity and differential geometry is assumed, especially the way manifolds and tensor fields are defined. I'll start from the very basics though, and hope to emphasize some conceptual points which a lot of textbooks simply omit. At the end you find some resources I used. If every

[1] See e.g. the papers by $John^3$, i.e. John Norton, John Stachel and John Earman, of which some are mentioned in the sources at the end.

now and then you feel like changing the title of these notes to 'how to interpret tensors for the mentally retarded', I don't mind; being excruciatingly explicit can uncover many misunderstandings. After all, the topic of these notes was Einstein's greatest stumbling block to his field equations of General Relativity, so I'm in good company being a retard. Let's get started.

2 Transformations . . . more than meets the eye

In a course on classical mechanics you often encounter coordinate transformations for the first time. A well-known example is the group of rotations. If we decompose a vector in a basis $\{e_{(i)}\}$, where the i labels a whole vector (not its components!) and runs from 1 to 3, we have[2]

$$\boldsymbol{V} = V^i e_{(i)} \,. \quad (\textit{summation convention}) \tag{1}$$

The vector *components* V^i transform in the opposite way compared to the basis *vectors* $\{e_{(i)}\}$, such that the combination (1) doesn't change. The lesson here is that the vector $\boldsymbol{V}$ doesn't care about your choice of basis, i.e. your choice of *coordinates*, while separately the components and basis vectors *do* care. In this view, the vector stays untouched, and we simply change the coordinate axes. We call this the *passive view*.

But we could also rotate the *vector* itself, keeping the coordinate axes fixed. After all, if a painting hangs tilded on the wall, a mathematician can't tell the difference between a straight painting on a tilded wall, or an oppositely tilded painting on a straight wall; the orientation of the painting with respect to the wall remains the same in both cases. So, a rotation of the axis over an angle of θ with a fixed vector, or a rotation of the vector over an angle of $-\theta$ in a fixed coordinate frame gives us the same vector components, because the vector's position with respect to the coordinate frame (or axes) is the same. We'll come back to this issue later on.

Let's think passively now. If you consider general coordinate transformations in General Relativity, you'll experience that most of these coordinate transformations are hard to interpret physically; they're mainly for mathematical convenience. But let's start simple: classical mechanics. If we switch from Cartesian to spherical coordinates (because our problem is spherically symmetric, for example), the origin stays the same and we can regard this transformation as the very same observer sitting in her origin simply relabeling the points in space. But in classical mechanics we can also perform a Galilei boost,

$$\begin{aligned} t' &= t \,, \\ x'^i &= x^i + v^i t \,, \end{aligned} \tag{2}$$

which we can interpret as a change of *observer*. This new observer, using coordinates $\{t', x'^i\}$, labels events in space differently then the original observer who uses the coordinates $\{t, x^i\}$. So a Galilei boost is also a relabeling of events, but physically we also have a change of observer in mind. Something similar goes

[2] I'll use boldface letters as coordinate-free notation for tensors.

for Lorentz boosts, and this is the whole point of relativity: how do different observers look at the same event?

Imagine we have a stick of length L at rest as measured by the observer with coordinate $\{t, x^i\}$. If we denote the coordinates of the left and right side of this stick as x_L and x_R, then the length is defined as

$$L = x_R - x_L \,. \tag{3}$$

The subtlety is that, because the stick is at rest for the observer using $\{t, x^i\}$, she can measure x_L and x_R at any time she wants. But a Galilei-boosted observer in the x-direction measures

$$x'_R - x'_L = (x_R + vt_R) - (x_L + vt_L) = x_R - x_L - v\Big(t_R - t_L\Big) = L - v\Big(t_R - t_L\Big) \,. \tag{4}$$

Here $t' = t_L$ and $t' = t_R$ are the moments where the boosted observer measures x'_L and x'_R respectively. If the boosted observer measures e.g. first the x-coordinate x'_R of the right side of the stick, and waits a few seconds $\Delta t' = t_R - t_L$ to measure the x-coordinate x'_L of the left side of the stick, the value of x'_L has increased with an amount of $v \cdot \Delta t$ because the observer is moving with respect to the stick. Whatever he's measuring, it doesn't resemble the length of the stick anymore. So 'length' has to be defined as the spatial distance between two events *at the same time*. Because 'the same time' has the same meaning for all observers, $\Delta t = \Delta t' = 0$, this length-definition gives the same value for all observers. Of course, in Special Relativity things become more subtle, with simultanity being frame dependent.

That in a mathematical abstract space we can relabel our points as we please and that geometrical objects like tensors don't care about our labeling of points is quite easy to understand: the formal definition of tensors in general don't mention any coordinate chart. But that the tensorial laws of physics, where we consider objects on space and *time*, also don't care about a relabeling of *events* is less trivial[3]... but, as we will see, more trivial than a first course on classical mechanics would make you think! In classical mechanics, we know that Newton's second law for a point particle traversing a trajectory in space with coordinate representation $x^i(t)$,

$$m\frac{d^2x^i}{dt^2} = F^i \tag{5}$$

only keeps the same form under the group of Galilei transformations, of which the Galilei boost (2) is just one. We also say that Newton's second law is *covariant* under the Galilei group. Mathematically, this is because a Galilei boost is linear in time, while Newton's second law is of second order. But if we apply a constant *acceleration*

$$t' = t \,,$$
$$x'^i = x^i + \frac{1}{2}a^i t^2 \,,$$

[3]At least, for me. I'm a dummie, remember?

to Newton's second law (5), we get

$$m\frac{d^2x'^i}{dt'^2} = F^i - ma^i \tag{6}$$

Physically, we say that the *non-inertial observer* with coordinates $\{t', x^{'i}\}$ experiences an inertial force $F^i_{inert} = -ma^i$ in her frame. We could also write

$$m\frac{d^2x'^i}{dt'^2} = F'^i = F^i - ma^i\,, \tag{7}$$

from which it is clear that, unlike for constant rotations, the force $\boldsymbol{F}$ does not transform as a vector under accelerations. Something similar can be done by applying a time-dependent rotation; in that case you'll find after some algebra (using properties of rotation matrices) that *two* extra inertial forces pop up:[4] the Coriolis force and the centrifugal force. If you followed a course on General Relativity, you know that Newton's second law will be replaced by the geodesic equation, which remains its form (or: is *covariant*) under general coordinate transformations. The Coriolis and centrifugal forces for example can then be identified as the components of the connection in the rotating frame. We'll encounter connections later on again; now we will encounter the difference between passive and active transformations in the context of the simplest of tensor fields: *scalars*.

3 Heating up the debate

Just as Newtonian mechanics without vectors is an akward business, in General Relativity we use tensors all the time. Tensors (or more precisely: tensor *fields*) are, formally, multilinear maps from products of (co)tangent spaces to the real numbers. This multilinear property, and the property that tensors are geometric objects defined without any reference to coordinates, gives us the famous transformation law for the tensor components under coordinate transformations:

$$T'^{\alpha\cdots'\beta}_{\mu\cdots\nu}(x') = \frac{\partial x^{'\alpha}}{\partial x^\lambda}\cdots\frac{\partial x^{'\beta}}{\partial x^\gamma}\frac{\partial x^\sigma}{\partial x'^\mu}\cdots\frac{\partial x^\rho}{\partial x'^\nu}T^{\lambda\cdots\gamma}_{\sigma\cdots\rho}(x)\,. \tag{8}$$

Because we are talking tensor *fields* here, a general tensor field $\boldsymbol{T}(x)$ of arbitrary rank depends on the coordinate value at which it is evaluated. We can evaluate it for different coordinate values like a normal function, e.g.

$$T_{\mu\nu}(x^\rho)\,,\quad T_{\mu\nu}(x^{'\rho})\,,\quad \ldots \tag{9}$$

with e.g. $x^\rho = (0,0,0,0)$ and $x^{'\rho} = (1,1,1,1)$. Here we use a prime simply to denote that the coordinates at which we evaluate the tensor field components $T_{\mu\nu}$ are different. If we work in one single coordinate chart, these different coordinate values x^ρ and $x^{'\rho}$ represent different *points* on the manifold. But the crucial difference between a tensor field $\boldsymbol{T}(x)$ and a normal function $f(x)$ is that if we interpret the transition from one coordinate to another as a coordinate *transformation* (!), the tensor components $T_{\mu\nu}$ also change *functionally*

[4] Two, because Newton's second law is of second order.

according to the transformation law (8). We denote that corresponding functional change with a prime on the field itself. E.g., for a scalar field Φ, the simplest of all tensors, one has per definition

$$\Phi(x^\nu) = \Phi'(x^{'\nu})\,. \tag{10}$$

So the *new* field Φ' at the *new* coordinate value $x^{'\nu}$ has the same numerical value as the *old* field Φ at the *old* coordinate value x^ν. If we are careful, we state under which group of transformations exactly this is a scalar; in e.g. textbooks on Quantum Field Theory, we would take the Poincaré transformations

$$x^{'\nu} = \Lambda^\nu{}_\mu x^\mu + \zeta^\mu\,, \tag{11}$$

with $\Lambda^\nu{}_\mu \in SO(3,1)$ being Lorentz transformations and ζ^μ spacetime translations. Let's take, as an explicit example, the Klein-Gordon equation

$$\Big(\partial_\mu\partial^\mu + m^2\Big)\Phi(x^\nu) = 0\,. \tag{12}$$

The transformed scalar field $\Phi'(x^{'\nu})$ then obeys

$$\Big(\partial'_\mu\partial^{'\mu} + m'^2\Big)\Phi'(x^{'\nu}) = 0\,, \tag{13}$$

where $m = m'$, i.e. mass is a scalar. Let's take the solution (C is a constant)

$$\Phi(x^\nu) = Ce^{i\cdot x^\nu p_\nu}\,, \tag{14}$$

which is a solution due to Einstein's energy-momentum relation $p_\nu p^\nu + m^2 = 0$. The transformed field $\Phi'(x^{'\nu})$ is completely determined by the condition (10) and the fact that it solves eqn. (13):

$$\Phi'(x^{'\nu}) = Ce^{i\cdot x'^\nu p'_\nu}\,. \tag{15}$$

The condition (10) is satisfied because the inner product $x^\nu p_\nu$ is a scalar. If you wish, you can read the prime on the field Φ' in the solution (15) as the prime on the transformed four-momentum p'_ν. Then it also becomes clear that $\Phi'(x^\nu) \neq \Phi(x^\nu)$, because

$$\Phi'(x^\nu) = Ce^{i\cdot x^\nu p'_\nu}\,, \tag{16}$$

and $x^\nu p_\nu \neq x^\nu p'_\nu$ in general. So with this example we made explicit that the functional dependence of Φ' on x^ν differs from that of Φ. If you want it even more explicit, you can consider a Lorentz transformation in some direction, calculate the p'_ν components and put them into Φ'.

