ENS M1 General Relativity - Lecture 3 - Tensor Calculus

Nick Kaiser

September 25, 2020

Contents

List of Figures

1 Introduction

This lecture develops the formalism of 'tensor calculus'. We closely follow the approach and terminology of Bernard Schutz's textbook.

The first section deals with special relativity in ordinary rectilinear coordinates. The goal here is to make more precise concepts we have already introduced. We introduce basis vectors and their equivalents for 1-forms and tensors and we show how these things transform under boosts and/or rotations. We also introduce the concept of 1-forms and tensors as scalar valued functions of vectors. We briefly discuss symmetries and finish with the definition the derivatives of vector and tensor fields, along with path derivatives, and the various different notations for these things.

The second section extends this to arbitrary curvilinear coordinate systems, where, in the definition of the derivative – the so-called 'covariant derivative' – we have to worry about how the basis vectors vary with position. We define the 'connection' appearing in the covariant derivative and we show how this is related to, and can be determined explicity from, the metric tensor. All of this is still in the domain of flat (i.e. Minkowskian) space-time. All of the results and constructs, however, are still valid on a curved manifold, and it is formally then a relatively straightforward jump – though a conceptually massive one – to differential geometry in curved space-time.

2 Tensor calculus in rectilinear coordinates

2.1 Frames of reference

We concentrated before on frames of reference of observers that are boosted with respect to each other. More generally, reference frames can be rotated with respect to each other also. So these reference frames form a 6-parameter family determined by the boost velocity $\mathbf{v}$ and the three 'Euler angles' defining the spatial rotations.

The physical quantities we deal with may be classified as scalars, vectors or tensors according to how they transform under such changes of coordinate system.

2.2 Lorentz scalars

Lorentz scalars are measurable quantities whose values are independent of the observer's coordinate frame.

2.3 Vectors

The *prototype 4-vector* is the displacement vector – the space-time separation $\vec{\Delta x}$ between two events – with components, in some observer's frame, $\{\Delta x^\alpha\} = (c\Delta t, \Delta x, \Delta y, \Delta z)$.

An entity $\vec{V}$ with 4 components $\{V^\alpha\}$ is a 4-vector if its components transform under coordinate transformations in the same way as the components of $\vec{\Delta x}$. We have seen various examples before such as the particle current density $\vec{n}$, the charge current density $\vec{j}$ and the electromagnetic potential $\vec{A}$.

2.3.1 Basis vectors

We can construct a 4-vector $\vec{V}$ – a frame-independent entity – as

$$\boxed{\vec{V} = V^\alpha \vec{e}_\alpha} \tag{1}$$

where the set $\{V^\alpha\}$ for $\alpha = 0, 1, 2, 3$, are the components and where the set $\{\vec{e}_\alpha\}$ are 4 *basis vectors*. The components may have physical dimensions other than length, but they must all be the same. For example, the components $(cn, \mathbf{v}n)$ of the particle current density are $[n^\alpha] = \mathrm{L}^{-2}\mathrm{T}^{-1}$. The basis vectors are dimensionless constant vectors.

Don't confuse the index α on $\vec{e}_\alpha$ with the index on the component V^α. It is in the wrong place, for one thing, being downstairs. Here α is a *label* that *identifies* which of the 4 basis vectors we are dealing with. Each one of the 4 basis vectors having, itself, 4 components.

We use the notation

$$\vec{V} \underset{\mathrm{O}}{\longrightarrow} V^\alpha \tag{2}$$

to indicate that $\vec{V}$ has components $\{V^\alpha\}$ in the frame O.

We introduce here the additional notation

$$V^\alpha = (\vec{V})^\alpha \tag{3}$$

where the frame involved is indicated by the superscript which may be α, α' etc..

There is, in general, some freedom in the choice of basis. Here we use only the *coordinate basis vectors.* These point along the coordinate axes and have unit length: $\vec{e}_0 \rightarrow (1, 0, 0, 0)$ and $\vec{e}_1 \rightarrow (0, 1, 0, 0)$ ..., or, more succinctly

$$\boxed{(\vec{e}_\alpha)^\beta = \delta^\beta_\alpha.} \tag{4}$$

So the coordinate basis vectors as just like the unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ that point along the x, y, z axes in 3-space, and, in terms of which we construct 3-vectors like $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Equation (4) says that these basis vectors are orthogonal like $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. One would not say that they are *orthonormal* like $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ since the norm of the 0^{th} basis vector is actually negative.

2.3.2 Transformation of coordinates and basis vectors

According to an observer who is boosted and/or rotated with respect to the un-primed observer the prototype vector has components $\{\Delta x^{\alpha'}\} = (c\Delta t', \Delta x', \Delta y', \Delta z')$ where

$$\boxed{\Delta x^{\alpha'} = \Delta x^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha} = \Lambda^{\alpha'}{}_\alpha \Delta x^\alpha} \tag{5}$$

where the matrix $\Lambda^{\alpha'}{}_\alpha$ is independent of position and is the generalisation of the x-directed Lorentz boost matrix we considered, for the most part, before. Note that *translations* of the reference frame – which *are* included in the Poincaré group – are not involved here as we are only concerned with displacements between events.

The components of a 4-vector $\vec{V}$ therefore transform as

$$\boxed{V^{\alpha'} = \frac{\partial x^{\alpha'}}{\partial x^\alpha} V^\alpha} \tag{6}$$

but we demand that $\vec{V}$ be *frame independent* so

$$\vec{V} = V^{\alpha'} \vec{e}_{\alpha'} = \frac{\partial x^{\alpha'}}{\partial x^\alpha} V^\alpha e_{\alpha'} = V^\alpha e_\alpha \tag{7}$$

where, in the last equality, we see that

$$\boxed{\vec{e}_{\alpha'} = \left(\frac{\partial x^{\alpha'}}{\partial x^{\alpha}}\right)^{-1} \vec{e}_{\alpha}} \tag{8}$$

so the basis vectors are 'transformed' using the matrix that is the inverse of that used to transform the components.

We put 'transformed' in quotes here to emphasis that this is *not* expressing the transformation of the *components* of the basis vectors. The components $(\vec{e}_0)^\alpha$ of the basis vector $\vec{e}_0$, for instance, transform under boosts/rotations just like the components of any vector, i.e. according to (6). Rather (8) is a rule for constructing a new, *different*, set of 4 basis vectors $\{\vec{e}_{\alpha'}\}$ from the original set $\{\vec{e}_\alpha\}$.

Now the inverse of the matrix $\partial x^{\alpha'}/\partial x^\alpha$ is just $\partial x^\alpha/\partial x^{\alpha'}$. If this is not obvious, it may help to consider $\{x^\alpha\} = (x, y)$ and $\{x^{\alpha'}\} = (\xi, \eta)$ and write out the matrix product $[\partial(x, y)/\partial(\xi, \eta)][\partial(\xi, \eta)/\partial(x, y)]$. The upper-right component, for example, is

$$\frac{\partial x}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial x}{\partial \eta}\frac{\partial \eta}{\partial y} = \left(\frac{\partial \xi}{\partial y}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial y}\frac{\partial}{\partial \eta}\right) x \tag{9}$$

but the operator in parentheses is just $\partial/\partial y$. So this is $\partial x/\partial y = 0$. And doing the same for the other components we get the matrix $[\partial(x, y)/\partial(x, y)]$, whose components are the identity matrix. QED.

Replacing $\partial x^{\alpha'}/\partial x^\alpha \Rightarrow \Lambda^{\alpha'}{}_\alpha$ and $(\partial x^{\alpha'}/\partial x^\alpha)^{-1} = \partial x^\alpha/\partial x^{\alpha'} \Rightarrow \Lambda^\alpha{}_{\alpha'}$ the transformation laws for the components and the basis vectors are

$$\boxed{V^{\alpha'} = \Lambda^{\alpha'}{}_\alpha V^\alpha} \qquad \text{and} \qquad \boxed{\vec{e}_{\alpha'} = \Lambda^\alpha{}_{\alpha'}\vec{e}_\alpha}. \tag{10}$$

What we are doing here is precisely analogous to what we do in Euclidean 3-space with a vector which we write as $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. If we rotate the coordinate system, nothing happens to any of the vectors here. But is is *convenient* to construct a new set of basis vectors $\mathbf{i}', \mathbf{j}', \mathbf{k}'$ that point along the new, rotated, axes and in terms of which the *same* vector $\mathbf{r}$ is given by $\mathbf{r} = x'\mathbf{i}' + y'\mathbf{j}' + x'\mathbf{k}'$. Thus it should be that the components of the transformed basis vectors are $(\vec{e}_{\alpha'})^{\beta'} = \delta^{\beta'}_{\alpha'}$. This is readily verified[1] since

$$(\vec{e}_{\alpha'})^{\beta'} = (\Lambda^\alpha{}_{\alpha'}\vec{e}_\alpha)^{\beta'} = \Lambda^\alpha{}_{\alpha'}(\vec{e}_\alpha)^{\beta'} = \Lambda^\alpha{}_{\alpha'}\Lambda^{\beta'}{}_\beta(\vec{e}_\alpha)^\beta = \Lambda^\alpha{}_{\alpha'}\Lambda^{\beta'}{}_\beta\delta^\beta_\alpha = \Lambda^{\beta'}{}_\alpha\Lambda^\alpha{}_{\alpha'} = \delta^{\beta'}_{\alpha'}. \tag{11}$$

So the transformed set of basis vectors are indeed an orthogonal set aligned with the new axes.

2.3.3 The norm and scalar product of vectors

Just as the invariant interval (proper distance squared) between two events is $\Delta s^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = \eta_{\alpha\beta}\Delta x^\alpha \Delta x^\beta$, where, as always $\eta_{\alpha\beta} = \mathrm{diag}\{-1, 1, 1, 1\}$ is the Minkowski metric, we can define the *norm* of a vector with components A^α to be the frame invariant (i.e. scalar)

$$\boxed{\vec{A}\cdot\vec{A} = \eta_{\alpha\beta}A^\alpha A^\beta} \tag{12}$$

and which we will also denote by $\vec{A}^2$.

The norm of the sum of two vectors is $\vec{A} + \vec{B}$ is $\vec{A}^2 + 2\vec{A}\cdot\vec{B} + \vec{B}^2$, which is a scalar, as are $\vec{A}^2$ and $\vec{B}^2$, so $\vec{A}\cdot\vec{B}$ is also a scalar – it is the called *scalar product* of $\vec{A}$ and $\vec{B}$

$$\boxed{\vec{A}\cdot\vec{B} = \eta_{\alpha\beta}A^\alpha B^\beta.} \tag{13}$$

We found earlier that the components of the Minkowski metric are invariant under boosts. The same is true for spatial rotations (as the spatial part of $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$ is just $\mathrm{diag}(1, 1, 1) = \delta_{ij}$).

[1] In the first step we apply the rule for transforming the basis vectors. In the second we invoke linearity. In the third we apply the rule for transforming the components of a vector. In the 5th we use $(\vec{e}_\alpha)^\beta = \delta^\beta_\alpha$ and in the last we use the fact that $\Lambda^{\alpha'}{}_\alpha$ and $\Lambda^\alpha{}_{\alpha'}$ are inverses of one another.

2.3.4 The scalar products of the basis vectors

Since the components of the basis vectors are $(\vec{e}_\mu)^\alpha = \delta^\alpha_\mu$ the 4x4 matrix of scalar products of the basis vectors is, from (4) and (13),

$$\boxed{\vec{e}_\mu \cdot \vec{e}_\nu = \eta_{\alpha\beta}\delta^\alpha_\mu\delta^\beta_\nu = \eta_{\mu\nu}.} \tag{14}$$

This is an illustration of something we will see later. If we write $\vec{A}\cdot\vec{B} = \mathbf{g}(\vec{A},\vec{B})$ where $\mathbf{g}(\ ,\ )$ is a linear function of it's arguments and is some geometric (i.e. frame independent) object, then we obtain its components – the matrix of numbers that define it – by applying it to the pairs of basis vectors. Here the components are $g_{\alpha\beta} = \eta_{\alpha\beta}$.

2.4 1-forms

Previously we drew the distinction between the contravariant components of a vector like V^α and the covariant components $V_\alpha = \eta_{\alpha\beta}V^\alpha$, which are the same except for a change of sign of the 0^{th} component. And we said that covariant vectors are sometimes called '1-forms'. Note that 'dotting' the covariant and contravariant components gives the scalar product $\vec{V}^2 = V_\alpha V^\alpha$.

Here we take a slightly different, but equivalent, approach which may seem at first to be a bit heavy handed, but which, especially once we come to curvilinear coordinates, helps by providing a precise language. This is the geometrical approach popularised by MTW.

2.4.1 Definition of 1-forms and the contraction

We define 1-forms as linear maps of a vectors to scalars. We denote them with a tilde (e.g. $\tilde{p}$ denotes a 1-form). As a function this is $\tilde{p}(\ )$, whose value, given a vector $\vec{V}$ as its argument, is a scalar. Like a vector, a 1-form is a geometric entity, which we think of as having a frame-independent existence.

We define the α^{th} component of $\tilde{p}$ to be the value of $\tilde{p}(\ )$ when applied to the α^{th} basis vector

$$\boxed{p_\alpha = \tilde{p}(\vec{e}_\alpha)} \tag{15}$$

And we also use the notation

$$\boxed{p_\alpha = (\tilde{p})_\alpha} \tag{16}$$

to denote components, much like $V^\alpha = (\vec{V})^\alpha$ for vectors.