Another often used example for a scalar is that of a temperature field $T(x)$. Let's say we have such a temperature field defined on a metal plate (which mathematically is a subset of the plane $\mathbb{R}^2$), which takes coordinate values $x^i \in \mathbb{R}^2$ and spits out a temperature $T(x^i) \in \mathbb{R}$:

$$T(x^i) : \mathbb{R}^2 \to \mathbb{R}\,. \tag{17}$$

Now imagine there is a heater underneath our (very large) plate, which causes a certain temperature distribution on the plate. If we shift the heater with

the vector v^i, the whole temperature distribution is shifted. So, if at a fixed coordinate value x_0^i the temperature is $T(x_0^i) = 100^oC$, then we have after the shift that the old temperature distribution T at the old coordinate value x_0^i has the same value as the new 'shifted along' temperature distribution T' at the new coordinate value $x_0^i + v^i$:

$$T'(x_0^i + v^i) = T(x_0^i) = 100^oC\,. \tag{18}$$

Notice the prime on T! In general, $T(x^i + v^i)$ has a different value then $T(x^i)$ because $T(x^i + v^i)$ denotes the *old* temperature distribution evaluated at the *shifted* coordinate $x^i + v^i$. We are not interested in that; we want to stress that, if we shift our heater, the temperature distribution is shifted along with it. That's why we interpret the temperature distribution $T(x^i)$ as a *scalar field*. Notice that we are talking about shifting the heater, i.e. the points which make up the temperature distribution on the plate. So different coordinate values mean different points on the plate! This makes sense: we stay in one and the same coordinate chart, so different points have different coordinate values per definition.

But this is not the only way we can interpret the coordinate shift

$$x'^i = x^i + v^i\,. \tag{19}$$

We could also interpret this as if we keep the heater at the same place, but merely shift the *coordinate grid* we layed upon our plate to label the points on the plate. Eqn.(18) would, in that case, mean that after this relabeling of the coordinates on the plate the *new* temperature distribution at the *new* coordinate value has the same numerical value as the old temperature distribution at the old coordinate value. But note that now, *different* coordinate values refer to the *very same* point on the plate! This makes also sense, because that's the very definition of a passive coordinate transformation: the same point on the manifold obtains a new coordinate value, i.e. is relabeled.

Because of these subtleties I talked about 'coordinate values' instead of 'points' earlier on. The reason for this carefulness is the two ways we can interpret coordinate transformations for tensors. And although these interpretations are conceptually quite different, they turn out to be the same at the calculational level. In the heating example, the first transformation was an active one: we shifted the points actively. The second interpretation was a passive one: we merely relabeled our coordinates. We also saw that there is a deep connection between these two interpretations. This is what I'll call the *passive-active duality*. To dive into that duality, we first repeat some basic differential geometry, with the assumption that you have all seen this before; see e.g. Wald, Carroll or Nakahara.

4 Differential geometry

An n-dimensional manifold $\mathscr{M}$ is per definition locally homeomorphic to flat $\mathbb{R}^n$, which means that we can use Cartesian coordinates to label points on this manifold, and we can apply all the calculus we love and hate. In General Relativity this $\mathscr{M}$ represents spacetime, and $\mathbb{R}^n$ represents Minkowski spacetime.

This means in particular that in an open neighbourhood of a point on the manifold, we can approximate this manifold with $\mathbb{R}^n$; in General Relativity this is the equivalence principle saying that locally in spacetime we can approximate a gravitational field (= spacetime curvature) by a flat spacetime; the corresponding coordinate transformation brings us to a freely falling observer. But such a coordinate system is just one choice; points on the manifold can be represented by many different coordinate systems. In formal treatments, like Nakahara chapter 5.1 or Wald chapter 2, you will see that the coordinates on a manifold are provided by a map $\psi : \mathscr{M} \to \mathbb{R}^n$, sucht that for a point $p \in \mathscr{M}$ we can write $\psi(p) = x^\mu$. Together with the submanifold on which this map is defined (often this is not the whole of the manifold $\mathscr{M}$) this map ψ is called a *chart*. If we have two such charts ψ_1 and ψ_2 of which the corresponding subsets of $\mathscr{M}$ overlap, then the composed map $\psi_2 \circ \psi_1^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ provides the transition from one coordinate chart to the other. We however shall abuse this notation. We will denote coordinate maps like ψ simply by x^μ, and if we go to another chart the transition function $\psi_2 \circ \psi_1^{-1}$ will be denoted by $x^{'\nu}(x^\mu)$.

So let's consider two different coordinate systems $\{x^\mu\}$ and $\{x^{'\nu}\}$, where $\mu, \nu = \{0, 1, \ldots, n-1\}$. Such a coordinate function is in our notation a map from the manifold to $\mathbb{R}^n$,

$$x^\mu : \mathscr{M} \to \mathbb{R}^n \tag{20}$$

and the coordinate value of a point $p \in \mathscr{M}$ will be denoted as $x^\mu(p)$. With these coordinate functions we're free to label our points on the manifold as we please, and we're also free to jump from one coordinate system to the other via the transition

$$x^{'\nu}(x^\mu) \; : \; \mathbb{R}^n \to \mathbb{R}^n \tag{21}$$

This function $x^{'\nu}(x^\mu)$ is actually a *diffeomorphism* on $\mathbb{R}^n$, meaning that every coordinate in the chart of x^μ is uniquely mapped to a coordinate in the chart of $x^{'\nu}$ and vice versa such that it is invertible.

We can also define tensor fields on this manifold, which are multilinear maps from the (co)tangent spaces of the manifold to the real line. E.g., at the point p the metric tensor $\boldsymbol{g}$ acts as a bilinear map at vectors $\boldsymbol{V}$ in the tangent space at p, $\boldsymbol{g}(\boldsymbol{V}, \boldsymbol{V})$, and returns its norm:

$$\boldsymbol{g} : T_p \otimes T_p \to \mathbb{R} \, . \tag{22}$$

If we choose a basis, we can denote $\boldsymbol{g}(\boldsymbol{V}, \boldsymbol{V})$ in component-form as

$$g_{\mu\nu} V^\mu V^\nu \in \mathbb{R} \, . \tag{23}$$

Under the transformation (21) the metric components transform as

$$g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x^{'\mu}} \frac{\partial x^\lambda}{\partial x^{'\nu}} g_{\rho\lambda}(x) \, . \tag{24}$$

This is the usual, *passive* view as most physicists are often exposed at if they follow their first course on General Relativity: we leave the points on the manifold untouched, but simply change the labeling. So, if we calculate an interval like

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu = g'_{\mu\nu}(x') dx^{'\mu} dx^{'\nu} = ds'^2 \, , \tag{25}$$

at one and the same point p on the manifold in two different coordinate charts, the scalar property of ds expresses that this geometric feature does not depend on how we choose to represent the differentials dx and the metric tensor $\boldsymbol{g}$ with coordinates.

We conclude that the passive point of view of coordinate transformations involves diffeomorphisms from $\mathbb{R}^n$ to itself. It turns out that the *active* point of view involves diffeomorphisms on the *manifold* to itself.

5 The passive-active duality

With a general map we can map points on one manifold to the other. If we choose these two manifolds to be the same, the map can be a diffeomorphism, and we can 'move points around'. Since in these discussions the coordinate chart is often not even mentioned, one then implicitly assumes a (one and the same!) coordinate chart, and as a result moving points around on the manifold changes the corresponding coordinate values. This is what we'll repeat here, although the usual discussion in textbooks is more general, by stressing that the diffemorphisms ed between two different manifolds $\mathscr{M}$ and $\mathscr{N}$. We'll restrict ourselves to one and the same manifold $\mathscr{M}$. We denote such a diffeomorphism as ϕ,

$$\phi : \mathscr{M} \to \mathscr{M} \tag{26}$$

and write for the two points $p \in \mathscr{M}$ and $q \in \mathscr{M}$ for example $\phi(p) = q$. The natural question then is how the *tangent spaces* $T_p\mathscr{M}$ and $T_q\mathscr{M}$ are related via this diffeomorphism ϕ. After all, these tangent spaces are the cosy homes of our tensor fields. It turns out there is a naturally induced map between these two tangent spaces, called the *differential map* ϕ^*:

$$\phi^* : T_p\mathscr{M} \to T_{\phi(p)}\mathscr{M} \,. \tag{27}$$

We can then consider maps $f : \mathscr{M} \to \mathbb{R}$ from the manifold to the real line, and note that also $f \circ \phi : \mathscr{M} \to \mathbb{R}$ is a mapping between $\mathscr{M}$ and $\mathbb{R}$. Whereas f brings us e.g. from a point p to a number in $\mathbb{R}$, the map $f \circ \phi$ brings us from another point $\phi(p) = q$ to another number in $\mathbb{R}$. This action is denoted as

$$f \circ \phi = \phi_* f \,. \tag{28}$$

Such a map is a nice thing to have, because vector fields are defined as differential operators on precisely such maps f from the manifold to the real line. So if we know how this diffeomorphism ϕ induces the map ϕ^* between the tangent spaces, we can write down a relation between the vectors sitting in these different tangent spaces T_p and $T_{\phi(p)}$, and from there go on to induce maps between dual vectors and higher rank tensor in these different tangent spaces. This generalization can be done because dual vectors are defined by their action on vectors, and higher rank tensors are defined as multilinear maps acting on vectors and dual vectors.

So what are these naturally induced maps ϕ_* and ϕ^*? For a vector $\boldsymbol{V} \in T_p\mathscr{M}$ we define the action of the transformed vector $(\phi^*\boldsymbol{V})$ on a map f as the action

of the original vector $\boldsymbol{V}$ on $f \circ \phi$:

$$(\phi^* \boldsymbol{V})(f) = \boldsymbol{V}(f \circ \phi), \qquad \phi^* \boldsymbol{V} \in T_{\phi(p)}\mathscr{N} \tag{29}$$

This is a very natural thing to do, which explains its name. Often the map ϕ^* is also called the *pushforward*, because if we would have stayed general, regarding the diffeomorphism as a map between two different manifolds, it would look like we 'pushed the vector from one manifold forwards to the other'. As you can guess, the map ϕ_* is then called the *pullback* for similar reasons, see e.g. Carroll. However, if we consider one and the same manifold, the pushforward can be regarded as the inverse of the pullback and vice versa:

$$\phi^* = (\phi^{-1})_* \,. \tag{30}$$

Now, up to this point we haven't mentioned the word 'coordinate'. Let's say we use coordinate functions x^μ to describe our point p. Coordinate functions are scalar functions (don't be fooled by the index μ!), so when we shift points on the manifold with a diffeomorphism ϕ the coordinates of this shifted point become

$$(\phi_* x)^\mu = (x \circ \phi)^\mu \equiv x'^\mu : \mathscr{M} \to \mathbb{R}^n \,. \tag{31}$$

In the first step I just applied the definition of ϕ_* on the (scalar!) coordinate functions x^μ, and in the second step I've renamed these $(\phi_* x)^\mu$ as x'^μ. But be aware: contrary to what this x'^μ-notation might suggest, we stay in the very same coordinate chart after the shifting of the point p! After all, if we apply the diffeomorphism ϕ to a point p, then

$$(x \circ \phi(p))^\mu = x^\mu\Big(\phi(p)\Big) = x^\mu(q) \,. \tag{32}$$

If we then apply ϕ_* to e.g. a second-rank tensor like the metric, we get

$$\Big[(\phi_* g)_{\mu\nu}\Big]|_{\phi(p)} = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\lambda}{\partial x'^\nu} g_{\rho\lambda}|_p \,. \tag{33}$$

The left hand side of eqn.(33) is evaluated at the point $\phi(p) = q$ with coordinates $x'^\alpha = (\phi_* x)^\alpha$, while the right hand side is evaluated at the point p with coordinates x^α. But this transformation has mathematically *exactly* the same form as the transformation (24), although the interpretation is different. In the active interpretation I just gave you, the diffeomorphism ϕ acts upon the metric $\boldsymbol{g}$ at the point p (with coordinates x^μ) and induces via ϕ_* a new metric at the new point $\phi(p) = q$ (with coordinates $(\phi_* x)^\mu$) on the manifold! This means that the very same transformation law for tensors can be interpreted in an active and a passive way. This is the passive-active duality.

To explore this passive-active duality further we note that a diffeomorphism ϕ on the manifold, i.e. an *active* transformation, can also *induce* a passive coordinate transformation. Wald explains this in his textbook on General Relativity (Appendix C). The picture which accompanies this claim is fig.1.

In fig.1 we shift the point p with coordinates $x^\mu(p)$ by a diffeomorphism to the new point $\phi(p)$ having coordinates $(\phi_* x)^\mu$. We can now can do the following: assign ($\equiv$) to the point p the *new* coordinate function $x'^\mu(p)$, such that

$$x^\mu(\phi(p)) \equiv x'^\mu(p) \,. \tag{34}$$

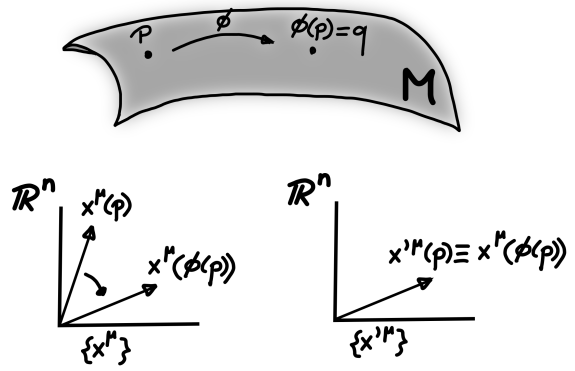

Figure 1: *How a diffeomorphism on the manifold can be used to induce a passive coordinate transformation.*

In words: we assign new coordinate values $x^{'\mu}$ to the point p in such a way that the numerical values equal the numerical values of the shifted point $\phi(p) = q$ in the old coordinate system. So beware: now the prime on $x^{'\mu}$ does mean a passive coordinate transformation again! With the very specific choice (34) the diffeomorphism ϕ *induces* the passive transformation $x^\mu(p) \rightarrow x^{'\mu}(p)$. For all tensors $\boldsymbol{T}$ we then have the following result: the components of the tensor $\phi_*\boldsymbol{T}$ at the point $\phi(p)$, in the coordinate system $x^\mu(\phi(p))$ in the *active viewpoint* have the *very same* numerical values as the components of the tensor $\boldsymbol{T}$ at the point p in the coordinate system $x^{'\mu}$ in the *passive viewpoint*.

So why mention these 'passive' coordinate transformations at all then? We can just as well stick to diffeomorphisms on the manifold, recognize that apart from shifting points actively they can also induce passive relabelings of our tensor components by imposing the condition (34), and as such avoid any long discussions and cluttered notation.[5] No wonder these smart mathematicians scarcely mention any coordinates in their formal textbooks on differential geometry and General Relativity.

6 Passive versus active: an explicit example

In this section I'll give you an explicit example of how a coordinate transformation can be interpreted both passively and actively. I stole this example and the accompanying graphs shamelessly from a user on *physics.stackexchange.com* who goes by the name *twistor59*. So *twistor59*, many thanks for your enlightening example and nice graphs.

Let's say we have the plane $\mathbb{R}^2$ with coordinates (x, y). We consider the following

[5] Or the other way around: we can use passive coordinate transformations to induce active ones.

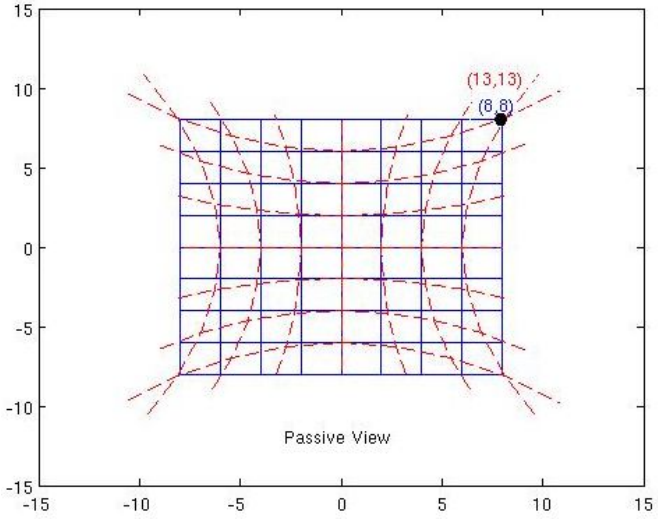

Figure 2: *The passive interpretation of the coordinate transformation eqn.*(35).

coordinate transformation to the new coordinates (X, Y):

$$X(x,y) = x\Big(1 + \frac{5}{512}y^2\Big)$$
$$Y(x,y) = y\Big(1 + \frac{5}{512}x^2\Big) \tag{35}$$

The blue coordinate lines are transformed to the red coordinate lines. In fig.2 you see the *passive* interpretation of the transformation (35): a point p first had coordinates $(x, y) = (8, 8)$ and now has coordinates $(X, Y) = (13, 13)$. The blue lines denote the coordinate lines of (x, y) and the red lines denote the coordinate lines of (X, Y).

So what happens to the geometry? Well, if we denote the coordinates (x, y) for the sake of argument as x^i, and if we denote (X, Y) as X^i, then according to (24) we have

$$g'_{km}(X) = \frac{\partial x^i}{\partial X^k}\frac{\partial x^j}{\partial X^m} g_{ij}(x)\,. \tag{36}$$

and

$$g'_{km}(X) dX^k dX^m = g_{ij}(x) dx^i dx^j\,. \tag{37}$$

In fig.3 however you see the *active* interpretation of eqn.(35): the point p is moved ('stretched out') from $(x, y) = (8, 8)$ to the new point q with coordinates $(x, y) = (13, 13)$. But after that, we *also* stretch along the coordinate lines with this transformation, giving us the coordinate system (X, Y) (the red lines in fig.3), such that the *new* point q in the *new* coordinate system (X, Y) now has the *old* coordinate values of p: $(X, Y) = (8, 8)$! In other words: the active transformation on the manifold is followed by a passive one!

The metric is also 'dragged along' from the point p to the point q, such that

$$g'_{km}(X)|_q\, dX^k dX^m = g_{ij}(x)|_p dx^i dx^j\,. \tag{38}$$

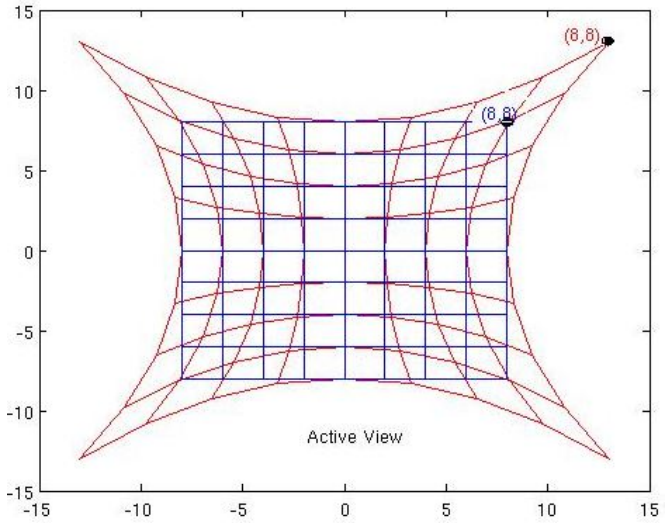

Figure 3: *The active interpretation of the coordinate transformation eqn.*(35).