Linearity implies $\tilde{p}(\vec{A}) = \tilde{p}(A^\alpha\vec{e}_\alpha) = A^\alpha\tilde{p}(\vec{e}_\alpha)$, or

$$\boxed{\tilde{p}(\vec{A}) = A^\alpha p_\alpha} \tag{17}$$

which we call the *contraction* of $\vec{A}$ with $\tilde{p}$. It is effectively the same as the scalar product, but it is actually a more *primitive* operation than the scalar product as it does not require that there be a metric.

Comparing this with the definition of the scalar product (13) we see that $\tilde{p}(\vec{A}) = \vec{p}\cdot\vec{A}$ if $p_\alpha = \eta_{\alpha\beta}p^\alpha$.

2.4.2 Visualising 1-forms and vectors

The 'prototype' 1-form is the 4-gradient of a scalar field $\phi(\vec{x})$ which we denote by $\tilde{d}\phi$:

$$\boxed{\tilde{d}\phi \longrightarrow \partial_\alpha\phi} \tag{18}$$

or, equivalently, $(\tilde{d}\phi)_\alpha = \partial_\alpha\phi = \phi_{,\alpha}$, and where the components of the derivative operator here are $\{\partial_\alpha\} = (c^{-1}\partial_t, \boldsymbol{\nabla})$.

If we make the contraction with some displacement vector we get $\tilde{d}\phi(\vec{\Delta x}) = \Delta x^\alpha\partial_\alpha\phi$ which is just the change in ϕ over the displacement $\vec{\Delta x}$.

We may visualise $\tilde{d}\phi$ as a stack of locally parallel 3-surfaces of constant ϕ. I.e. the analogue of 1D contours on a 2D map, with the 'magnitude' of $\tilde{d}\phi$ being proportional to density of, or the inverse spacing between, the 3-surfaces. This is illustrated in figure 1.

If we visualise the prototypical vector $\vec{\Delta x}$ as an arrow then the scalar $\tilde{d}\phi(\vec{\Delta x}) = \phi_{,\alpha}\Delta x^\alpha$ is the number of iso-ϕ surfaces 'pierced' by the vector $\vec{\Delta x}$

It is very natural to represent the 4-momentum of a particle, or a wave-packet, as a 1-form. For example, for a family of particles emerging from a common starting point, the Hamilton-Jacobi formalism tells us that the relativistic 3-momentum and energy (in the 'lab'-frame say) are

$$\mathbf{p} = \boldsymbol{\nabla} S(t, \mathbf{x}) \qquad \text{and} \qquad E = -\partial S(t, \mathbf{x})/\partial t \tag{19}$$

where $S(t, \mathbf{x})$ is the action. But the action is a Lorentz scalar. So $p_\alpha = (-E/c, \mathbf{p}) = \partial_\alpha S$ are the components of a 1-form:

$$\tilde{p} = \tilde{d}S \underset{\text{lab}}{\longrightarrow} (-E/c, \mathbf{p}). \tag{20}$$

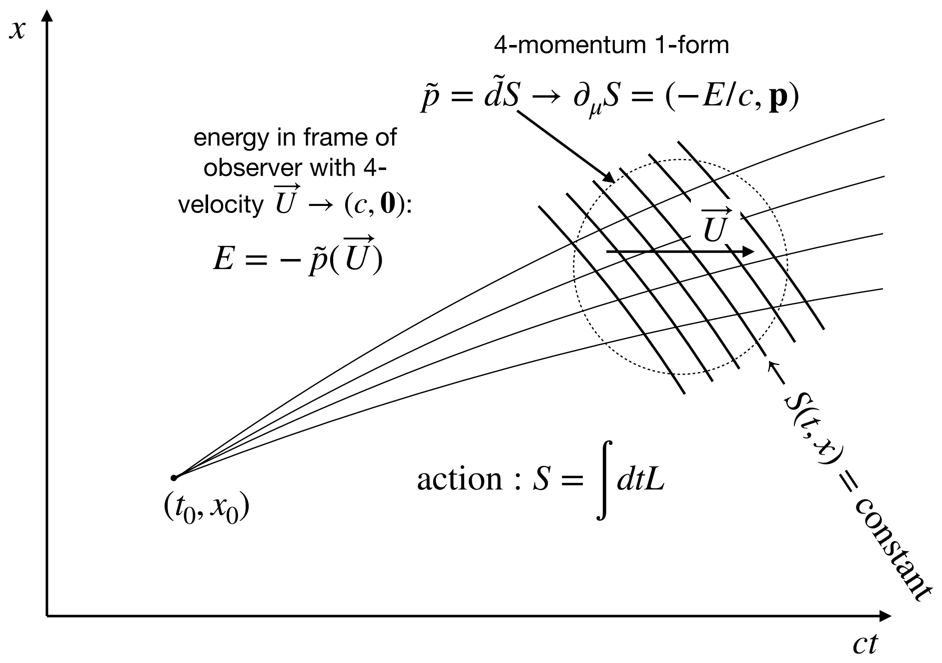

Figure 1: The 4-momentum 1-form. The thin curves are trajectories of particles emanating from a common starting point. The heavy curves are surfaces of constant action. The Hamilton-Jacobi formalism tells us that $\mathbf{p} = \boldsymbol{\nabla} S$ and $E = -\partial_t S$ where the action S, for a relativistic particle, is a Lorentz scalar. The arrow indicates the world-line of an observer who happens to be at rest in the chosen reference frame. The particle energy that that observer would measure is $E = -\tilde{p}(\vec{U})$ where $\tilde{p}(\vec{U})$ is the number of iso-S surfaces pierced by the arrow $\vec{U}$.

If we take this 1-form and apply it to the 4-velocity of an observer who is at rest in the lab frame (i.e. for whom $\vec{U} \underset{\text{lab}}{\longrightarrow} (c, \mathbf{0})$) we get

$$\boxed{\tilde{p}(\vec{U}) = -E} \tag{21}$$

the lab-frame energy.

But this is a scalar equation so it tells us that, were we to use the 4-velocity of some other observer here, we would get minus the energy that that observer would see.

The same result applies quantum mechanically for the particle energy measured by an observer from the wave-function of a particle prepared in this state (well defined initial location but somewhat uncertain momentum). The wave function here being $\psi \sim e^{iS/\hbar}$ and the energy and momentum operators being $E = \hbar\partial_t$ and $\mathbf{p} = -\hbar\boldsymbol{\nabla}$.

Also, if we have a classical nearly monochromatic wave packet with $\phi(\vec{x}) = \phi_0 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)$, the total energy and momentum for the packet are proportional to $\tilde{k} \rightarrow (-\omega/c, \mathbf{k})$ and the phase is $\mathbf{k} \cdot \mathbf{x} - \omega t = \tilde{k}(\vec{x})$.

2.4.3 The basis for 1-forms

In a similar manner to what we did for vectors, we can express a 1-form $\tilde{p}$ as a sum of components times basis 1-forms:

$$\boxed{\tilde{p} = p_\alpha \tilde{\omega}^\alpha} \tag{22}$$

where we have a set of 4 basis 1-forms $\{\tilde{\omega}^\alpha\}$.

As with vectors, there is, in principle, freedom for how we choose the basis, but we have already fixed the components by stipulating that they are given by letting $\tilde{p}$ act on the basis vectors: $p_\alpha \equiv \tilde{p}(\vec{e}_\alpha)$. So with that choice, we may ask: what are the basis 1-forms?

To answer this, we let $\tilde{p} = p_\alpha \tilde{\omega}^\alpha$ act on $\vec{A} = A^\beta \vec{e}_\beta$ to obtain $\tilde{p}(\vec{A}) = p_\alpha A^\beta \tilde{\omega}^\alpha(\vec{e}_\beta)$. But we also have $\tilde{p}(\vec{A}) = p_\alpha A^\alpha$ which implies that $\tilde{\omega}^\alpha(\vec{e}_\beta) = \delta^\alpha_\beta$. But any 1-form acting on the β^{th} basis vector is the β^{th} component of that 1-form. I.e. $\tilde{\omega}^\alpha(\vec{e}_\beta) = (\tilde{\omega}^\alpha)_\beta$ so the components of the 1-form basis are $\tilde{\omega}^0 \rightarrow (1,0,0,0)$ and $\tilde{\omega}^1 \rightarrow (0,1,0,0)$ etc. or, more succinctly

$$\boxed{(\tilde{\omega}^\alpha)_\beta = \delta^\alpha_\beta} \tag{23}$$

so the same as the components of the basis vectors.

2.4.4 Transformation of 1-form components and basis 1-forms

Obtaining the transformation for 1-form components is straightforward. A 1-form (a frame independent entity) acting on a vector $\vec{A}$ is a scalar so $\tilde{p}(\vec{A}) = p_\alpha A^\alpha = p_{\alpha'} A^{\alpha'}$. The components of $\vec{A}$ transform according to $A^{\alpha'} = \Lambda^{\alpha'}{}_\alpha A^\alpha$ so the components p_α must transform according to the inverse of the matrix:

$$\boxed{p_{\alpha'} = \Lambda^\alpha{}_{\alpha'} p_\alpha} \tag{24}$$

which looks a lot like the transformation of the basis vectors: $\vec{e}_{\alpha'} = \Lambda^\alpha{}_{\alpha'} \vec{e}_\alpha$, but it really, of course, quite different as the things being transformed here are numbers, not vectors.

To get the transformation of the basis 1-forms we simply appeal to the fact that $\tilde{p} = p_\alpha \tilde{\omega}^\alpha = p_{\alpha'} \tilde{\omega}^{\alpha'}$ from which, with the transformation law for the components above gives

$$\boxed{\tilde{\omega}^{\alpha'} = \Lambda^{\alpha'}{}_\alpha \tilde{\omega}^\alpha} \tag{25}$$

i.e. opposite to the transformation for the components p_α of the 1-form and also opposite to the transformation for the basis vectors $\vec{e}_\alpha$ but the same as for the vector components V^α.

A simple trick for getting the transformation laws for components or bases for either vectors or 1-forms is to remember that the first (i.e. row) index on the Λ-matrix always goes upstairs. The indices are then fixed by the requirement that repeated indices must be upstairs-downstairs pairs and non-repeated indices should appear in the same place on both sides of an equation.

2.4.5 Magnitude and scalar products of 1-forms

- we define the *squared magnitude* or norm of a 1-form $\tilde{p}^2 \equiv \tilde{p} \cdot \tilde{p}$ to be the same as that of the corresponding vector
  - $\tilde{p}^2 = \vec{p}^2 = \eta_{\alpha\beta} p^\alpha p^\beta$
- or, since $p^\alpha = \eta^{\alpha\nu} p_\nu$,
  - $\tilde{p}^2 = \eta^{\alpha\beta} p_\alpha p_\beta$
- We define the *scalar product of 1-forms:*
  - $\tilde{p} \cdot \tilde{q} \equiv [(\tilde{p} + \tilde{q})^2 - \tilde{p}^2 - \tilde{q}^2]/2$
  - in terms of components, $\tilde{p} \cdot \tilde{q} = -p_0 q_0 + p_1 q_1 \ldots$
  - which is the same as the contraction $\tilde{p}(\vec{q}) = p_0 q^0 + p_1 q^1 \ldots$

2.5 Tensors

Previously, we said that e.g. a matrix $T^{\alpha\beta}$ is the matrix of (contravariant) components of a tensor $\mathbf{T}$ if the $T^{\alpha\beta}$ transform as $T^{\alpha'\beta'} = \Lambda^{\alpha'}{}_\alpha \Lambda^{\beta'}{}_\beta T^{\alpha\beta}$. That means, for example, that dotting this with the (covariant) components $V_\alpha = \eta_{\alpha\beta} V^\beta$ of a vector gives the components of a contravariant vector $U^\alpha = T^{\alpha\beta} V_\beta$.

Here we will take an alternative, but equivalent, approach which is to define tensors as scalar valued linear functions of vectors (or 1-forms as appropriate). The main thing we get from this is bases for tensors as outer products of vector or 1-form bases. Let's start with covariant tensors.

2.5.1 The definition of a $\binom{0}{N}$ tensor

A $\binom{0}{N}$ tensor is a *function* of N vectors that returns a scalar and is *linear* in each of its arguments.

A 1-form has these properties (for $N = 1$) and so is a $\binom{0}{1}$ tensor.

The scalar product also has these properties (for $N = 2$) and is therefore an example of a $\binom{0}{2}$ tensor. We call this function the *metric tensor*, and denote it by $\mathbf{g}$. Note that this is not to be confused with a 3-vector, and is and definitely not the Newtonian gravity vector.

The metric is defined by the rule

$$\boxed{\mathbf{g}(\vec{A}, \vec{B}) = \vec{A} \cdot \vec{B}.} \tag{26}$$

We can think of $\mathbf{g}(\ ,\ )$, with empty 'slots' for two vectors, as another frame-independent entity.

As with 1-forms, we define the *components* of a tensor to be its value when applied to arguments that are basis vectors. As there are 4 basis vectors, a $\binom{0}{N}$ tensor has 4^N components, and these are labelled by N greek subscript 'indices'.