Compare this with eqn.(33); the notation here is that $\phi(p) = q$, and $X^i = (\phi_* x)^i$. If the metric is the good old Euclidean metric in the Cartesian coordinates x^i, $g_{ij}(x^k) = \delta_{ij}$, then eqn.(38) states that $g'_{ij}(X^k)$ *also* equals δ_{ij}:

$$dx^2 + dy^2 = dX^2 + dY^2 \,. \tag{39}$$

Note: this equation equates two expressions at two *different* points p and $\phi(p) = q$ on the manifold! If you take a look at figure 4, this means that the angle between the two basis vector (∂_x, ∂_y) at p, which is 90^o, is the same as the angle between the basis vectors (∂_X, ∂_Y) at q. Notice however that the red vectors in fig.4 surely don't *look* orthogonal, but at q we use the new metric $\phi_* \boldsymbol{g}$ with metric components $g'_{km}(X)|_q$ to measure angles.

7 The Lie derivative

You probably know about the subtleties to introduce a derivative operator on a manifold. After all, a derivative consists of comparing two different points, and on a manifold there is not a unique way to compare two different points on a manifold. More technically: if we want to take derivatives of a tensor, we have to compare the tensor at two infinitesimally separated points on the manifold, but that means comparing two different tangent spaces. There is not a unique way to drag the tensors between those two points if there is curvature. In order to compare two tangents spaces at different points, we need to introduce a connection. But without such a connection we're still not lost: we can use curves which are induced by diffeomorphisms!

So here's the idea. Instead of a connection we will use a curve. For that we first introduce so-called one-parameter families of diffeomorphisms ϕ_t. These

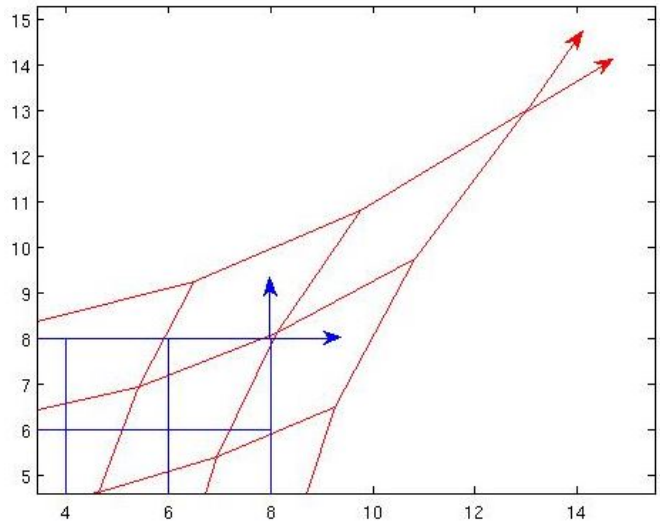

Figure 4: *The angle between the basis vectors at the two different points.*

diffeomorphisms are smooth (i.e. infinitely differentiable) maps,

$$\phi_t : \mathbb{R} \times \mathscr{M} \to \mathscr{M} \,, \tag{40}$$

such that for every value of t the map ϕ_t is a diffeomorphism. This map satisfies certain conditions, like $\phi_t \circ \phi_s = \phi_{s+t}$ and ϕ_0 being the identity operator. An example would be the following families of diffeomorphisms on the plane $\mathscr{M} = \mathbb{R}^2$ defined by

$$\phi_t(x, y) = (x + t, y) \,, \tag{41}$$

i.e. a 'flow' along the x-direction. As such these ϕ_t can be regarded as curves on the manifold, and we can also think about these curves as arising from vector fields $\boldsymbol{\xi}$. After all, a vector field is defined by its action on precisely such curves! If we denote a coordinate representation of such a curve as $x^\mu(t)$, this means

$$\xi^\mu = \frac{dx^\mu}{dt} \,. \tag{42}$$

For a given vector field $\boldsymbol{\xi}$ a solution could be parametrised as

$$x^\mu(t) = x^\mu(0) + t\,\xi^\mu \,. \tag{43}$$

Hence every (equivalence class of) curve(s) defines a vector and vice versa. Now we can use these curves to pushforward and pullback tensors as in figure 5.

Before I give you my definition of the Lie derivative, a word of warning: you should be aware of different definitions of the Lie derivative; see e.g. Nakahara chapter 5.2. One reason is that already a normal derivative can be written down in different ways. E.g., for a function $f(x)$ we define

$$\frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \,, \tag{44}$$

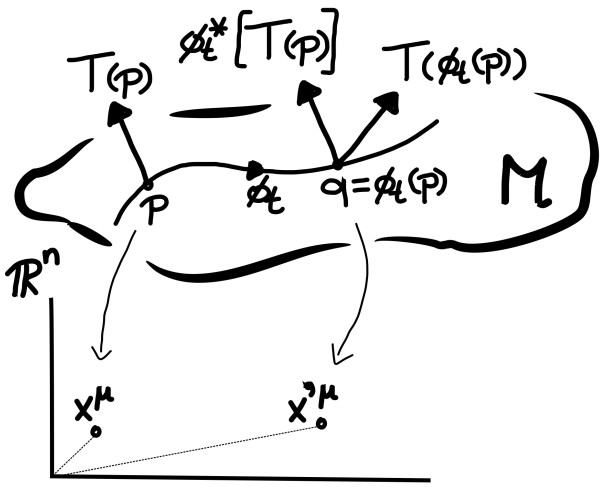

Figure 5: *The idea of the Lie derivative.*

but we can just as well replace h by $-h$ to get

$$\frac{df}{dx} = \lim_{-h\to 0} \frac{f(x-h)-f(x)}{-h} = \lim_{-h\to 0} \frac{f(x)-f(x-h)}{h} = \lim_{h\to 0} \frac{f(x)-f(x-h)}{h}\,. \tag{45}$$

I.e., now we compare the function in the points x and $x-h$. For the Lie derivative you can similarly encounter different expressions. So having said that, we will definie the Lie derivative with the help of figure 6 as follows:

$$\mathscr{L}_\xi \boldsymbol{T} = \lim_{t\to 0} \frac{\Big(\boldsymbol{T}(\phi_t(p)) - \phi_t^*\Big[\boldsymbol{T}(p)\Big]}{t}\Big)\,. \tag{46}$$

You can compare this definition with the one of e.g. Carroll. What we do mathematically, is to pushforward the tensor from the old point p towards the new point $\phi_t(p)$, and compare that result with the (old) tensor simply evaluated at the new point $\phi_t(p)$. Note that $\phi_t^*\Big[\boldsymbol{T}(p)\Big]$ is a tensor evaluated at the point $\phi_t(p)$. After all, that's what our naturally induced map ϕ_t^* does: it's a map from the tangent space at p to the tangent space at $\phi_t(p)$. If we want to apply this definition to a covariant tensor, then we have to use the map ϕ_{t*}, but being a diffeomorphism this map equals $[\phi_t^*]^{-1} = \phi_{-t}^*$; see eqn.(30).

The simplest way to calculate an explicit expression for this Lie derivative is given by the coordinate representations of the tensor field $\boldsymbol{T}$ of interest and the vector field $\boldsymbol{\xi}$. We stick to one single chart, and then we give the point p the coordinates x^μ. The point $\phi_t(p)$ then has different coordinate values, which we will call x'^μ. Doing things infinitesimally in t, we can then write the solution (43) as

$$x'^\mu = x^\mu + t\,\xi^\mu\,. \tag{47}$$

Note that we now write x'^μ for the coordinate $x^\mu(t)$ which denotes the point $\phi_t(p)$, and x^μ for the coordinate $x^\mu(0)$ of the point p. With this we can then

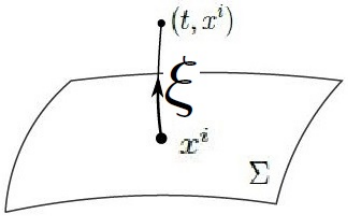

Figure 6: *The hypersurface Σ with adapted coordinates, with the vector field $\boldsymbol{\xi}$. Adapted from Harvey Reall's lecture notes on General Relativity.*

also define our Lie derivative of a tensor $\boldsymbol{T}$ with respect to $\boldsymbol{\xi}$ as

$$\mathscr{L}_\xi \boldsymbol{T} = \lim_{t\to 0}\Big(\frac{\boldsymbol{T}(x') - \boldsymbol{T'}(x')}{t}\Big)\,. \tag{48}$$

You should compare this expression with eqn.(46). For e.g. a rank 2 contravariant tensor with components $T^{\mu\nu}(x^\alpha)$ this becomes

$$\mathscr{L}_\xi T^{\mu\nu} = \lim_{t\to 0}\Big(\frac{T^{\mu\nu}(x'^\alpha) - T'^{\mu\nu}(x'^\alpha)}{t}\Big)\,. \tag{49}$$

We can then apply the diffeomorphism (47) and a Taylor expansion to get

$$\begin{aligned} T^{\mu\nu}(x') - T'^{\mu\nu}(x') &= T^{\mu\nu}(x+t\xi) - \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\lambda}T^{\rho\lambda}(x) \\ &= T^{\mu\nu}(x+t\xi) - [\delta^\mu_\rho + t\partial_\rho\xi^\mu][\delta^\nu_\lambda + t\partial_\lambda\xi^\nu]T^{\rho\lambda}(x) \\ &= T^{\mu\nu}(x) + t\xi^\rho\partial_\rho T^{\mu\nu}(x) - [\delta^\mu_\rho + t\partial_\rho\xi^\mu][\delta^\nu_\lambda + t\partial_\lambda\xi^\nu]T^{\rho\lambda}(x) \\ &= t\Big(\xi^\rho\partial_\rho T^{\mu\nu}(x) - \partial_\rho\xi^\mu T^{\rho\nu}(x) - \partial_\rho\xi^\nu T^{\mu\rho}(x)\Big) + \mathcal{O}(t^2)\,. \end{aligned} \tag{50}$$

So

$$\mathscr{L}_\xi T^{\mu\nu}(x) = \xi^\rho\partial_\rho T^{\mu\nu}(x) - \partial_\rho\xi^\mu T^{\rho\nu}(x) - \partial_\rho\xi^\nu T^{\mu\rho}(x)\,, \tag{51}$$

and similar expressions can be derived for other types of tensors. In particular, a similar calculation applied to the metric tensor gives us

$$\mathscr{L}_\xi g_{\mu\nu} = \xi^\rho\partial_\rho g_{\mu\nu} + \partial_\mu\xi^\rho g_{\rho\nu} + \partial_\nu\xi^\rho g_{\mu\rho}\,. \tag{52}$$

As you know, without torsion we can simply replace the partial derivatives by covariant ones, and given metric compatibility we can write the Lie derivative of the metric then also explicitly covariant as

$$\mathscr{L}_\xi g_{\mu\nu} = 2\nabla_{(\mu}\xi_{\nu)}\,. \tag{53}$$

Let's take a concrete example. We consider a vector field $\boldsymbol{\xi}$ and choose a spatial hypersurface Σ such that $\boldsymbol{\xi}$ is not tangent to Σ. We also assign coordinates x^i to this hypersurface. Then we can use $\boldsymbol{\xi}$ to define a diffeomorphism ϕ_t.