For example, the components of the metric tensor are

$$g_{\alpha\beta} = \mathbf{g}(\vec{e}_\alpha, \vec{e}_\beta) = \vec{e}_\alpha \cdot \vec{e}_\beta = \eta_{\alpha\beta}. \tag{27}$$

2.5.2 Outer products and bases for tensors

Another example of a $\binom{0}{2}$ tensor is the *outer product* of two 1-forms $\tilde{p} \otimes \tilde{q}$, which is defined such that, acting on two vectors $\vec{A}$ and $\vec{B}$, it returns the scalar $\tilde{p}(\vec{A})\tilde{q}(\vec{B})$. Previously, we had the stress-energy tensor for particles: $\mathbf{T} = \int \frac{d^3p}{p^0} f(\mathbf{r}, \mathbf{p}) \vec{p}\vec{p}$, the combination $\vec{p}\vec{p}$ being an outer product of two vectors. Using the notation here, we could write $\mathbf{T} = \int \frac{d^3p}{p^0} f(\mathbf{r}, \mathbf{p}) \tilde{p} \otimes \tilde{p}$ which would be a $\binom{0}{2}$ tensor.

We can use the (4 by 4 matrix of) outer products of the basis 1-forms $\tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta$ as a 'basis' for $\binom{0}{2}$ tensors. I.e. we can express a general $\binom{0}{2}$ tensor $\mathbf{T}(\ ,\ )$ as a linear double sum of these

$$\mathbf{T}(\ ,\ ) = T_{\alpha\beta}\, \tilde{\omega}^\alpha(\ ) \otimes \tilde{\omega}^\beta(\ ) \tag{28}$$

where the μ, ν component of $\mathbf{T}$ – defined as usual as the (scalar) value of the result of applying this function to two basis vectors $\vec{e}_\mu$ and $\vec{e}_\nu$ – is

$$T_{\alpha\beta}\, \tilde{\omega}^\alpha(\vec{e}_\mu)\tilde{\omega}^\beta(\vec{e}_\nu) = T_{\alpha\beta}\delta^\alpha_\mu \delta^\beta_\nu = T_{\mu\nu} \tag{29}$$

The transformation of the $\binom{0}{2}$ basis tensor $\tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta$ is straightforward: we simply replace the basis 1-forms by their primed versions $\tilde{\omega}^{\alpha'} = \Lambda^{\alpha'}{}_\alpha \tilde{\omega}^\alpha$.

The transformation of the components comes from the requirement that $\mathbf{T}$ be frame independent, so $T_{\alpha'\beta'}\tilde{\omega}^{\alpha'} \otimes \tilde{\omega}^{\beta'} = T_{\alpha\beta}\tilde{\omega}^\alpha \otimes \tilde{\omega}^\beta$ which gives

$$\boxed{T_{\alpha'\beta'} = \Lambda^\alpha{}_{\alpha'} \Lambda^\beta{}_{\beta'} T_{\alpha\beta}.} \tag{30}$$

It is straightforward to extend this to $N > 2$.

2.5.3 The metric as a mapping of a vector to a 1-form and vice versa

Consider $\mathbf{g}(\vec{V},\ )$. It is evidently a 1-form, since, given another vector as the missing argument, it returns a scalar, and it is linear function of that argument. Let's call it $\tilde{V} = \mathbf{g}(\vec{V},\ )$ and ask: what are its components?

Its α^{th} component is $V_\alpha = \tilde{V}(\vec{e}_\alpha) = \mathbf{g}(\vec{V}, \vec{e}_\alpha) = \vec{e}_\alpha \cdot \vec{V} = \vec{e}_\alpha \cdot V^\beta \vec{e}_\beta = \eta_{\alpha\beta} V^\beta$.

So $V_\alpha = \eta_{\alpha\beta} V^\beta$ and $V^\alpha = \eta^{\alpha\beta} V_\beta$, where $\eta^{\alpha\beta}$ is the inverse of $\eta_{\alpha\beta}$ and which also has components $\text{diag}(-1, 1, 1, 1)$.

So this agrees with what we found previously: matrix multiplication by $\eta_{\alpha\beta}$ and $\eta^{\alpha\beta}$ are respectively index lowering and raising operators: If $\vec{V}$ has components (a, b, c, d) (i.e. $\vec{V} \to (a, b, c, d)$) then $\tilde{V} = \mathbf{g}(\vec{V},\ ) \to (-a, b, c, d)$.

We can similarly formalise the mapping affected by the inverse of the metric as follows. Consider the function $\mathbf{g}^{-1}(\ ,\ )$ defined such that

$$\boxed{\mathbf{g}^{-1}(\tilde{p}, \tilde{q}) = \tilde{p} \cdot \tilde{q}.} \tag{31}$$

Now $\tilde{p}\cdot\tilde{q} = \vec{p}\cdot\vec{q} = \eta_{\alpha\beta}p^\alpha q^\beta = \eta_{\alpha\beta}\eta^{\alpha\mu}p_\mu\eta^{\beta\nu}q_\nu$ so the components are $\mathbf{g}^{-1} \to \eta_{\alpha\beta}\eta^{\alpha\mu}\eta^{\beta\nu} = \eta^{\mu\nu}$.

Given a 1-form $\tilde{p}$, consider $\mathbf{g}^{-1}(\tilde{p},\ )$. This is a vector, since, given another 1-form to fill the empty slot, we obtain a scalar. Let's call it $\vec{p}$. Letting that act on the α^{th} basis 1-form $\tilde{\omega}^\alpha$ gives its α^{th} component: $p^\alpha = \vec{p}(\tilde{\omega}^\alpha) = \mathbf{g}^{-1}(\tilde{p},\tilde{\omega}^\alpha) = \mathbf{g}^{-1}(p_\beta\tilde{\omega}^\beta,\tilde{\omega}^\alpha) = p_\beta\mathbf{g}^{-1}(\tilde{\omega}^\beta,\tilde{\omega}^\alpha) = p_\beta(\mathbf{g}^{-1})^{\beta\alpha} = \eta^{\alpha\beta}p_\beta$. So $\mathbf{g}^{-1}$ is indeed the index raising operator.

2.5.4 The definition of a $\binom{M}{0}$ tensor

A $\binom{M}{0}$ tensor is a linear map from M 1-forms to a scalar. We've just seen an example: the inverse of the metric tensor $\mathbf{g}^{-1}(\ ,\ )$ is a $\binom{2}{0}$ tensor.

For example, $\mathbf{T} = \int \frac{d^3p}{p^0} f(\mathbf{r},\mathbf{p})\vec{p}\otimes\vec{p}$ is a $\binom{2}{0}$ tensor.

Another example of a $\binom{2}{0}$ tensor is the outer product of two vectors $\vec{U}\otimes\vec{V}$, which is defined such that its value, given the two 1-form arguments $\tilde{p}$ and $\tilde{q}$ is $\vec{U}\otimes\vec{V}(\tilde{p},\tilde{q}) = \vec{U}(\tilde{p})\vec{V}(\tilde{q})$.

A general $\binom{2}{0}$ tensor $\mathbf{T}$ can be expressed as a linear combination of the 4x4 array of bases $\vec{e}_\alpha\otimes\vec{e}_\beta$ with coefficients $T^{\alpha\beta}$:

$$\mathbf{T}(\ ,\ ) = T^{\alpha\beta}\vec{e}_\alpha(\ )\otimes\vec{e}_\beta(\ ) \tag{32}$$

and the value of the function $\mathbf{T}(\ ,\ )$ given arguments $\tilde{\omega}^\mu$ and $\tilde{\omega}^\nu$ – i.e. the μ,ν *component* of $\mathbf{T}$ – is $T^{\mu\nu}$.

The transformation law for the components is obtained just as before by writing the same tensor in terms of the primed frame tensor basis:

$$\boxed{T^{\alpha'\beta'} = \Lambda^{\alpha'}{}_\alpha\Lambda^{\beta'}{}_\beta T^{\alpha\beta}.} \tag{33}$$

And the extension to $M > 2$ is also straightforward.

2.5.5 The definition of a $\binom{M}{N}$ tensor

Finally, a 'mixed' $\binom{M}{N}$ tensor is a linear map of M 1-forms and N vectors to a scalar.

Consider, for example, a $\binom{2}{1}$ tensor $\mathbf{T}(\ ,\ ,\ )$ that takes 2 1-form arguments in the first two slots and a vector in the third. The ordering is significant, but unfortunately it is hidden in the geometric notation.

This can be expressed as a linear combination of outer products of the 4^3 bases $\vec{e}_\alpha\otimes\vec{e}_\beta\otimes\tilde{\omega}^\gamma$:

$$\mathbf{T} = T^{\alpha\beta}{}_\gamma\vec{e}_\alpha\otimes\vec{e}_\beta\otimes\tilde{\omega}^\gamma \tag{34}$$

with components $T^{\alpha\beta}{}_\gamma = \mathbf{T}(\tilde{\omega}^\alpha,\tilde{\omega}^\beta,\vec{e}_\gamma)$ which transform as

$$\boxed{T^{\alpha'\beta'}{}_{\gamma'} = \Lambda^{\alpha'}{}_\alpha\Lambda^{\beta'}{}_\beta\Lambda^\gamma{}_{\gamma'}T^{\alpha\beta}{}_\gamma.} \tag{35}$$

2.5.6 Raising and lowering tensors with $\mathbf{g}(\ ,\ )$ and $\mathbf{g}^{-1}(\ ,\ )$

Just as the metric maps a vector $\vec{V}$ to a corresponding 1-form $\tilde{V} = \mathbf{g}(\vec{V},\ )$ and we can convert it back with $\vec{V} = \mathbf{g}^{-1}(\tilde{V},\ )$ we should be able to do the same kind of thing with tensors; i.e. we should be able to map a $\binom{M}{N}$ tensor to a corresponding $\binom{M-1}{N+1}$ tensor. This is done in the following manner:

Consider the $\binom{1}{2}$ tensor $\mathbf{S}$ constructed from a $\binom{2}{1}$ tensor $\mathbf{T}$ as

$$\mathbf{S}(\ ,\ ,\ ) = \mathbf{T}(\ ,\mathbf{g}(\ ,\ ),\ ) \tag{36}$$

with the understanding that the second argument of $\mathbf{g}(\ ,\ )$ is left empty.

Thus, when $\mathbf{S}$ is fed 3 arguments $\tilde{A},\vec{B},\vec{C}$, the second slot of $\mathbf{T}$ gets fed $\tilde{B} = \mathbf{g}(\vec{B},\ )$, which an appropriate argument for this slot.

So $\mathbf{S}(\ ,\vec{V},\ )$ and $\mathbf{T}(\ ,\tilde{V},\ )$ are identical $\binom{1}{1}$ tensors for any $\vec{V}$.

The tensor $\mathbf{S}$ can be expressed as a linear combination of the appropriate tensor bases:

$$\mathbf{S} = S^\alpha{}_{\beta\gamma}\vec{e}_\alpha\otimes\tilde{\omega}^\beta\otimes\tilde{\omega}^\gamma \tag{37}$$

where the components are $S^\alpha{}_{\beta\gamma} = \mathbf{S}(\tilde{\omega}^\alpha,\vec{e}_\beta,\vec{e}_\gamma) = \mathbf{T}(\tilde{\omega}^\alpha,\mathbf{g}(\vec{e}_\beta,\ ),\vec{e}_\gamma)$

But $\mathbf{g} = \eta_{\mu\nu}\,\tilde{\omega}^\mu \otimes \tilde{\omega}^\nu$ so $\mathbf{g}(\vec{e}_\beta, \;) = \eta_{\mu\nu}\tilde{\omega}^\mu(\vec{e}_\beta)\tilde{\omega}^\nu = \eta_{\beta\nu}\tilde{\omega}^\nu$, and $S^\alpha{}_{\beta\gamma} = \mathbf{T}(\tilde{\omega}^\alpha, \eta_{\beta\nu}\tilde{\omega}^\nu, \vec{e}_\gamma)$ so, invoking linearity,

$$\boxed{S^\alpha{}_{\beta\gamma} = \eta_{\beta\nu}T^{\alpha\nu}{}_\gamma.} \tag{38}$$

So the components of **S** are indeed obtained from those of **T** simply by applying the index lowering matrix – the Minkowski metric – to the appropriate index (the second one here).

While **S** and **T** are different tensors they are clearly completely determined from each other, and we use the same symbol for their components and say, for example, $T^\alpha{}_{\beta\gamma} = \eta_{\beta\nu}T^{\alpha\nu}{}_\gamma$.

2.5.7 Symmetries

- a $\binom{0}{2}$ tensor $\mathbf{T}(\;,\;)$ for which $\mathbf{T}(\vec{A}, \vec{B}) = \mathbf{T}(\vec{B}, \vec{A}) \quad \forall \quad \vec{A}, \vec{B}$ is *symmetric*
- and given an arbitrary tensor $\mathbf{H}(\;,\;)$ one can define a *symmetrised* version $\mathbf{H}_{(S)}(\;,\;)$ by
  - $\mathbf{H}_{(S)}(\vec{A}, \vec{B}) = [\mathbf{H}(\vec{A}, \vec{B}) + \mathbf{H}(\vec{B}, \vec{A})]/2$
- and an antisymmetrised version $\mathbf{H}_{(A)}(\;,\;)$ by
  - $\mathbf{H}_{(A)}(\vec{A}, \vec{B}) = [\mathbf{H}(\vec{A}, \vec{B}) - \mathbf{H}(\vec{B}, \vec{A})]/2$
- in component form these expressions are
  - $H_{(S)\alpha\beta} = [h_{\alpha\beta} + h_{\beta\alpha}]/2 = h_{\{\alpha\beta\}}/2$
  - $H_{(A)\alpha\beta} = [h_{\alpha\beta} - h_{\beta\alpha}]/2 = h_{[\alpha\beta]}/2$
- where we have introduced the commutator $h_{[\alpha\beta]}$ and anti-commutator $h_{\{\alpha\beta\}}$
- one can express the original tensor as $\mathbf{H} = \mathbf{H}_{(S)} + \mathbf{H}_{(A)}$
- Symmetry properties are frame invariant.

2.6 Derivatives and path derivatives of vectors and tensors

Given a scalar field $\phi(\vec{x})$, its gradient is the 1-form field $\tilde{d}\phi$. So the derivative of ϕ (a scalar being a $\binom{0}{0}$ tensor if you will) is the $\binom{0}{1}$ tensor $\tilde{d}\phi \rightarrow (c^{-1}\phi_{,t}, \boldsymbol{\nabla}\phi)$.

Given a parameterised path $\vec{x}(\lambda) \rightarrow (ct(\lambda), \mathbf{x}(\lambda))$ with *tangent vector* $\vec{U} \equiv d\vec{x}/d\lambda \rightarrow (cdt/d\lambda, d\mathbf{x}/d\lambda)$, the rate of change of ϕ along the path is

$$d\phi/d\lambda = \phi_{,\mu} d\vec{x}/d\lambda = \phi_{,\mu}U^\mu = \tilde{d}\phi(\vec{U}). \tag{39}$$

This would apply for an observer, with $\lambda = \tau$, the proper time, and $\vec{U}$ being the 4-velocity.

The same is true for tensor (or vector) fields. The derivative of a $\binom{M}{N}$ tensor field is a $\binom{M}{N+1}$ tensor field which, operating on the tangent vector $\vec{U}$, gives the $\binom{M}{N}$ tensor that is the rate of change of the original tensor wrt λ along the path.

Consider, for example, the $\binom{1}{1}$ tensor field

$$\boxed{\mathbf{T}(\vec{x}) = T^\alpha{}_\beta(\vec{x})\vec{e}_\alpha \otimes \tilde{\omega}^\beta,} \tag{40}$$

the vector and 1-form bases being independent of location in flat space-time.

We define its *path derivative* as

$$\frac{d\mathbf{T}}{d\lambda} = \lim_{\Delta\lambda \to 0} \frac{\mathbf{T}(\vec{x}(\lambda + \Delta\lambda)) - \mathbf{T}(\vec{x}(\lambda))}{\Delta\lambda}. \tag{41}$$

This is evidently a $\binom{1}{1}$ tensor like **T** itself:

$$\frac{d\mathbf{T}}{d\lambda} = \frac{dT^\alpha{}_\beta}{d\lambda}\vec{e}_\alpha \otimes \tilde{\omega}^\beta. \tag{42}$$

But $dT^{\alpha}{}_{\beta}/d\lambda = T^{\alpha}{}_{\beta,\gamma}U^{\gamma}$, so $T^{\alpha}{}_{\beta,\gamma}$ are the components of a $\binom{1}{2}$ tensor, which we will denote by $\nabla\mathbf{T}$ (even though $\tilde{d}\mathbf{T}$ would seem more logical):

$$\boxed{\nabla\mathbf{T} \equiv T^{\alpha}{}_{\beta,\gamma}\vec{e}_{\alpha} \otimes \tilde{\omega}^{\beta} \otimes \tilde{\omega}^{\gamma}} \tag{43}$$

which, acting on (or, in component form, contracted with) the tangent vector $\vec{U}$ gives $d\mathbf{T}/d\lambda$.

Rather that writing $d\mathbf{T}/d\lambda = \nabla\mathbf{T}(\vec{U})$, which is somewhat ambiguous, as it is not clear what index of $\nabla\mathbf{T}$ is getting contracted with $\vec{U}$, we tend to use the notation

$$d\mathbf{T}/d\lambda = \nabla_{\vec{U}}\mathbf{T} \to U^{\gamma}T^{\alpha}{}_{\beta,\gamma} \tag{44}$$

where we note that if we work in terms of components there is no ambiguity.

The above equation defines the scalar differential operator

$$\boxed{\nabla_{\vec{U}} \to U^{\gamma}\partial_{\gamma}} \tag{45}$$

which gives $d/d\lambda$ along the path $\vec{x}(\lambda)$, or, in the case of an observer, along the world-line with 4-velocity $\vec{U}$, of any entity.

A key assumption in the above is that the basis vectors and basis 1-forms are independent of location. This is valid in flat space (or flat space-time) *and* using Cartesian (or Minkowski) coordinates and their related coordinate basis vectors and 1-forms.

In the following section we see how things change – particularly the derivative operator – if we work in arbitrary curvilinear coordinates.

3 Tensor calculus in curvilinear coordinates

In this section we will generalise the foregoing to arbitrary curvilinear coordinate systems. I.e. we will define vectors, 1-forms and tensors as linear combinations of bases times components. The important change here is that the coordinate transformation matrices and basis vectors and basis 1-forms become, in general, position dependent. We will describe the roles of the metric and its inverse as a way of measuring lengths or amplitudes of vectors and 1-forms and as a way to convert between them. And we will construct the so-called 'covariant derivative' which is somewhat more complicated than in rectilinear coordinates because we need to account for the variation of the bases as well as the components.

We should stress at the outset that we still consider here only special relativity; i.e. flat rather than curved space-time. That is in order to keep things relatively straightforward at a conceptual level. But essentially all of the results and machinery carry over to curved manifolds as they are concerned with what happens *locally* – in the immediate vicinity of some point on the manifold where space-time is locally flat, with curvature appearing only at 2nd order in position (for instance in 2nd derivatives of basis vectors).

In fact, most of the salient features are not highly dependent on the locally Minkowskian geometry of 4D space-time and are equally well illustrated by considering ordinary 2D Euclidean space in curvilinear coordinates. Following Schutz, we will use 2D polar coordinates as an illustrative example.

Another difference with the previous section is that while there we were considering a family of coordinate systems related to one and other by boost and/or rotations, here we just consider two coordinate systems; one being the usual Cartesian system where the basis vectors are aligned with the coordinate axes and where components measure physical distances, and the other being a single curvilinear coordinate system.

3.1 Vectors and 1-forms in curvilinear coordinates

Consider a flat Euclidean 2-space – a sheet of paper – on which we may have points (like events in special relativity) and fields such as a grey-scale image. Over this we can lay a transparency on which are drawn lines of a Cartesian, or 'physical', (x, y) coordinate grid, from which we can assign physical coordinates to points, and we can also overlay a transparency covered with lines of constant curvilinear coordinates $\xi = \xi(x, y)$ and $\eta = \eta(x, y)$. We shall sometimes refer to these systems as 'frames', by analogy with reference frames in SR.

3.1.1 Displacement vectors

Consider two neighbouring points with relative physical displacement vector $(\Delta x, \Delta y)$. These have a corresponding displacement $(\Delta\xi, \Delta\eta)$ in curvilinear coordinates which, for small Δx and Δy, will be linear in $(\Delta x, \Delta y)$. Representing these displacements as column vectors, we have

$$\begin{bmatrix} \Delta\xi \\ \Delta\eta \end{bmatrix} = \begin{bmatrix} \partial\xi/\partial x & \partial\xi/\partial y \\ \partial\eta/\partial x & \partial\eta/\partial y \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{46}$$

Or, letting $\Delta x^\alpha = (\Delta x, \Delta y)$ and $\Delta x^{\alpha'} = (\Delta\xi, \Delta\eta)$

$$\boxed{\Delta x^{\alpha'} = \Lambda^{\alpha'}{}_\alpha \Delta x^\alpha} \tag{47}$$

where $\Lambda^{\alpha'}{}_\alpha$ represents the matrix above and where, as always, the first (second) index labels the rows (columns):

$$\boxed{\Lambda^{\alpha'}{}_\alpha = \partial x^{\alpha'}/\partial x_\alpha.} \tag{48}$$

We will assume that the inverse mapping is single valued and that if $\Delta\xi$ and $\Delta\eta$ are both zero then the same is true for Δx and Δy – this requires that the determinant of the matrix $\Lambda^{\alpha'}{}_\alpha$ be non-vanishing.

An entity $\vec{V}$ with two components $V^\alpha = (V^x, V^y)$ and $V^{\alpha'} = (V^\xi, V^\eta)$ is a vector if the components transform in the same way as does the displacement vector.

3.1.2 1-forms as gradients of scalar fields

Consider now a scalar function of position $\phi(x, y)$. We *define* the 1-form $\tilde{d}\phi$ to be the gradient, whose physical components are $(\partial_x\phi, \partial_y\phi)$, or

$$(\tilde{d}\phi)_\alpha = \phi_{,\alpha}. \tag{49}$$

The change in ϕ over the interval $\vec{\Delta x}$ is the contraction

$$\Delta\phi = \Delta x^\alpha \phi_{,\alpha}. \tag{50}$$

which we also denote by $\tilde{d}\phi(\vec{\Delta x})$.

We take the components, in the (ξ, η) coordinate system to be $(\partial_\xi\phi, \partial_\eta\phi)$, or $\tilde{d}\phi \underset{\xi,\eta}{\longrightarrow} (\tilde{d}\phi)_{\alpha'} = (\phi_{,\xi}, \phi_{,\eta})$, in terms of which

$$\Delta\phi = \Delta x^{\alpha'} \phi_{,\alpha'}. \tag{51}$$

This defines a set of 1-form fields, one for each possible scalar field, the transformation law for any one of which may be expressed as matrix multiplication of *row* vectors

$$[\phi_{,\xi}, \phi_{,\eta}] = [\phi_{,x}, \phi_{,y}] \begin{bmatrix} \partial x/\partial\xi & \partial x/\partial\eta \\ \partial y/\partial\xi & \partial y/\partial\eta \end{bmatrix} \tag{52}$$

or as

$$\boxed{\phi_{,\alpha'} = \Lambda^\alpha{}_{\alpha'} \phi_{,\alpha}} \tag{53}$$

with

$$\boxed{\Lambda^\alpha{}_{\alpha'} \equiv \partial x^\alpha/\partial x^{\alpha'}.} \tag{54}$$

As we saw previously, $\Lambda^{\alpha'}{}_\alpha = \partial x^{\alpha'}/\partial x^\alpha$ and $\Lambda^\alpha{}_{\alpha'} = \partial x^\alpha/\partial x^{\alpha'}$ are matrix inverses of one another. This ensures that the contraction $\Delta\phi = \tilde{d}\phi(\vec{\Delta x})$ is frame invariant.

3.1.3 Curves and path derivatives of scalars

We define a *curve* to be a parameterised path $\vec{x}(\lambda)$ which in the x, y system has components $x^\alpha(\lambda) = (x(\lambda), y(\lambda))$ and in the ξ, η system has components $x^{\alpha'}(\lambda) = (\xi(\lambda), \eta(\lambda))$. If the path is sufficiently smooth we can define the *tangent vector* to be $\vec{V} = d\vec{x}/d\lambda$. The *path derivative* of the field $\phi(\vec{x})$ along the curve is

$$d\phi/d\lambda = \phi_{,\alpha} dx^\alpha/d\lambda = \phi_{,\alpha'} dx^{\alpha'}/d\lambda \tag{55}$$

or

$$d\phi/d\lambda = \tilde{d}\phi(\vec{V}) = (\tilde{d}\phi)_\alpha (\vec{V})^\alpha \tag{56}$$

i.e. the (coordinate sytem independent) contraction of the components of $\tilde{d}\phi$ and $\vec{V}$.

3.1.4 The coordinate basis vectors

Quite generally we want to write a vector $\vec{V}$ as equal to a sum of basis vectors times components:

$$\boxed{\vec{V} = V^\alpha \vec{e}_\alpha} \tag{57}$$

and we want this to hold in an arbitrary frame. There is some freedom in how, exactly, to do this. What we will describe here is the use of *coordinate basis vectors.*

As illustrated in figure 2, we take the basis vectors $\vec{e}_x, \vec{e}_y$ in the (x, y) frame (the un-primed frame here) to be of unit length and aligned with the x and y axes. That means that the components of $\vec{\Delta x}$ in the (x, y) frame directly measure physical distances projected onto the axes. While it is somewhat tautological we can say that these basis vectors have components $\vec{e}_x \to (1, 0)$ and $\vec{e}_y \to (0, 1)$ or, equivalently,

$$\boxed{(\vec{e}_\alpha)^\beta = \delta^\beta_\alpha.} \tag{58}$$

These basis vectors are evidently orthonormal. If we were doing this in Minkowski space the basis vectors would be orthogonal, but the norm of the time basis vector would be negative.