In the neighbourhood of Σ we can then use the coordinate (t_p, x^i_p) for a point p. This means that the vector $\boldsymbol{\xi}$ is given by

$$\boldsymbol{\xi} = \partial_t \quad \text{or} \quad \xi^\mu = (1, 0, \ldots, 0)\,. \tag{54}$$

The action of the accompanying diffeomorphism sends a point p with coordinates $x^\mu_p = (t_p, x^i_p)$ to a point $\phi_t(p) = q$ with coordinates

$$(x \circ \phi_t)^\mu \equiv x^{'\mu} = x^\mu + t\xi^\mu = (t_p + t, x^i_p)\,, \tag{55}$$

i.e. it generates a 'shift in time'. Then with the fact that the transformation matrices (Jacobians) are given by Kronecker delta's and $\xi^\mu = \delta^\mu_0$, the Lie derivative (51) becomes the partial derivative of the tensor with respect to t:

$$\mathscr{L}_\xi \boldsymbol{T} = \partial_t T_{\mu\nu}(t_p, x^i_p)\,. \tag{56}$$

Of course, we can also derive this result by using the coordinate-free definition (46); the matrix-representation of the differential map ϕ^*_t describing the shift (55) is simply the identity matrix, and hence (56) follows immediately.[6] So if the tensor components don't depend explicitly on time, the Lie derivative along the vector flow adapted to the time coordinate is zero.

This brings us to the subject of symmetries of spacetime. For that we compare the line element at two different points in spacetime with coordinates x^μ and $x^{'\mu}$, where the relation (47) holds. So we want to consider the difference

$$g_{\mu\nu}(x')dx^{'\mu}dx^{'\nu} - g_{\mu\nu}(x)dx^\mu dx^\nu\,. \tag{57}$$

But we know (see also eqn.(25)) that the line element transforms as a scalar under eqn.(47):

$$g_{\mu\nu}(x)dx^\mu dx^\nu = g'_{\mu\nu}(x')dx^{'\mu}dx^{'\nu}\,. \tag{58}$$

So eqn.(57) becomes (watch those primes carefully!)

$$\begin{aligned} g_{\mu\nu}(x')dx^{'\mu}dx^{'\nu} - g_{\mu\nu}(x)dx^\mu dx^\nu &= g_{\mu\nu}(x')dx^{'\mu}dx^{'\nu} - g'_{\mu\nu}(x')dx^{'\mu}dx^{'\nu} \\ &= \Big(g_{\mu\nu}(x') - g'_{\mu\nu}(x')\Big)dx^{'\mu}dx^{'\nu} \\ &= \Big(\mathscr{L}_\xi g_{\mu\nu}(x)\Big)dx^{'\mu}dx^{'\nu}\,. \end{aligned} \tag{59}$$

We conclude that whenever the Lie derivative of the metric (53) vanishes, $\mathscr{L}_\xi g_{\mu\nu}(x) = 0$, we have a symmetry: the geometry then doesn't change if we move into the direction of the vector field $\boldsymbol{\xi}$.

8 Down the rabbit hole

I'll now introduce you to the thing that confused me the most during my first encounters with this whole active versus passive discussion. I call it a *rabbit*

[6] A potential confusion: aren't we subtracting tensor components at two different points on the manifold here? Wasn't that ill-defined without a connection? Remember that $\phi^*_t[\boldsymbol{T}(p)]$ is actually a tensor in the (co)tangent space of $\phi(p)$, but that its value there is related to its value at p.

hole, and it concerns Lie derivatives. Let's start easy, with a scalar field $\Phi(x)$. By now, you know how to compute the Lie derivative of such a scalar field $\Phi(x)$ with respect to a vector field $\boldsymbol{\xi}$: just apply the definition (48), using eqn.(47):

$$\begin{aligned}\mathscr{L}_\xi \Phi &= \lim_{t\to 0}\Big(\frac{\Phi(x') - \Phi'(x')}{t}\Big) \\ &= \lim_{t\to 0}\Big(\frac{\Phi(x+t\xi) - \Phi(x)}{t}\Big) \\ &= \lim_{t\to 0}\Big(\frac{\Phi(x) + t\xi^\rho\partial_\rho\Phi(x) - \Phi(x)}{t}\Big) \\ &= \xi^\rho\partial_\rho\Phi(x)\,. \end{aligned} \tag{60}$$

To go to the second line, we used $\Phi'(x') = \Phi(x)$. Now let's, for the moment, just stare at the difference

$$\Phi(x') - \Phi'(x')\,. \tag{61}$$

We know with the help of figure 5 what this expression means in the active sense. My confusion, which threw me deep (*deep*!) down into the rabbit hole, arose because my first encounters with field theory all treated coordinate transformations in a passive way. So my *passive* interpretation of eqn.(61) was the following: if we use a coordinate transformation to go from $\Phi(x^\mu)$ to $\Phi'(x^{'\mu})$, the coordinates x^μ and $x^{'\mu}$ both describe the *same* point p on the manifold. However, this must mean that in the expression $\Phi(x^{'\mu})$ the coordinate $x^{'\mu}$ must describe a *different* point (say, q) on the manifold then the point that is described by $x^{'\mu}$ in $\Phi'(x^{'\mu}$! So in the passive interpretation of eqn.(61) we seem to compare tensor components at two *different* points p and q on the manifold! This seemed very strange to me, after all those warnings that to subtract tensors evaluated at two different points on the manifold we need a connection. So what's going on? Does the expression (61) simply not make sense passively?

It turns out that we *can* interpret eqn.(61) in a sensible way. In the Lie derivative, see figure 5, we made use of a curve to compare tensors. In the second term of the Lie derivative (46) we pushforward the tensor to the new point, but then we make use of the fact that its value there is related to its value at the original point; see (50). So where's the curve to make sense of eqn.(61)? Well, it's not there explicitly. But there is another trick we used, which is implicit in the passive interpretation of (61), and hence caused my confusion. Let's be very (*very*) slow here. We start with a scalar field $\Phi(x)$ in a coordinate chart $\{x^\mu\}$. The coordinate x is the value we assign to the point p. Then we perform a coordinate transformation,

$$\{x^\mu\} \to \{x^{'\mu}\} \quad , \quad \Phi(x) \to \Phi'(x')\,. \tag{62}$$

The coordinate x' in $\Phi'(x')$ still refers to the same point p. Now we take our original scalar field Φ in our original chart $\{x^\mu\}$, and evaluate the field in some point q on the manifold with coordinate y,

$$\Phi(y)\,, \tag{63}$$

where q differs from p. But we choose this new point q with coordinate y in the chart $\{x^\mu\}$ in such a way, that

$$y^\mu \equiv x^{'\mu}(p)\,. \tag{64}$$

Now my lousy notation shows up; the value y^μ is simply the value of the coordinate *function* x^μ in the point p! So eqn.(64) then tells us

$$x^\mu(q) = y^\mu \equiv x^{'\mu}(p)\,, \tag{65}$$

and hence

$$\Phi(y^\mu) = \Phi(x^{'\mu})\,. \tag{66}$$

So in $\Phi'(x')$, the coordinate $x^{'\mu}(p)$ has exactly the same numerical value as $x^\mu(q)$! But according to eqn.(34) this is *exactly* how a passive coordinate transformation can induce an active one! So there is a curve secretly in the difference (61) after all.

Confusing? Well, I warned you; it would be a deep rabbit hole. But this also explains how in some lecture notes, like the ones of e.g. Bertschinger or Andersson and Comer, the Lie derivative is defined: as a combination of an active *and* a passive transformation! We basically did this combination already in section 6. E.g, in the notes of Andersson and Comer you will see the following definition of the Lie derivative (adjusted to my notation). Consider two infinitesimally separated spacetime points p and q connected by a curve $x^\mu(t)$. Let $x^\mu(t=0) = x^\mu_0$ describe p and $x^\mu(t=\epsilon) = x^\mu_\epsilon$. We can then write infinitesimally

$$x^\mu_\epsilon = x^\mu_0 + \epsilon\xi^\mu \tag{67}$$

where

$$\xi^\mu = \frac{dx^\mu}{dt}|_{t=0}\,. \tag{68}$$

So far nothing new: we move actively from p to q following the flow of the vector field ξ^μ. But after we arrive at q, we subsequently perform the *passive* coordinate transformation

$$x^{'\mu} = x^\mu - \epsilon\xi^\mu\,. \tag{69}$$

See that minus sign? This minus sign is chosen such that when you apply this transformation on our curve,

$$x^{'\mu}_\epsilon = x^\mu_0\,. \tag{70}$$

I.e. the *new* coordinate of the *new* point q ($t=\epsilon$) has the same numerical value as the *old* coordinate of the *old* point p ($t=0$)! We also call this 'Lie dragging of the coordinate system'. The Lie derivative of a tensor field is then defined as

$$\mathscr{L}_\xi \boldsymbol{T} = \lim_{\epsilon\to 0}\Big(\frac{\boldsymbol{T}'(q) - \boldsymbol{T}(p))}{t}\Big)\,. \tag{71}$$

And this gives exactly the same expressions for the Lie derivatives. See e.g. Bertschinger's notes.