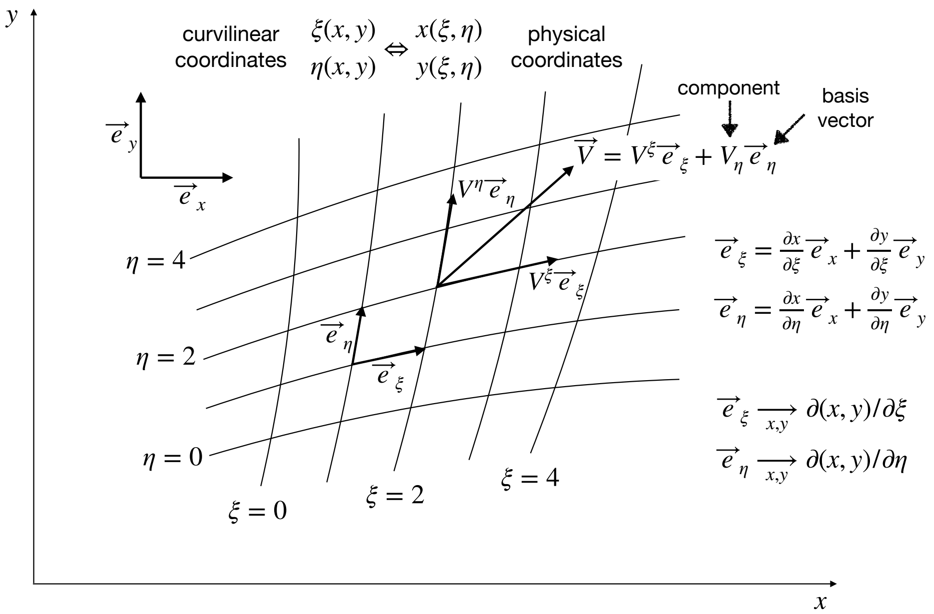

Figure 2: The curved lines are contours of constant curvilinear coordinates (ξ, η). In 3-dimensions they would be surfaces. An arbitrary vector like $\vec{V}$ can be expressed either as a sum of (ξ, η)-frame basis vectors $(\vec{e}_\xi, \vec{e}_\eta)$ times (ξ, η)-frame components, as indicated, or as a sum of the physical basis vectors $(\vec{e}_x, \vec{e}_y)$ times physical components. As shown in the text, the physical components of $\vec{e}_\xi$ are $\partial(x, y)/\partial\xi$ and similarly for $\vec{e}_\eta$.

Since $\vec{V} = V^\alpha \vec{e}_\alpha$ is coordinate system independent, while the components transform like $V^{\alpha'} = \Lambda^{\alpha'}{}_\alpha V^\alpha$, the basis vectors must transform as

$$\boxed{\vec{e}_{\alpha'} = \Lambda^\alpha{}_{\alpha'} \vec{e}_\alpha} \tag{59}$$

so the matrix that affects this transformation is the inverse of that involved in the transformation for the components, as, of course, it has to be if the combination $V^{\alpha'} \vec{e}_{\alpha'}$ be frame invariant.

This formula gives either of the pair of basis vectors $\vec{e}_{\alpha'} = (\vec{e}_\xi, \vec{e}_\eta)$ as linear combinations of the physical basis vectors $\vec{e}_\alpha = (\vec{e}_x, \vec{e}_y)$. As discussed previously, this is not telling us how components of a vector transform, it is telling us how to construct new vectors out of old ones.

One might ask: what are the components of these new basis vectors? The answer is again somewhat tautological. For any vector $\vec{V} = V^\xi \vec{e}_\xi + V^\eta \vec{e}_\eta$, so the components of $\vec{V} = \vec{e}_\xi$ are obviously $(1, 0)$ and those of $\vec{V} = \vec{e}_\eta$ are just $(1, 0)$ or

$$\boxed{(\vec{e}_{\alpha'})^{\beta'} = \delta^{\beta'}_{\alpha'}.} \tag{60}$$

These may look orthonormal, like the physical basis vectors, but they are not, because these are components in the (ξ, η) frame and the components of the metric tensor in this frame are not in general diag$(1,1)$ as they are in the (x, y) frame.

It is more interesting is to ask: what are the components of $\vec{e}_\xi$ and $\vec{e}_\eta$ in the (x, y) frame? I.e. what are the *physical* components of these basis vectors? The answer is

$$(\vec{e}_{\alpha'})^\beta = (\Lambda^\alpha{}_{\alpha'}\vec{e}_\alpha)^\beta = \Lambda^\alpha{}_{\alpha'}(\vec{e}_\alpha)^\beta = \Lambda^\alpha{}_{\alpha'}\delta^\beta_\alpha = \Lambda^\beta{}_{\alpha'} = \partial x^\beta/\partial x^{\alpha'}. \tag{61}$$

so

$$\vec{e}_\xi \underset{x,y}{\longrightarrow} \partial(x,y)/\partial\xi \qquad \text{and} \qquad \vec{e}_\eta \underset{x,y}{\longrightarrow} \partial(x,y)/\partial\eta \tag{62}$$

which means that $\vec{e}_\xi$ points along a line of constant η and $\vec{e}_\eta$ points along a line of constant ξ. In higher dimensional space, each basis vector points along the line where the other coordinates are constant.

3.1.5 Coordinate basis 1-forms

As mentioned, the prototype 1-form is the gradient of a scalar $\tilde{p} = \tilde{d}\phi \underset{x,y}{\longrightarrow} \partial_\alpha\phi$. We can write this as $\tilde{p} = p_x\tilde{\omega}^x + p_y\tilde{\omega}^y$ where $(\tilde{\omega}^x, \tilde{\omega}^y)$ are the (x, y)-frame basis 1-forms. Or

$$\boxed{\tilde{p} = p_\alpha\tilde{\omega}^\alpha.} \tag{63}$$

Just as we visualise $\tilde{d}\phi$ as a stack of iso-ϕ contours, we can visualise the basis 1-forms as stacks of iso-x and iso-y contours, as illustrated in figure 3. The *coordinate basis 1-forms* are $\tilde{\omega}^x = \tilde{d}x$ and similarly $\tilde{\omega}^y = \tilde{d}y$ or

$$\tilde{\omega}^\alpha = \tilde{d}x^\alpha \tag{64}$$

with components $\tilde{\omega}^x \to (1, 0)$ and $\tilde{\omega}^y \to (0, 1)$, or

$$(\tilde{\omega}^\alpha)_\beta = \delta^\alpha_\beta. \tag{65}$$

That's all in the physical (x, y)-frame. Invariance of the scalar $\tilde{d}\phi(\vec{V}) = (\tilde{d}\phi)_\alpha V^\alpha = \phi_{,\alpha}V^\alpha$ along with the transformation law for the components of a vector implies that the components of a 1-form must transform as

$$\boxed{p_{\alpha'} = \Lambda^\alpha{}_{\alpha'}p_\alpha} \tag{66}$$

so, as $\tilde{p}$ is invariant also, the basis 1-forms in the curvilinear frame must be

$$\boxed{\tilde{\omega}^{\alpha'} = \Lambda^{\alpha'}{}_\alpha\tilde{\omega}^\alpha = \partial x^{\alpha'}/\partial x_\alpha\tilde{\omega}^\alpha.} \tag{67}$$

The (x, y)-frame basis 1-forms are just $\{\tilde{\omega}^x, \tilde{\omega}^y\} = \{\tilde{d}x, \tilde{d}y\}$ so $\tilde{\omega}^\xi = \partial\xi/\partial x^\alpha\tilde{d}x^\alpha$, which is just $\tilde{d}\xi$, just the gradient of ξ considered as a function of x and y. Similarly $\tilde{\omega}^\eta = \tilde{d}\eta$.

The physical (i.e. (x, y)-frame) components of these basis 1-forms are

$$(\tilde{\omega}^{\alpha'})_\beta = (\Lambda^{\alpha'}{}_\alpha\tilde{\omega}^\alpha)_\beta = \Lambda^{\alpha'}{}_\alpha(\tilde{\omega}^\alpha)_\beta = \Lambda^{\alpha'}{}_\alpha\delta^\alpha_\beta = \Lambda^{\alpha'}{}_\beta = \partial x^{\alpha'}/\partial x^\beta \tag{68}$$

so

$$\tilde{\omega}^\xi \underset{x,y}{\longrightarrow} (\partial\xi/\partial x, \partial\xi/\partial y) \qquad \text{and} \qquad \tilde{\omega}^\eta \underset{x,y}{\longrightarrow} (\partial\eta/\partial x, \partial\eta/\partial y) \tag{69}$$

3.1.6 Visualising basis vectors and basis 1-forms:

One should keep in mind that while we talk about the (x, y) 'frame' and the (ξ, η) frame, there is only one *space* – a flat sheet of paper perhaps. On that sheet we can *overlay* a Cartesian grid of lines of constant x and y on a transparency and we can also overlay a warped grid of lines of constant ξ and η. We can think of ξ and η as functions of x and y and vice versa. On the sheet of paper we can draw physical dots (events in SR) and physical vectors like $\vec{V}$ as little arrows. These are real geometrical objects. The basis vectors $\vec{e}_\xi$ and $\vec{e}_\eta$ are two *fields* of such vectors; $\vec{e}_\xi(x, y)$ and $\vec{e}_\eta(x, y)$. The (x, y)-frame components of $\vec{e}_\xi(x, y)$ are $\partial(x, y)/\partial\xi$, and those of $\vec{e}_\eta(x, y)$ are $\partial(x, y)/\partial\eta$. The values of these vectors are determined by, and describe in a differential manner, the mapping of the coordinates $(x, y) \leftrightarrow (\xi, \eta)$, and they are labelled by the names of the coordinates ξ, η. But they do not 'live' in (ξ, η) space; they are real vectors that live

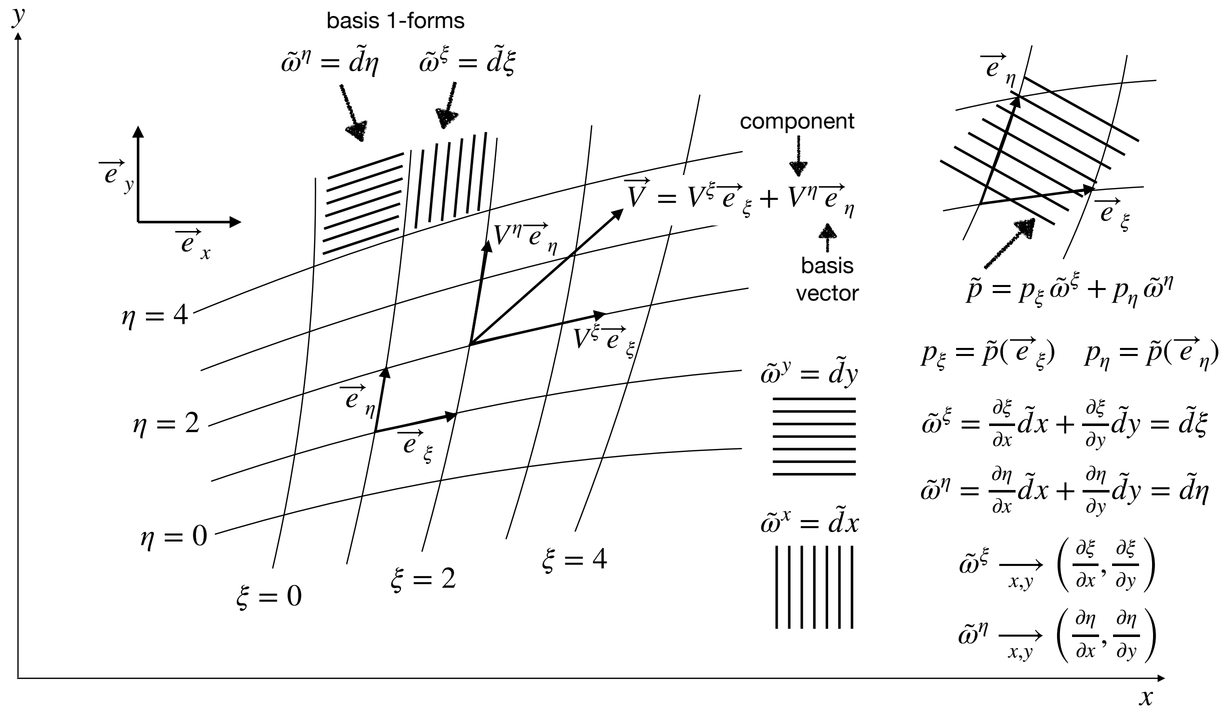

Figure 3: Basis 1-forms in curvilinear coordinates. Surfaces of constant curvilinear coordinates and basis vectors are as in figure 2. Also shown are basis 1-forms as stacks of iso-coordinate surfaces. At the upper right is shown a 1-form $\tilde{p} = \tilde{d}\phi$ and how it is decomposed, much like a vector, as a sum of bases times components $\tilde{p} = p_\xi \tilde{\omega}^\xi + p_\eta \tilde{\omega}^\eta$. The components of a 1-form are $p_{\alpha'} = \tilde{p}(\vec{e}_{\alpha'})$. I.e. the number of iso-ϕ contours pierced by the basis vector $\vec{e}_{\alpha'}$. Basis 1-forms are given by the operator $\tilde{d}$ acting on the appropriate coordinate. The physical components of the curvilinear basis 1-forms are $(\tilde{\omega}^{\alpha'})_\beta = \partial x^{\alpha'}/\partial x^\beta$.

on the sheet of paper. On the sheet we can also draw a 1-form $\tilde{p}$ as a little stack of equally spaced parallel lines whose magnitude we understand to be the inverse spacing of the lines. The basis 1-forms are two 1-form *fields* $\tilde{\omega}^\xi(x, y)$ and $\tilde{\omega}^\eta(x, y)$. The (x, y)-frame components of $\tilde{\omega}^\xi$ are $(\partial\xi/\partial x, \partial\xi/\partial y)$ and those of $\tilde{\omega}^\eta$ are $(\partial\eta/\partial x, \partial\eta/\partial y)$.

This is illustrated in figure 3 where the basis vector $\vec{e}_\xi$ points along the contour of constant η and reaches from one contour of ξ to the next one. Similarly for $\vec{e}_\eta$. And the basis 1-form $\tilde{\omega}^\xi$ is visualisable as a stack of (perhaps more finely spaced) contours of ξ and similarly for $\tilde{\omega}^\eta$.

Both the basis vectors and the basis 1-forms describe, in a differential manner, the mapping between the two coordinate systems:

$$\{\vec{e}_\alpha\} \underset{x,y}{\longrightarrow} \partial(x, y)/\partial(\xi, \eta) \qquad \text{and} \qquad \{\tilde{\omega}^\alpha\} \underset{x,y}{\longrightarrow} \partial(\xi, \eta)/\partial(x, y) \tag{70}$$

Their physical components are both partial derivatives of one set of coordinates with respect to the other and they therefore encode identical information.

But they are useful for different things. If we have a separation $\Delta r, \Delta\theta$ between two points in polar coordinates, for example, and form the vector $\Delta r \vec{e}_r + \Delta\theta \vec{e}_\theta$ then the (x, y) components of this are useful for computing things like the squared modulus of the separation $\Delta l^2 = \Delta x^2 + \Delta y^2$. The basis 1-forms are more useful if we are dealing with e.g. a tabulated scalar function whose differences on a grid of r, θ have been computed. If we have the physical components of the basis 1-forms tabulated also then the (x, y) components of $\tilde{d}\phi = \phi_{,r}\tilde{d}r + \phi_{,\theta}\tilde{d}\theta$ are the physical components of the gradient.

3.1.7 Orthogonality of the basis vectors and basis 1-forms

The basis vectors $\vec{e}_{\alpha'}$ associated with a curvilinear frame are, in general, neither normal nor orthogonal to one another. But the basis vectors and the basis 1-forms are orthogonal to each other in the sense that

$$\tilde{\omega}^{\alpha'}(\vec{e}_{\beta'}) = \delta^{\alpha'}{}_{\beta'} \tag{71}$$

I.e. the same as $\tilde{\omega}^\alpha(\vec{e}_\beta) = \delta^\alpha{}_\beta$ in the Cartesian frame.

This is *not* simply saying that, for any α and β, $\tilde{\omega}^{\alpha}(\vec{e}_{\beta})$ is a scalar and so is frame invariant, because $\tilde{\omega}^{\alpha'}(\vec{e}_{\beta'}) = \delta^{\alpha'}{}_{\beta'}$ is expressing a relation between members of *different* sets of basis vectors and basis 1-forms.

But it is valid nonetheless, since

$$\tilde{\omega}^{\alpha'}(\vec{e}_{\beta'}) = \Lambda^{\alpha'}{}_{\alpha}\tilde{\omega}^{\alpha}(\Lambda^{\beta}{}_{\beta'}\vec{e}_{\beta}) = \Lambda^{\alpha'}{}_{\alpha}\Lambda^{\beta}{}_{\beta'}\tilde{\omega}^{\alpha}(\vec{e}_{\beta}) = \Lambda^{\alpha'}{}_{\alpha}\Lambda^{\beta}{}_{\beta'}\delta^{\alpha}{}_{\beta} = \Lambda^{\alpha'}{}_{\mu}\Lambda^{\mu}{}_{\beta'} = \delta^{\alpha'}{}_{\beta'} \tag{72}$$

and is obvious in terms from figure 3.

3.2 The metric in curvilinear coordinates

3.2.1 The metric as a way to measure magnitudes of vectors

Just as in rectilinear coordinates, the metric $\mathbf{g}(\ ,\ )$ takes 2 vector arguments. Given the same vector twice, it gives the squared length:

$$\Delta l^2 = \mathbf{g}(\vec{\Delta x}, \vec{\Delta x}) = \vec{\Delta x} \cdot \vec{\Delta x} \tag{73}$$

and its components, in the frame with basis vectors $\{\vec{e}_{\alpha}\}$, are defined to be $g_{\alpha\beta} = \mathbf{g}(\vec{e}_{\alpha}, \vec{e}_{\beta}) = \vec{e}_{\alpha} \cdot \vec{e}_{\beta}$.

In (x, y) coordinates $\vec{e}_x \cdot \vec{e}_x = \vec{e}_y \cdot \vec{e}_y = 1$ and $\vec{e}_x \cdot \vec{e}_y = 0$, so

$$g_{\alpha\beta} = \mathrm{diag}(1, 1) = \delta_{\alpha\beta}. \tag{74}$$

We may also write $\mathbf{g}$ in terms of the outer product bases: $g_{\alpha\beta}\tilde{\omega}^{\alpha} \otimes \tilde{\omega}^{\beta} = g_{\alpha\beta}\tilde{dx}^{\alpha} \otimes \tilde{dx}^{\beta}$, so

$$\boxed{\mathbf{g} = \tilde{dx} \otimes \tilde{dx} + \tilde{dy} \otimes \tilde{dy}.} \tag{75}$$

Similarly, in (r, θ) coordinates $g_{\alpha'\beta'} = \vec{e}_{\alpha'} \cdot \vec{e}_{\beta'}$, which, with $\vec{e}_r = \cos\theta\vec{e}_x + \sin\theta\vec{e}_y$ and $\vec{e}_{\theta} = r(-\sin\theta\vec{e}_x + \cos\theta\vec{e}_y)$ (see figure 4 and appendix A), gives $g_{\alpha'\beta'} = \mathrm{diag}(1, r^2)$. So we can express $\mathbf{g}$ as a sum of outer products of pairs of (r, θ)-frame basis 1-forms as $\mathbf{g} = g_{\alpha'\beta'}\tilde{\omega}^{\alpha'}\tilde{\omega}^{\beta'}$ or

$$\boxed{\mathbf{g} = \tilde{dr} \otimes \tilde{dr} + r^2\tilde{d\theta} \otimes \tilde{d\theta}.} \tag{76}$$

The length dl^2 of an infinitesimal vector $dr\vec{e}_r + d\theta\vec{e}_{\theta}$ is

$$\boxed{dl^2 = dr^2 + r^2 d\theta^2} \tag{77}$$

which is often called the *'line element'*, and from which we can read off the components of the metric.

The expressions (76) and (77) for $\mathbf{g}$ and for dl^2 above look similar but are somewhat different things. The expression for $\mathbf{g}$ defines an *operator* or *function* to which we feed vectors to get their lengths and scalar products. And $\tilde{dr}$ and $\tilde{d\theta}$ are not in any meaningful sense small. Whereas dr and $d\theta$ are infinitesimals.

Though if we feed $\mathbf{g}$ two copies of $dr\vec{e}_r + d\theta\vec{e}_{\theta}$, then we get $dl^2 = \ldots$. And we could also feed $\mathbf{g}$ two *different* vectors – in SR, for example, the four-momentum of a photon $\vec{p}$ and the 4-velocity of an observer $\vec{U}$ – in which case $\mathbf{g}(\vec{p}, \vec{U}) = \tilde{p}(\vec{U})$ gives (minus) the energy of the photon that the observer would measure.

But both (76) and (77) are equally good for reading off the components of the metric and indicating at the same time the coordinates being used.

The (x, y)-frame expression (75), while equivalent to (76), reveals something obvious but which will be crucial later: The metric is a sum of outer product bases – which are composed from position independent 1-form bases $\tilde{\omega}^{\alpha} = \tilde{dx}^{\alpha}$ – with coefficients that are also constant. Thus the metric $\mathbf{g}(\ ,\ )$ is independent of position. This is obvious as we are just dealing with a sheet of paper which is flat. But it might not be obvious – while true – if we were given (76).

3.2.2 The metric as a mapping between vectors and 1-forms

The other thing we have used the metric for is mapping vectors to 1-forms and vice versa. We noted that $\mathbf{g}(\vec{V},\ )$ is a 1-form and $\mathbf{g}^{-1}(\tilde{p},\ )$ is a vector since filling the empty slots with a vector and 1-form, respectively, give scalars.

We can write $\tilde{V} = \mathbf{g}(\vec{V},\ )$ and $\vec{V} = \mathbf{g}^{-1}(\tilde{V},\ )$, so $\mathbf{g}$ and $\mathbf{g}^{-1}$ are respectively the index lowering and raising operators.

In component form $V_{\alpha'} = g_{\alpha'\beta'}V^{\beta'}$, which implies $V^{\alpha'} = (g_{\alpha'\beta'})^{-1}V_{\beta'}$, so the components of the 'index raising' operator $\mathbf{g}^{-1}$ is the matrix inverse of $g_{\alpha'\beta'}$.

For example, in polar coordinates, the inverse exists and is:

$$g^{\alpha'\beta'} = (g_{\alpha',\beta'})^{-1} = \mathrm{diag}(1, 1/r^2). \tag{78}$$

Note however that here, unlike in flat Euclidean space (or flat Minkowski space), the raising and lowering matrices are not the same

We noted that it was *natural* to think of the 4-momentum of a particle as a 1-form $\tilde{p}$, because of the close relationship between the conceptual picture of the 4-momentum 1-form and the wave-crests of the matter- (or radiation-) wave packet that represents such an object. But you are free to use the corresponding, and equivalent, vector $\vec{p}$, though it is important to keep track of the minus sign in $p_0 = -p^0$ (in a rectilinear frame)

3.2.3 The metric in terms of transformation matrices

The components of the metric are $g_{\alpha'\beta'} = \mathbf{g}(\vec{e}_{\alpha'}, \vec{e}_{\beta'}) = \vec{e}_{\alpha'} \cdot \vec{e}_{\beta'}$ so

$$g_{\alpha'\beta'} = \Lambda^{\alpha}{}_{\alpha'}\Lambda^{\beta}{}_{\beta'}\vec{e}_{\alpha} \cdot \vec{e}_{\beta} = \delta_{\alpha\beta}\Lambda^{\alpha}{}_{\alpha'}\Lambda^{\beta}{}_{\beta'} = \delta_{\alpha\beta}\frac{\partial x^{\alpha}}{\partial x^{\alpha'}}\frac{\partial x^{\beta}}{\partial x^{\beta'}}. \tag{79}$$

3.3 The covariant derivative

We will now construct the *covariant derivative* operator, which differs from that in rectilinear coordinates as the basis vectors and basis 1-forms are position dependent.

First we will introduce the *connection* or Christoffel symbols that give the components of the derivative of the basis vectors.

We then consider the derivative of a vector and later that of tensors.

3.3.1 The derivative of the basis vectors: the connection $\Gamma^{\mu}{}_{\alpha\nu}$

In the (x, y) frame – which will be the primed frame here – the basis vectors are position independent, so $\partial\vec{e}_{\alpha'}/\partial x^{\beta'} = 0$. While in the curvilinear frame $\partial\vec{e}_{\alpha}/\partial x^{\beta} \neq 0$.

As an example, consider polar coordinates, in which (see figure 4 and appendix A)

$$\begin{aligned} \vec{e}_r &= \cos\theta\vec{e}_x + \sin\theta\vec{e}_y \\ \vec{e}_\theta &= r(-\sin\theta\vec{e}_x + \cos\theta\vec{e}_y) \end{aligned} \tag{80}$$

As the basis vectors $\vec{e}_x$ and $\vec{e}_x$ are constant, the derivatives of $\vec{e}_r$ are

$$\begin{aligned} \partial\vec{e}_r/\partial r &= 0 \\ \partial\vec{e}_r/\partial\theta &= -\sin\theta\vec{e}_x + \cos\theta\vec{e}_y = \vec{e}_\theta/r \end{aligned} \tag{81}$$

and the derivatives of $\vec{e}_\theta$ are readily found to be

$$\begin{aligned} \partial\vec{e}_\theta/\partial r &= \vec{e}_\theta/r \\ \partial\vec{e}_\theta/\partial\theta &= -r\vec{e}_r. \end{aligned} \tag{82}$$

The derivative of the basis vector $\partial\vec{e}_{\alpha}/\partial x^{\beta}$ is *not*, in fact, a tensor. It vanishes in the (x, y)-frame. But if the components of a tensor vanish in one frame they vanish in all frames, which is not the case.

Nonetheless, we can write the derivative of $\vec{e}_{\alpha}$ along the β^{th} coordinate direction, as a linear combination of the basis vectors:

$$\boxed{\partial\vec{e}_{\alpha}/\partial x^{\beta} = \Gamma^{\mu}{}_{\alpha\beta}\vec{e}_{\mu}} \tag{83}$$

where the coefficients $\Gamma^{\mu}{}_{\alpha\beta}$ are called the *Christoffel symbols* and are the *components* of $\partial\vec{e}_{\alpha}/\partial x^{\beta}$.

They are also called the *connection coefficients* or just the *connection*, and are sometimes also denoted by $\left\{ {\mu \atop \alpha\beta} \right\}$.