9 Digging other holes

The essence of the hole argument is that we can generate 'new solutions' of fields by using covariance, 'new' in the sense that these newly obtained solutions depend functionally different on the old coordinates. So let's take the simplest example, a toy model, to illustrate this phenomenon. Let's say we have a scalar

field $\Phi(x)$ defined on a one-dimensional manifold ('space'). We use coordinates $x \in \mathbb{R}$ to label the points on this manifold. The 'equation of motion' for this scalar field reads

$$\frac{d^2\Phi}{dx^2}(x) = 0\,. \tag{72}$$

Yes, apart from algebraic equations of motion it almost doesn't get any simpler than that. Now we perform the coordinate transformation

$$x' = 2x\,. \tag{73}$$

Per construction, our scalar field transforms as $\Phi'(x') = \Phi(x)$. And by the chaine rule,

$$\frac{d}{dx} = \frac{dx'}{dx}\frac{d}{dx'} = 2\frac{d}{dx}, \quad \rightarrow \quad \frac{d^2}{dx'^2} = 4\frac{d^2}{dx^2}\,. \tag{74}$$

So, our equation of motion (72) is covariant, i.e.

$$\frac{d^2\Phi'}{dx'^2}(x') = 4\frac{d^2\Phi}{dx^2}(x) = 0\,. \tag{75}$$

Now we look at a specific solution to eqn.(72). Employing all your differential equations solving skills, you come up with the solution

$$\Phi(x) = x - 1\,. \tag{76}$$

Then we apply our coordinate transformation (73). Similary to how we arrived at the solution eqn.(15), we conclude that then

$$\Phi'(x') = \frac{x'}{2} - 1\,. \tag{77}$$

You can check for yourself that indeed $\Phi'(x') = \Phi(x)$. If you plot eqn.(76) and (77), you find two functions in two different coordinate systems x and x', resembling the very same solution from two different points of view. But now we can do something clever: we take the solution (77), and just replace the new coordinate x' by the old coordinate x! This new solution,

$$\Phi'(x) = \frac{x}{2} - 1\,, \tag{78}$$

solves the equation of motion

$$\frac{d^2\Phi'}{dx^2}(x) = 0\,. \tag{79}$$

After all, we just relabeld Φ as Φ' in eqn.(72); who will care about the naming of the field? So in the very same coordinate system $\{x\}$ we obtain two mathematically different solutions to our equations of motion! They differ because the solutions (76) and (78), i.e. $\Phi(x)$ and $\Phi'(x)$, both depend functionally different on x. You should compare this discussion to the field eqn.(16). This field is similarly a solution of the Klein-Gordon equation (12), as you can check via $p'_\nu p'^\nu = 0$.

The upshot is that when you assign a physical meaning to x *before* you consider

the field Φ, then you have two physically different solutions to Φ. For instance, $\Phi'(x) = 0$ for $x = 2$, while $\Phi(x) = 0$ for $x = 1$. Also, $\frac{d\Phi'}{dx} = \frac{1}{2}$ while $\frac{d\Phi}{dx} = 1$. But if you *don't* assign any physical meaning to x before you have a field Φ on hand, this generation of new solutions is not giving you any new information. You could regard the generation of new solutions as mere gauge transformations, giving solutions which *look* new and which *are* new mathematically, but resemble the same physics. This last point of view is exactly the one you should take in General Relativity according to the *hole argument*.

10 Rabbit black holes

To illustrate Einstein's problems which culminated in his (in)famous hole argument, we will focus on the vacuum Einstein equations of General Relativity without cosmological constant.[7] These equations determine the time-evolution of the metric in the vacuum:

$$G_{\mu\nu}[g_{\rho\lambda}(x)] = 0\,. \tag{80}$$

The metric is a tensor under general coordinate transformations $x^\mu \to x'^\mu(x^\nu)$, which is expressed by eqn.(24). We regard this transformation in the active sense: the coordinates x^μ and x^ν in eqn.(24) are defined in the same chart and as such refer to *different* points on the manifold. Under the general coordinate transformation (24) the Einstein equations are covariant:

$$G'_{\mu\nu}[g'_{\rho\lambda}(x')] = 0\,. \tag{81}$$

Now imagine one has found a solution $g_{\mu\nu}(x)$ of (80). By covariance the transformed metric $g'_{\mu\nu}(x')$ can be constructed via the coordinate transformation (24), which solves eqn.(81). However, we can reset x' in $g'_{\mu\nu}(x')$ to its old value x, giving $g'_{\mu\nu}(x)$. This new metric also solves (81) exactly in the same way as we arrived at eqn.(79):

$$G'_{\mu\nu}[g'_{\rho\lambda}(x)] = 0\,. \tag{82}$$

The following question now arises: as $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x)$ seem to be two different metrics in the same coordinate system, just as the fields $\Phi'(x)$ and $\Phi(x)$ in the former section, what then is the precise relation between them? If $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x)$ are physically different, general covariance allows one to construct an infinite amount of physically new solutions $g'_{\mu\nu}(x)$ from $g_{\mu\nu}(x)$, but with the same initial data. For Einstein it was tempting to think that $g'_{\mu\nu}(x)$ and $g_{\mu\nu}(x)$ are physically different, because they *look* different. He also happened to fail in his quest for his field equations, so it was even more tempting for him to see this hole argument as a confirmation of his failure. So for the moment let's give in with Einstein's temptation and consider a spacetime manifold $\mathscr{M}$ with a region $H \subset \mathscr{M}$ which is non-empy: $H \neq \emptyset$. The points of $\mathscr{M}$ are interpreted as events. Now consider a general coordinate transformation, such that

- outside H one has $x^\mu = x'^\mu$,
- inside H one has $x^\mu \neq x'^\mu$,

[7] This section is based on a section from my PhD-thesis *Newton-Cartan theory revisited*.

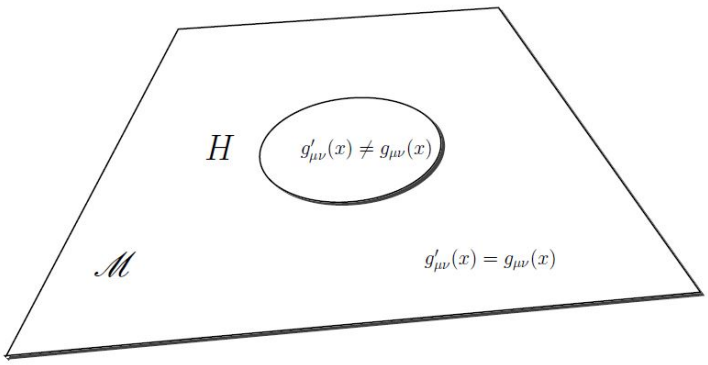

Figure 7: *The hole in the argument.*

- on the boundary of H these two transformations are smoothly connected.

As such the region H is called a "hole". Note that this construction can only be made because the coordinate transformations involved are general. With Poincaré- or Galilei transformations we wouldn't be able to do this!

The following subtlety then arised for Einstein: his equations describe the evolution of the metric, and a set of initial data should suffice to determine the metric $g_{\mu\nu}(x)$ uniquely through spacetime. Everything is fine outside the hole. But once the hole is entered, one can suddenly use covariance to obtain from the metric $g_{\mu\nu}(x)$ the mathematically different metric $g'_{\mu\nu}(x)$, as is shown in figure 7. Remember, $g'_{\mu\nu}(x)$ has a different functional dependence on the coordinate x; compare this to our discussion of the scalar field in eqn.(16). If these two metrics are also different *physically*, then covariance implies that Einstein's field equations are not deterministic. Namely, the same initial data results in different solutions inside the hole. You see the situation[8] in fig.8, where the left depicts the geometry of $g_{\mu\nu}(x)$ and the right depicts the geometry for $g'_{\mu\nu}(x)$. On the left, a light ray goes through the point P outside the hole and inside the hole crosses the point Q. But after our transformation to $g'_{\mu\nu}(x)$, the light ray doesn't cross the point Q inside the hole anymore! You can compare this situation with the red and blue vectors in fig.4: the red vectors don't *look* orthogonal anymore. But that's no surprise, because this confusion arises because we don't measure orthogonality anymore with the old metric at p, but with the new metric at the point q.

The solution to save General Relativity is clear: $g'_{\mu\nu}(x)$ and $g_{\mu\nu}(x)$ must be physically the same. One must conclude that *mathematically, points on a manifold can be distinguished without a metric, but physically they cannot.* Points (events) and their coordinates can only be physically interpreted *after* one introduces a metric, and as such a spacetime always consists of a manifold $\mathscr{M}$ equipped with a metric structure. But in the hole argument one tacitly assumes that the points, labeled by $\{x^\mu\}$ and $\{x'^\mu\}$, have a meaning *before* the metric is considered. This is deceiving and simply wrong. In figure 3 you see a similar deceit: we seem move the point with respect to a fixed background, but physically all we can do is to regard the shift of the point p with respect

[8] Inspired by the discussion of Tim Maudlin in his 'Philosophy of Physics: Space and Time'.

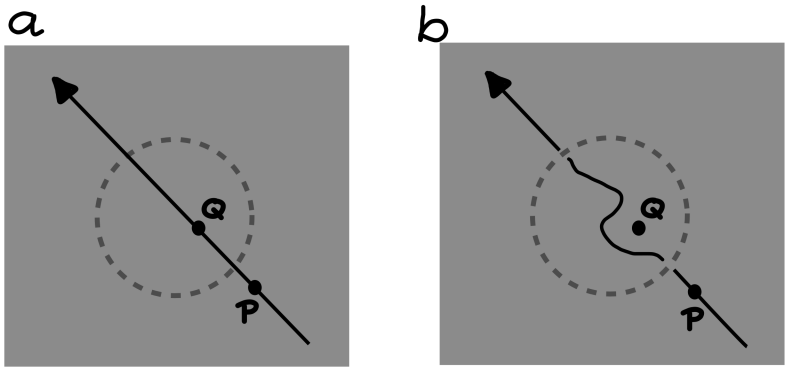

Figure 8: *A light ray inside and outside the hole, before and after the transformation. Taken from my book* Ruimte, Tijd, Materie.

to p itself; there is no 'fixed background'. This is 'relativity' at its best, I suppose, which goes under the name of *background independence*. In this sense General Relativity must be regarded as a gauge theory. If we write the general coordinate transformation as the Lie derivative eqn.(53), we get[9]

$$\delta_\xi g_{\mu\nu}(x) \equiv g'_{\mu\nu}(x) - g_{\mu\nu}(x) = 2\nabla_{(\mu}\xi_{\nu)}\,. \tag{83}$$

Under this gauge transformation the vacuum Einstein equation $G_{\mu\nu} = 0$ is invariant.