The connection is related to the coordinate transformation matrix as follows: The curvilinear bases are

$$\vec{e}_{\alpha} = (\partial x^{\alpha'}/\partial x^{\alpha})\vec{e}_{\alpha'} \tag{84}$$

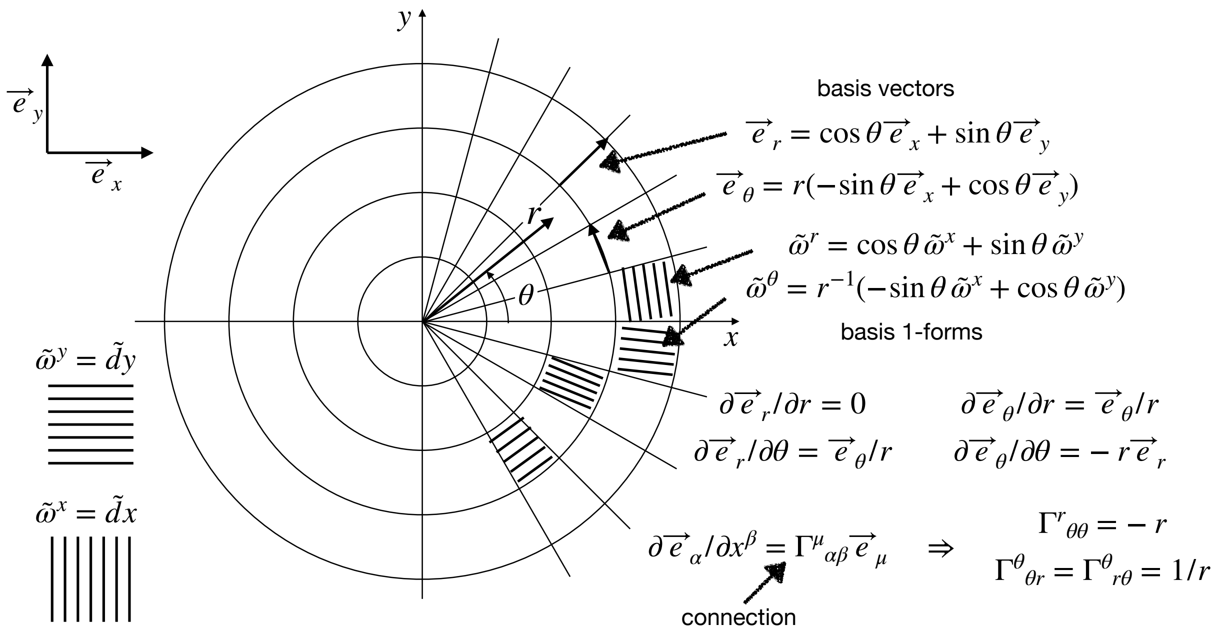

Figure 4: Basis vectors in polar coordinates are indicated by the arrows. The (x, y)-frame basis vectors $\vec{e}_x$ and $\vec{e}_y$ are just unit vectors pointing along the x and y coordinate axes. The basis 1-forms $\tilde{\omega}^r$ and $\tilde{\omega}^\theta$ are also indicated as stacks of iso-r and iso-θ surfaces. All of the basis vectors and basis 1-forms rotate if we change θ. The magitudes of $\vec{e}_r$ and $\tilde{\omega}^r$ are constant while those of $\vec{e}_\theta$ and $\tilde{\omega}^\theta$ vary linearly and inversely with r. The partial derivatives of the basis vectors are shown, along with the formula defining the connection $\Gamma^\mu{}_{\alpha\beta}$, and at bottom right, the non-vanishing components of the connection are shown.

so $\partial\vec{e}_\alpha/\partial x^\beta = (\partial^2 x^{\alpha'}/\partial x^\beta \partial x^\alpha)\vec{e}_{\alpha'}$ as the $\vec{e}_{\alpha'}$ are constant. So

$$\partial\vec{e}_\alpha/\partial x^\beta = (\partial^2 x^{\alpha'}/\partial x^\beta \partial x^\alpha)\partial x^\mu/\partial x^{\alpha'}\vec{e}_\mu \tag{85}$$

Comparing with the definition of the Christoffel symbols above we have

$$\boxed{\Gamma^\mu{}_{\alpha\beta} = \frac{\partial x^\mu}{\partial x^{\alpha'}}\frac{\partial^2 x^{\alpha'}}{\partial x^\beta \partial x^\alpha}} \tag{86}$$

which is useful as it shows that $\Gamma^\mu{}_{\alpha\beta}$ is symmetric under exchange of its lower indices $\alpha \Leftrightarrow \beta$.

This shows that $\Gamma^\mu{}_{\alpha\beta}$ is the derivative with respect to x^β of the matrix $\partial x^{\alpha'}/\partial x^\alpha$ multiplied by its inverse, so $\Gamma^\mu{}_{\alpha\beta} = \partial_\beta \log\left(\partial x^{\alpha'}/\partial x^\alpha\right)$.

Finally, we can also write the connection in terms of the Λ-matrices as

$$\boxed{\Gamma^\mu{}_{\alpha\beta} = \Lambda^\mu{}_{\alpha'}\Lambda^{\alpha'}{}_{\alpha,\beta}.} \tag{87}$$

3.3.2 The covariant derivative of a vector

Consider a vector field $\vec{V}(\vec{x}) = V^\alpha(\vec{x})\vec{e}_\alpha(\vec{x})$ where, as before, the un-primed frame is the curvilinear frame, and take the difference between its values at two neighbouring points:

$$\Delta\vec{V} = V^\alpha(\vec{x} + \Delta\vec{x})\vec{e}_\alpha(\vec{x} + \Delta\vec{x}) - V^\alpha(\vec{x})\vec{e}_\alpha(\vec{x}). \tag{88}$$

Taking the limit of $\Delta\vec{V}/\Delta x^\beta$ for a separation along the β^{th} coordinate axis we have

$$\frac{\partial\vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\vec{e}_\alpha + V^\alpha\frac{\partial\vec{e}_\alpha}{\partial x^\beta} \tag{89}$$

or, invoking the definition of $\Gamma^\mu{}_{\alpha\beta}$,

$$\frac{\partial\vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\vec{e}_\alpha + V^\alpha\Gamma^\mu{}_{\alpha\beta}\vec{e}_\mu \tag{90}$$

or, exchanging the dummy indices $\alpha \Leftrightarrow \mu$ in the last term,

$$\frac{\partial \vec{V}}{\partial x^\beta} = \left(\frac{\partial V^\alpha}{\partial x^\beta} + V^\mu \Gamma^\alpha{}_{\mu\beta} \right) \vec{e}_\alpha. \tag{91}$$

Switching notation $\partial X / \partial x^\alpha \Rightarrow X_{,\alpha}$, this says $\vec{V}_{,\beta}$, i.e. *the rate at which the entire vector – bases <u>and</u> components – changes wrt x^β, has components* $V^\alpha{}_{,\beta} + V^\mu \Gamma^\alpha{}_{\mu\beta}$.

We call this the *covariant derivative*, and we denote its components by

$$\boxed{V^\alpha{}_{;\beta} \equiv V^\alpha{}_{,\beta} + V^\mu \Gamma^\alpha{}_{\mu\beta}} \tag{92}$$

or, denoting the entity whose components are $V^\alpha{}_{;\beta}$ by $\nabla \vec{V}$, we have

$$\boxed{\nabla \vec{V} = V^\alpha{}_{;\beta} \vec{e}_\alpha \otimes \tilde{\omega}^\beta.} \tag{93}$$

Now $\nabla \vec{V} \to V^\alpha{}_{;\beta}$, despite involving the non-tensorial connection, is a tensor (see appendix B). That means that $V^\alpha{}_{,\beta}$ is not a tensor.

A variety of notation is used for covariant derivatives of vectors. We have $(\nabla \vec{V})^\alpha{}_\beta = V^\alpha{}_{;\beta}$, which is also denoted by $(\nabla_\beta \vec{V})^\alpha$. And we denote $(\nabla \vec{V})(\vec{U})$ – the path derivative of $\vec{V}$ along the path with tangent vector $\vec{U}$ – by $\nabla_{\vec{U}} \vec{V}$, which is the vector with components $V^\alpha{}_{;\beta} U^\beta$.

Here is a question to test your understanding of the covariant derivative: Given a vector field $\vec{V}(\vec{x})$, in rectilinear coordinates we can obtain, to first order in the displacement, the components of the vector at a displaced location $\vec{x} + \vec{\Delta x}$ as $V^\alpha(\vec{x}) + V^\alpha{}_{,\beta} \Delta x^\beta$. What about the analogous quantity in curvilinear coordinates: $V^\alpha(\vec{x}) + V^\alpha{}_{;\beta} \Delta x^\beta$? Does this provide the components of $\vec{V}(\vec{x} + \Delta \vec{x})$? If not, why not? What *does* it give us.

3.3.3 Parallel transport of a tangent vector along a curve

One use of the covariant derivative is to provide the variation of the components of the tangent vector $\vec{U} = d\vec{x}/d\lambda$ to a parameterised curve $\vec{x}(\lambda)$ in the case that the tangent vector is unchanging:

$$\nabla_{\vec{U}} \vec{U} = 0. \tag{94}$$

An important application of this is a free-particle, whose 4-momentum $\vec{p} = m\vec{U}$ is unchanging as it moves. We say that such a particle *'parallel transports'* its 4-momentum.

It is worth mentioning, at this point, that we done something potentially questionable in that we have here the operator that gives the derivative of a vector *field* and we have applied it to a tangent vector that is defined only along a *line.* If this worries you, you might want to think about the field $\vec{U}(\vec{x})$ that gives the 4-velocity of a collection of particle à la Hamilton and Jacobi.

Since $\nabla_{\vec{U}} \vec{U} \to U^\beta U^\alpha{}_{;\beta}$ we have, from (92), for $dU^\alpha / d\lambda = U^\beta U^\alpha{}_{,\beta}$

$$\boxed{dU^\alpha / d\lambda = -\Gamma^\alpha{}_{\mu\beta} U^\mu U^\beta} \tag{95}$$

or equivalently

$$\boxed{d^2 x^\alpha / d\lambda^2 = -\Gamma^\alpha{}_{\mu\beta} (dx^\mu / d\lambda)(dx^\beta / d\lambda)} \tag{96}$$

which is called the *geodesic equation*. It is the equation obeyed by the curvilinear coordinates $\vec{x}(\lambda)$ of a straight line.

Given the initial location and tangent vector (or 4-velocity) and knowledge of the connection we can integrate this to give the path and tangent or 4-momentum vector along the path. The parameter λ, in the case of a particle, being proper time.

As an example, consider 2D polar coordinates, for which the non-vanishing connection coefficients are

$$\begin{gathered} \Gamma^r{}_{\theta\theta} = -r \\ \Gamma^\theta{}_{\theta r} = \Gamma^\theta{}_{r\theta} = 1/r \end{gathered} \tag{97}$$

(see figure 4). The geodesic equations are

$$\begin{aligned}\ddot{r} &= -\Gamma^{r}{}_{\theta\theta}\dot{\theta}^2 = r\dot{\theta}^2 \\ \ddot{\theta} &= -\Gamma^{\theta}{}_{\theta r}\dot{\theta}\dot{r} - \Gamma^{\theta}{}_{r\theta}\dot{r}\dot{\theta} = -2\dot{\theta}\dot{r}/r\end{aligned} \tag{98}$$

The latter implies that $d(r^2\dot{\theta})/d\lambda = r^2\ddot{\theta} + 2r\dot{r}\dot{\theta} = 0$. That is sensible as $r^2\dot{\theta}$ is the angular momentum of a particle. This is also the θ component of the 1-form $\tilde{U} = \mathbf{g}(\vec{U},\ \ )$, and its conservation, as we will see later, follows from the invariance of the metric components with respect to θ. This is another reason why it is nice to think about the momentum as a 1-form as the physical content of conservation laws are often explicit, while somewhat masked in terms of the vector components.

3.3.4 Covariant derivatives of scalars, 1-forms and other tensors

– Covariant derivative of a scalar

A scalar ϕ does not depend on any basis vectors, so its covariant derivative is just the usual partial derivative, $\nabla_\beta \phi = \phi_{,\beta}$, hence

$$\boxed{\phi_{;\beta} = \phi_{,\beta}} \tag{99}$$

– Covariant derivative of a 1-form

One can use the same line of argument as used in the previous section to compute the covariant derivative of a 1-form (denoted in component form by $p_{\alpha;\beta}$).

Alternatively, one can use the fact that the scalar product $\phi = \tilde{p}(\vec{q}) = p_\alpha q^\alpha$ is also, of course, a scalar, and being the sum over α of products we have

$$\phi_{,\beta} = p_\alpha q^\alpha{}_{,\beta} + p_{\alpha,\beta} q^\alpha \tag{100}$$

On the other hand, from its definition as a limit of a difference, the covariant derivative must obey the analogous rule:

$$\phi_{;\beta} = p_\alpha q^\alpha{}_{;\beta} + p_{\alpha;\beta} q^\alpha \tag{101}$$

but $\phi_{;\beta} = \phi_{,\beta}$ and $q^\alpha{}_{;\beta} = q^\alpha{}_{,\beta} + q^\mu \Gamma^\alpha{}_{\mu\beta}$ so $p_{\alpha;\beta} q^\alpha = p_{\alpha,\beta} q^\alpha - p_\alpha q^\mu \Gamma^\alpha{}_{\mu\beta}$. Swapping the dummy indices $\alpha \leftrightarrow \mu$ in the last term gives $p_{\alpha;\beta} q^\alpha = (p_{\alpha,\beta} - p_\mu \Gamma^\mu{}_{\alpha\beta}) q^\alpha$ which, q^α being arbitrary, gives us the components of the covariant derivative $\nabla\tilde{p}$:

$$\boxed{p_{\alpha;\beta} \equiv p_{\alpha,\beta} - p_\mu \Gamma^\mu{}_{\alpha\beta}} \tag{102}$$

and, just as with $\nabla\vec{V}$, we can also denote these components as $(\nabla_\beta \tilde{p})_\alpha$ or $(\nabla\tilde{p})_{\alpha\beta}$.

One must pay attention to the placing of the indices and parentheses since, for example, $\nabla_\beta p_\alpha$ is not the same as $(\nabla_\beta \tilde{p})_\alpha$. Do you see why?