Alan Macdonald gave a very nice explicit example of the hole argument in the *American Journal of Physics*. Let's consider the Schwarzschild metric, being a solution to the vacuum Einstein equations (80). In spherical coordinates $(t, r, \Omega) = (t, r, \theta, \phi)$ the spacetime interval with $G = c = 1$ is written as

$$ds^2 = -\Big(1 - \frac{2M}{r}\Big)dt^2 + \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 + r^2 d\Omega^2\,. \tag{84}$$

Then the following transformation is chosen:

$$\begin{aligned} t &\to t' = t\,, \\ r &\to r' = f^{-1}(r)\,, \\ \Omega &\to \Omega' = \Omega\,, \end{aligned} \tag{85}$$

where the inverse is for notational convenience. The function $f^{-1}(r)$ has the following properties:

- $f^{-1}(r) = r$ outside H,
- $f^{-1}(r) \neq r$ inside H,
- on the boundary of H these two transformations are smoothly connected.

[9]Note the switching of the primes; see the remarks after eqn.(44).

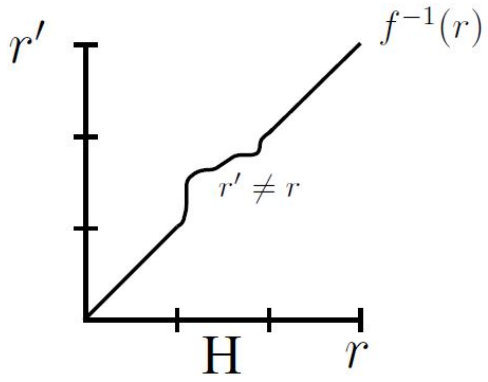

Figure 9: *The coordinate transformation* (85) *which defines our hole.*

As such our hole H is defined only by spatial coordinate transformations. This is just a choice to keep the argument as simple as possible; one could of course also involve the time coordinate in the transformations. Under the transformation (85) the spacetime interval (84) becomes

$$ds'^2 = -\Big(1 - \frac{2M}{f(r')}\Big)dt^2 + \Big(1 - \frac{2M}{f(r')}\Big)^{-1}(\frac{\partial f}{\partial r'})^2 dr'^2 + f^2(r')d\Omega^2\,, \qquad (86)$$

which by covariance equals ds^2.

Now choose $r' = r$ in eqn.(86) to get

$$d\tilde{s}^2 = -\Big(1 - \frac{2M}{f(r)}\Big)dt^2 + \Big(1 - \frac{2M}{f(r)}\Big)^{-1}(\frac{\partial f}{\partial r})^2 dr^2 + f^2(r)d\Omega^2\,. \qquad (87)$$

The spacetime interval (84) corresponds to $g_{\mu\nu}(x)$, whereas the spacetime interval (87) corresponds to $g'_{\mu\nu}(x)$. Comparison shows that they are mathematically different inside the hole,

$$ds^2 \neq d\tilde{s}^2 \qquad r \in H\,. \qquad (88)$$

If we now stick to the coordinates (t, r, θ, ϕ) and consider an event with coordinate r inside the hole H, we could naively think that for the metric with interval (84) the event is on a sphere with area $4\pi r^2$, while for (87) the event is instead on a sphere with area $4\pi f(r)^2$. Also, for (84) the horizon seems to be located at $r = 2M$, while for (87) the horizon is at $f(r) = 2M$. This surely doesn't look like the good old Schwarzschild radius! Did we, contrary to the uniqueness theorems, discover a whole new set of static and spherically symmetric black holes?

Well, no. Only *after* writing the metric (87) we can interpret the coordinate $r' = f^{-1}(r)$ and the corresponding points on the manifold. The two metrics (84) and (87) must be associated to two diffeomorphic spacetime manifolds, describing the same physics. This becomes even more clear when we write eqn.(86) as

$$ds'^2 = -\Big(1 - \frac{2M}{f}\Big)dt^2 + \Big(1 - \frac{2M}{f}\Big)^{-1} df^2 + f^2 d\Omega^2\,, \qquad (89)$$

with the dependency of f on r implicit. You should compare this interval with the one of eqn.(39): only *after* writing down the metric and the corresponding

interval we can interpret the coordinates, and hence we see that the two metrics are really equal. So, the moral of the story is:

"*Thou shalt not speculate about an event before the metric is on hand.*"

Historically, we can conclude that Einstein was troubled because he didn't recognize the metric to be a gauge field under general coordinate transformations. To paraphrase Anthony Zee in his *Nutshell* book on page 404: Einstein was confused because he thought that the metric was analogous to the electric field $\vec{E}$ and magnetic field $\vec{B}$ instead of the vector potential A_μ. These errors of thought costed him, in his own words, 'two years of excessively hard work'. In the end he solved his problem by the so-called *point coincidence argument*, which you can find in papers by e.g. Norton. Philosophers of science often write about the hole argument when it comes down to the so-called *ontology of spacetime*, i.e. the question what spacetime really 'is'. Does it exist independently from its content, energy and matter? The idea that spacetime exists independently from its content is called *spacetime substantialism*, i.e. the idea that spacetime is an independent 'substance'. The hole-argument shows that this view is problematic; after all, spacetime is given by a manifold *plus* a metric (as we saw, 'no metric, no nothing'), but this metric is a gauge field. What's observable is not this metric, but an equivalence class of spacetimes connected by diffeomorphisms.

11 General covariance and diffeomorphism invariance

As you may know, Einstein thought highly of the principle of general covariance. For years it served as his guiding principle for finding his field equations. It was the German physicist Erich Kretschmann[10] who corrected Einstein. Kretschmann claimed that general covariance on its own is a vacuous requirement for a theory. For example, the Klein-Gordon equation in Special Relativity, eqn. (12),

$$\Big(\eta^{\mu\nu}\partial_\mu\partial_\nu + m^2\Big)\Phi(x^\nu) = 0\,, \tag{90}$$

is defined on Minkowski spacetime with the Minkowski metric $\boldsymbol{\eta}$. So this metric provides us with a 'fixed background', and applying a general coordinate transformation will change the form of the metric $\boldsymbol{\eta}$. Therefore eqn.(90) is only covariant under the Poincaré transformations (11). But we can promote this group to the group of general coordinate transformations quite trivially by replacing the equations of motion (90) with

$$\begin{aligned} R_{\mu\nu\rho\sigma} &= 0\,, \\ \Big(g^{\mu\nu}\nabla_\mu\nabla_\mu + m^2\Big)\Phi &= 0\,. \end{aligned} \tag{91}$$

The vanishing Riemann tensor provides us with a general-covariant statement that spacetime is flat, e.g. $g_{\mu\nu} = \eta_{\mu\nu}$ such that $\nabla_\mu = \partial_\mu$. But of course we

[10] Yes, the guy after which the scalar quantity $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$ is named.

can choose whatever coordinate representation we want for the flat spacetime; we could also choose spherical coordinates, or go to an accelerating observer by choosing Rindler coordinates. The second equation provides us with the general covariant Klein-Gordon equation. If we choose inertial observers such that $g_{\mu\nu} = \eta_{\mu\nu}$, the second equation of eqn.(91) reduces to eqn.(90). As you can imagine, the condition $R_{\mu\nu\rho\sigma} = 0$ on spacetime is not trivially provided by an action principle; you need an extra structure, e.g. an extra field $C^{\mu\nu\rho\sigma}$ (playing the rôle of Lagrange multiplier), such that the action for the Riemann tensor becomes

$$S = \int \sqrt{|g|} d^4x \Big(C^{\mu\nu\rho\sigma} R_{\mu\nu\rho\sigma} \Big) \,. \tag{92}$$

The algebraic 'equation of motion' for this extra field $C^{\mu\nu\rho\sigma}$ then gives $R_{\mu\nu\rho\sigma} = 0$. However, the general-covariant form for our equations of motion (91) is not something to be excited about; after all, it comes at the expense of introducing a very strange equation of motion for the Riemann tensor, which must be implemented by an auxiliary field $C^{\mu\nu\rho\sigma}$ in the action. The theory becomes exciting however when we replace this 'equation of motion' $R_{\mu\nu\rho\sigma} = 0$ with the Einstein field equations (with e.g. the scalar field coupled to the Ricci scalar)!

A covariantization such as eqns.(91) is not limited to general coordinate transformations. We can also do it for other kinds of symmetries. Take e.g. a complex scalar field Φ obeying the Klein Gordon equation

$$\Big(\eta^{\mu\nu} \partial_\mu \partial_\nu + m^2 \Big) \Phi = 0 \,. \tag{93}$$

This equation of motion is invariant under a *global* $U(1)$ transformation with constant parameter Λ,

$$\Phi' = e^{-ie\Lambda} \Phi \,. \tag{94}$$

Now imagine you have a deep desire, an irresistable craving, to covariantize eqn.(93) with respect to *local* $U(1)$ transformations. This can be done by introducing a connection in the form of a vector field A_μ, and replacing partial derivatives with the covariant derivative

$$D_\mu = \partial_\mu + ieA_\mu \,. \tag{95}$$

With this covariant derivative we can covariantize the equations of motion for the scalar field (93) with respect to local $U(1)$ transformations again rather trivially:[11]

$$\begin{aligned} F_{\mu\nu} &= 2\partial_{[\mu} A_{\nu]} = 0 \,, \\ \Big(\eta^{\mu\nu} D_\mu D_\nu + m^2 \Big) \Phi &= 0, \,. \end{aligned} \tag{96}$$

These equations of motion are invariant under the local $U(1)$ transformations on Φ and A_μ in the following way:

$$A'_\mu = A_\mu + \partial_\mu \Lambda, \qquad \Phi' = e^{-ie\Lambda} \Phi \,, \tag{97}$$

with $\Lambda = \Lambda(x^\rho)$ now a function of the spacetime coordinates. But just as our former example this is not a very exciting theory; after all, the equations of

[11] We use the convention $\partial_{[\mu} A_{\nu]} = \frac{1}{2} \Big(\partial_\mu A_\nu - \partial_\nu A_\mu \Big)$.

motion for the vector potential A_μ read $F_{\mu\nu} = 0$, rendering the vector potential pure gauge: $A_\mu = \partial_\mu f$ for some function $f(x^\rho)$. So we can pick different gauge choices for A_μ, and we could say that our equations of motion for the scalar field don't depend on the chosen 'vector potential background'. But up to gauge transformations the solution for the vector field is simply $A_\mu = 0$ (just as the solution of eqn(91) for the metric is, up to 'gauge transformations', $g_{\mu\nu} = \eta_{\mu\nu}$). The theory would become interesting if we would promote A_μ from a mere background structure to a truly dynamical field by introducing dynamics for the vector potential, i.e. the Maxwell equations of motion $\partial_\mu F^{\mu\nu} = j^\nu$. These Maxwell equations are the analogue of the Einstein equations for the system (91).

Trivial general-covariantization is not limited to Special Relativistic field theories. To illustrate this, let's take the (rescaled) diffusion equation for a scalar field $\Phi(t, x^j)$:

$$\partial_t \Phi - \delta^{ij}\partial_i\partial_j\Phi = 0\,. \tag{98}$$

This equation surely is not covariant with respect to general coordinate transformations; it's not even covariant with respect to the non-relativistic Galilei boosts (2)! The time derivative ruins the fun: you can check this by boosting $\partial'_t = \partial_t - v^i\partial_i$ and $\delta'^{ij}\partial'_i\partial'_j = \delta^{ij}\partial_i\partial_j$.[12] It *is* covariant with respect to constant rotations, constant temporal and spatial translations, and the rescaling

$$t' = \lambda^2 t, \quad x'^i = \lambda x^i \tag{99}$$

for arbitrary constants λ. The sophisticated call this group of rotations, translations and rescalings the *Lifshitz-group*. But we can make eqn.(98) general covariant by adding some extra background structure in the form of a vector field $\boldsymbol{n}$:[13]

$$\begin{aligned} &\nabla_{[\mu}n_{\nu]} = 0\,,\\ &R_{\mu\nu\rho\sigma} = 0\,,\\ &n^\mu\nabla_\mu\Phi - \Big(n^\mu n^\nu + g^{\mu\nu}\Big)\nabla_\mu\nabla_\nu\Phi = 0\,. \end{aligned} \tag{100}$$

The second equation enables us to choose Cartesian coordinates such that $g_{\mu\nu} = \eta_{\mu\nu}$ in which the connection coefficients vanish and $\nabla_\mu = \partial_\mu$. The first equation then becomes $\partial_{[\mu}n_{\nu]} = 0$ which basically states that the components n_μ are those of an exact one-form, i.e.

$$n_\mu = \partial_\mu N \tag{101}$$

for some function $N(t, x^i)$. Choosing so-called *adapted coordinates* $N = t$ (similarly to eqn.(54) and figure 6), we get

$$n_\mu = \delta^0_\mu = (1, 0, 0, 0), \qquad n^\mu = \eta^{\mu\nu}n_\nu = (-1, 0, 0, 0)\,. \tag{102}$$

[12] The Schrödinger equation circumvents this problem because the wave function is not a scalar under Galilei boosts. It transforms instead with an extra phase factor. This follows from the fact that the covariance group of the Schrödinger equation is given by the *Schrödinger group*, which has the so-called *Bargmann group* as subgroup. This Bargmann group on its turn is just the Galilei group with a central extension, which describes mass. See e.g. Ballentine's textbook on Quantum Mechanics *Quantum Mechanics: A modern Development* section 3.2 − 3.4 for more details.

[13] The first equation can be derived from an action of the form $S = \int \sqrt{|g|}d^4x\Big(C^{\mu\nu}\nabla_\mu n_\nu\Big)$, where $C^{\mu\nu} = C^{[\mu\nu]}$ being antisymmetric.

The term $\left(n^\mu n^\nu + g^{\mu\nu}\right)$ in the third equation of (100) then acts as a spatial projection operator, becoming effectively the Cartesian metric δ^{ij}. Because the diffusion equation contains a first order time derivative but a second order spatial derivative, i.e. time and spare are 'thorn apart', we need the vector field $\boldsymbol{n}$ on top of the metric $\boldsymbol{g}$. With these two extra structures eqn.(100) constitutes an honest, general covariant version of the non-relativistic diffusion equation. Kretschmann would probably have been delighted with this example.

But perhaps the best example of the fact that General Relativity is not unique in being general covariant came with Elie Cartan around 1923 in the form of so-called *Newton-Cartan theory*, a general-covariant form of ... good old Newtonian gravity, the very theory which Einstein replaced with General Relativity! Cartan achieved this not by merely adding some ad-hoc extra structure as in the last two examples. In Newton-Cartan theory Newtonian gravity is truely described as the curvature of a spacetime, albeit with a non-degenerate metric structure.[14] The theory consists of an 'Einstein field equation' for the geometry and a geodesic equation for point particles. For the details you can consult Misner, Thorne and Wheeler or my PhD-thesis *Newton-Cartan theory revisited*. With all these examples the question now rises: what exactly *is* the meaning of general covariance in the theory of General Relativity?

Well, on its own it's not so special, as we saw. But the theory of General Relativity is not just general covariant, but also *background independent*. It doesn't have extra non-dynamical structure, like the 'fields' or Lagrange multipliers $C^{\mu\nu\rho\sigma}$ and $C^{\mu\nu}$ in the examples above. It *only* uses the metric to describe the background geometry, and this metric is a dynamical field in its own right, obeying the Einstein field equations! Even in Newton-Cartan theory, the whole metrical structure can, after a lot of gauge-fixing (choosing coordinates), *always* be brought back to one single field Φ: the Newtonian potential. After this gauge-fixing one ends up with the group of Galilei-transformations supplementent with accelerations. As such this whole business of making theories general-covariant reminds us a bit of the Stückelberg trick of introducing extra gauge degrees of freedom which later on can be fixed (see e.g. Hinterbichler's notes on Massive Gravity, chapter four). General Relativity is unique in the sense that the general-covariantization *only* entails the introduction of the dynamical metric field. If one approaches General Relativity as a self-interacting theory of spin-2 fields, in which one adds higher order derivative terms to the so-called *Fierz-Pauli theory*, then the gauge transformations (83) pop up as a consistency condition like the gauge transformations pop up in the quantization of spin-1 fields. This perfectly shows us that general covariance is not a 'defining property' of General Relativity! Also, this is a perfect example of how a background *in*dependent theory can be obtained from a background dependent theory like Fierz-Pauli theory. I guess this is how a string theorist would hope to make String Theory explicitly background independent some day.

[14] Actually, this metric structure was added only later, and not by Cartan. Some people who developed Cartan's theory further are Kurt Friedrichs, Georg Dautcourt, Andrzej Trautman and Jürgen Ehlers.

So what then is the exact meaning of *diffeomorphism invariance*?[15] The true marvel of diffeomorphism invariance is revealed in the case where there is no fixed background structure (like Minkowski spacetime in Quantum Field Theory), or 'no prior geometry'. The geometry which is present in General Relativity is determined by the metric, and this is a truly dynamical field on its own. The Einstein field equations have different solutions for the metric which are not related by mere diffeomorphisms. In Newton-Cartan on the other hand, we also have spacetime geometry which is determined by the matter distribution, but there we can *always* choose coordinates such that *all* of that geometry boils down to one single field Φ in a Euclidean space. So all the 'different' solutions tot the Newton-Cartan field equations are related by mere diffeomorphisms, or gauge transformations as field theorists would call it. In this 'no priori geometry with physically different background solutions'-sense General Relativity (or its extensions with higher order derivative terms, supersymmetry etc.) is truly special. Pictorially, if we apply a diffeomorphism to all the dynamical fields in General Relativity, these fields are shifted with respect to spacetime, but spacetime *itself* is represented by the (dynamical!) metric and as such 'shifts along with the ride'. There is no fixed background to hold on to or with respect to shift your fields; the stage is shifted along. You should compare that to applying a diffeomorphism on e.g. the scalar field Φ in a *Minkowski* background. There the Minkowski metric plays the rôle of a fixed stage, with respect to which all tensor fields can be shifted. To rephrase Shakespeare in his play *As You Like it*:

"*All of spacetime is a dynamical stage*
and all the tensor fields merely players."

12 Lecture notes, books, papers and websites

- Sean Carroll, 'Lecture Notes on General Relativity', *https://arxiv.org/abs/gr-qc/9712019*.
- Matthias Blau, 'General Relativity Lecture Notes'.
- Harvey Reall, 'Part 3: General Relativity'.
- Edmund Bertschinger, 'Symmetry Transformations, the Einstein-Hilbert Action, and Gauge Invariance'.
- N. Andersson and G. Comer, 'Relativistic fluid dynamics: physics for many different scales', *https://arxiv.org/abs/gr-qc/2008.12069v1*
- Kurt Hinterbichler, 'Theoretical Aspects of Massive Gravity'.
- Robert M. Wald, 'General Relativity'.
- Anthony Zee, 'Einstein Gravity in a Nutshell'.
- Ray d'Inverno, 'Introducing Einstein's Relativity'.
- Misner, Thorne and Wheeler, 'Gravitation'.

[15] As we saw, passive coordinate transformations are diffeomorphisms on $\mathbb{R}^n$. But that's not the manifold we're usually interested in; that's spacetime!

- M. Nakahara, 'Geometry, Topology and Physics'.
- Tim Maudlin, 'Philosophy of Physics: Space and Time'.
- John Norton, 'General covariance and the foundations of general relativity: eight decades of dispute'.
- John Norton, 'Did Einstein Stumble? The debate over general covariance'.
- Oliver Pooley, 'Background Independence, Diffeomorphism Invariance, and the Meaning of Coordinates' *https://arxiv.org/abs/1506.03512v1*
- Domenico Giulini, 'Some remarks on the notions of general covariance and background independence', *https://arxiv.org/abs/gr-qc/0603087*.
- Roel Andringa-Boxum, 'Newton-Cartan theory revisited'.
- Roel Andringa-Boxum, 'Ruimte, Tijd, Materie' (Dutch for 'Space, Time, Matter')
- John Norton, 'The Hole argument', *https://plato.stanford.edu/entries/spacetime-holearg/*.
- Alan Macdonald, 'Einsteins hole argument,' (American Association of Physics Teachers, Februari 2001).
- 'The role of Active and Passive Diffeomorphism Invariance in GR', *https://physics.stackexchange.com*.
- 'Active/Passive Diffeomorphisms – clarification on Rovelli', *https://www.physicsforums.com*