– Covariant derivative of a tensor

The argument above can be readily generalised to tensors of arbitrary rank and we find, for instance,

$$\boxed{\nabla_\beta B^\mu{}_\nu = B^\mu{}_{\nu;\beta} = B^\mu{}_{\nu,\beta} + B^\alpha{}_\nu \Gamma^\mu{}_{\alpha\beta} - B^\mu{}_\alpha \Gamma^\alpha{}_{\nu\beta}.} \tag{103}$$

So for each index there is a term involving a Christoffel symbol. For a covariant (downstairs) index we contract with the upstairs index on $\Gamma^\mu{}_{\alpha\beta}$ and for a contravariant index we contract with a downstairs index, and the other indices go where they have to go given that we always place the index denoting the coordinate with respect to which we are differentiation (β here) last

This is a convenient place to remind ourselves that since the operator $U^\beta \nabla_\beta$ is the same as $d/d\lambda$, which can defined in the usual way as the limit as $\Delta\lambda \to 0$ of the difference of the argument at two points on the path, all the usual rules of differentiation apply to the ∇ operator – so, for example, $\nabla(AB) = A\nabla B + B\nabla A$ whatever objects A and B might be and no matter how they may be contracted with one another

3.3.5 Computing the Christoffel symbols from the metric

The covariant derivate is the basis of Einstein's *'generalised covariance'*. If we have a law of physics written in relativistically covariant form, such as Maxwell's equations, for example, then we can obtain the form of the equation in an arbitrary coordinate system simply by replacing commas by semicolons. Another example already discussed is the law of motion for a free particle: it maintains constant 4-momentum so the 4-velocity obeys the *geodesic equation* (96).

But to apply this we need the connection. In the foregoing, we obtained this in terms of the matrix $\Lambda^{\alpha}{}_{\alpha'}$ that affects the transformation from the rectilinear to curvilinear coordinates: $\Gamma^{\mu}{}_{\alpha\beta} = \Lambda^{\mu}{}_{\alpha'}\Lambda^{\alpha'}{}_{\alpha,\beta}$.

But recall that the components of the metric are $g_{\alpha\beta} = \Lambda^{\alpha'}{}_{\alpha}\Lambda^{\beta'}{}_{\beta}\eta_{\alpha'\beta'}$, which, if we were to differentiate it, would give us something very similar. This suggests that we may be able to find an explicit expression for the connection directly in terms of the metric in the curvilinear frame. If so, we would not need to bother about the rectilinear coordinate system at all. That would be very nice. The metric is something that, in principle, is directly measurable from distances between events with known curvilinear coordinates. And we also obtain it, as we shall see, as the solution of Einstein's equations.

The desired relation comes, with a little algebra, from the fact that, as we saw, the metric, which we can express as $\mathbf{g} = \eta_{\alpha'\beta'}\tilde{d}x^{\alpha'} \otimes \tilde{d}x^{\beta'}$ is independent of position. That might appear to be involve the rectilinear (primed) coordinates, but it implies that

$$\boxed{\nabla \mathbf{g} = 0} \tag{104}$$

which is a tensor equation and is therefore true in an arbitrary frame.

Applying the rule for covariantly differentiating a tensor this says

$$g_{\alpha\beta,\mu} - \Gamma^{\nu}{}_{\alpha\mu}g_{\nu\beta} - \Gamma^{\nu}{}_{\beta\mu}g_{\alpha\nu} = 0. \tag{105}$$

To solve this equation for the connection, we write down the possible permutations of this (there being 3 as $g_{\alpha\beta}$ is symmetric)

$$\begin{aligned} g_{\alpha\beta,\mu} &= \Gamma^{\nu}{}_{\alpha\mu}g_{\nu\beta} + \Gamma^{\nu}{}_{\beta\mu}g_{\alpha\nu} \\ g_{\alpha\mu,\beta} &= \Gamma^{\nu}{}_{\alpha\beta}g_{\nu\mu} + \Gamma^{\nu}{}_{\mu\beta}g_{\alpha\nu} \\ -g_{\beta\mu,\alpha} &= -\Gamma^{\nu}{}_{\beta\alpha}g_{\nu\mu} - \Gamma^{\nu}{}_{\mu\alpha}g_{\beta\nu} \end{aligned} \tag{106}$$

where we have changed the sign in the last. Adding these, and using the symmetry of the metric and of the Christoffel symbols, we find that 4 of the 6 terms cancel, while the other two are equal, and this gives

$$g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha} = 2g_{\alpha\nu}\Gamma^{\nu}{}_{\beta\mu} \tag{107}$$

and dividing by 2 and multiplying by $g^{\alpha\gamma}$ gives the final result

$$\boxed{\Gamma^{\gamma}{}_{\beta\mu} = \tfrac{1}{2}g^{\alpha\gamma}(g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha}).} \tag{108}$$

To illustrate the utility of this, recall that Einstein's rocket and tower thought experiments suggest that clocks down in a potential well run slow compared to reference clocks at infinity $d\tau = 1 + \phi(\mathbf{x})/c^2)dt$. So the metric, in coordinates tied to steadily accelerated observers, is

$$g_{\alpha\beta} \simeq \mathrm{diag}(-(1 + 2\phi(\mathbf{x})/c^2), 1, 1, 1) \tag{109}$$

where $\phi(\mathbf{x})$ is the gravitational potential. The only coordinate dependent metric component is g_{00} (the time-time component) and that only depends on spatial position $\mathbf{x}$.

The geodesic equation is $d^2x^{\alpha}/d\tau^2 = -\Gamma^{\alpha}{}_{\mu\nu}(dx^{\mu}/d\tau)(dx^{\nu}/d\tau)$. We can use this to compute the i^{th} component of the acceleration of a particle by setting $\alpha = i$. But if $\phi \ll c^2$ – an excellent approximation for the Earth, for instance – and work to 1st order accuracy in ϕ/c^2 we can take $g^{\alpha\delta} = \eta^{\alpha\delta}$ above and the only non-vanishing Christoffel symbols appearing are $\Gamma^{i}{}_{00} = \phi_{,i}/c^2$. For a non-relativistic particle $\vec{dx}/d\tau \longrightarrow (c, \mathbf{o})$ and we obtain

$$d^2\mathbf{x}/d\tau^2 = -\boldsymbol{\nabla}\phi \tag{110}$$

which agrees with the Newtonian result.

3.3.6 Parallel transport of the momentum 1-form

Just as we obtained the geodesic equation telling us how the curvilinear components of the tangent vector (or equivalently momentum 4-vector) evolve for a freely moving particle (one for which $\nabla_{\vec{U}}\vec{p} = 0$) one can obtain the equation for the evolution of the components of the momentum 1-form for such a particle. In a rectilinear frame the components of the momentum 1-form are constant, so $U\beta p_{\alpha,\beta} = 0$. But the connection vanishes in the rectilinear frame so that is equivalent to the 1-form equation $\nabla_{\vec{U}}\tilde{p} = 0$. From this, and (102) we obtain

$$dp_\alpha/d\lambda = U^\beta p_{\alpha,\beta} = U^\beta \Gamma^\gamma{}_{\alpha\beta} p_\gamma. \tag{111}$$

Using (108) and $\tilde{p} = m\tilde{U}$ this says

$$dp_\alpha/d\lambda = \tfrac{1}{2} m U^\beta U_\gamma g^{\gamma\nu}(g_{\nu\alpha,\beta} + g_{\nu\beta,\alpha} - g_{\beta\alpha,\nu}) = \tfrac{1}{2} m U^\beta U^\nu (g_{\nu\alpha,\beta} + g_{\nu\beta,\alpha} - g_{\beta\alpha,\nu}). \tag{112}$$

But the 1st and last terms in the parentheses combined are anti-symmetric under $\beta \Leftrightarrow \nu$ and so vanish when contracted with the symmetric $U^\beta U^\nu$ and we are left with

$$\boxed{dp_\alpha/d\lambda = \tfrac{1}{2} m U^\beta U^\nu g_{\nu\beta,\alpha}} \tag{113}$$

from which we see that if the components of the metric are independent of one of the coordinates then the corresponding component of the momentum 1-form is a constant of the motion. This proves to be very useful.

3.4 Concluding comments

In this lecture we have developed the mathematical machinery of flat-space (or flat-space time) tensor calculus in arbitrary coordinate systems.

To recapitulate, we started with coordinate transformation $x^\alpha \Leftrightarrow x^{\alpha'}$. We then obtained the basis vectors and 1-forms as partial derivatives of x^α with respect to $x^{\alpha'}$ and vice versa. This allowed us to construct vectors and tensors (including 1-forms) and fields of such entities. We then constructed the 'covariant derivative' that gives the derivatives of vectors and tensors, properly taking into account the way the basis vectors vary with position. At the heart of the covariant derivative operator is the 'connection' $\Gamma^\mu{}_{\alpha\beta}$ which, we showed, can be computed explicitly from the metric in the curvilinear coordinate system.

This allows us to 'do' special relativity in arbitrary curvilinear coordinates and to translate the laws of physics from their form in the more restricted rectilinear Minkowski coordinates to obtain their 'generally covariant' form.

But the results we have obtained are of much greater generality. In Einstein's general relativity, space-time is not flat; the Universe is a curved 4-dimensional 'manifold'. But it is *locally* flat, in the Minkowskian sense, with a light-cone structure built into it and everything we have done here still applies. So if we have somehow obtained, perhaps by direct measurement, or have hypothesised, the metric for this manifold (expressed in some coordinate system) we can use (108) to get the connection and then, for example, solve (96) for trajectories of particles (e.g. planets in the solar system). Or we can solve things like Maxwell's equation or the Klein-Gordon equation to obtain the behaviour of EM radiation or scalar fields in this curved space-time.

A An example: Polar coordinates

A simple concrete example that illustrates many of the concepts developed here is the mapping on the plane from Cartesian to polar coordinates, as illustrated in figure 4.

The mappings between Cartesian $x^\alpha = (x, y)$ and polar $x^{\alpha'} = (r, \theta)$ coordinates are

$$\begin{aligned} (x(r,\theta), y(r,\theta)) &= r(\cos\theta, \sin\theta) \\ (r(x,y), \theta(x,y)) &= (\sqrt{x^2+y^2}, \tan^{-1}(y/x)) \end{aligned} \tag{114}$$

$$\begin{aligned} \Lambda^\alpha{}_{\alpha'} &= \frac{\partial x^\alpha}{\partial x^{\alpha'}} = \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{bmatrix} \partial x/\partial r & \partial x/\partial\theta \\ \partial y/\partial r & \partial y/\partial\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = \begin{bmatrix} x/r & -y \\ y/r & x \end{bmatrix} \\ \Lambda^{\alpha'}{}_\alpha &= \frac{\partial x^{\alpha'}}{\partial x^\alpha} = \frac{\partial(r,\theta)}{\partial(x,y)} = \begin{bmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial\theta/\partial x & \partial\theta/\partial y \end{bmatrix} = \begin{bmatrix} x/r & y/r \\ -y/r^2 & x/r^2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -(\sin\theta)/r & (\cos\theta)/r \end{bmatrix} \end{aligned} \tag{115}$$

$$\vec{e}_{\alpha'} = \Lambda^{\alpha}{}_{\alpha'}\vec{e}_{\alpha} \to [\vec{e}_r \ \vec{e}_\theta] = {}^{[\vec{e}_x \ \vec{e}_y]}\begin{bmatrix}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{bmatrix} \tag{116}$$

or

$$\begin{aligned}\vec{e}_r &= \cos\theta\vec{e}_x + \sin\theta\vec{e}_y \\ \vec{e}_\theta &= r(-\sin\theta\vec{e}_x + \cos\theta\vec{e}_y)\end{aligned} \tag{117}$$

and that the basis 1-forms are

$$\tilde{\omega}^{\alpha'} = \Lambda^{\alpha'}{}_{\alpha}\tilde{\omega}^{\alpha} = \Lambda^{\alpha'}{}_{\alpha}\tilde{d}x^{\alpha} \to \begin{bmatrix}\tilde{\omega}^r \\ \tilde{\omega}^\theta\end{bmatrix} = \begin{bmatrix}\cos\theta & \sin\theta \\ -(\sin\theta)/r & (\cos\theta)/r\end{bmatrix}\begin{bmatrix}\tilde{d}x \\ \tilde{d}y\end{bmatrix} \tag{118}$$

or

$$\begin{aligned}\tilde{\omega}^r &= \cos\theta\tilde{d}x + \sin\theta\tilde{d}y \\ \tilde{\omega}^\theta &= (-\sin\theta\tilde{d}x + \cos\theta\tilde{d}y)/r\end{aligned} \tag{119}$$

In trying to visualise the meaning of these equations, it may be useful to keep in mind that $(\cos\theta, \sin\theta)$ is a unit vector in the r-direction while $(-\sin\theta, \cos\theta)$ is – a unit vector in the θ-direction – is the same vector rotated anti-clockwise by 90 degrees.

B Transformation of the covariant derivative of a vector $\nabla\vec{V}$

The covariant derivative of a vector field is

$$\nabla\vec{V} \to V^{\alpha}{}_{;\beta} = V^{\alpha}{}_{,\beta} + \Gamma^{\alpha}{}_{\mu\beta}V^{\mu} \tag{120}$$

We would like to show that this transforms as a tensor:

$$V^{\alpha'}{}_{;\beta'} = \Lambda^{\alpha'}{}_{\alpha}\Lambda^{\beta}{}_{\beta'}V^{\alpha}{}_{;\beta}. \tag{121}$$

We start with the fact that the connection vanishes in the rectilinear (here primed) frame:

$$V^{\alpha'}{}_{;\beta'} = V^{\alpha'}{}_{,\beta'} = \partial_{\beta'}V^{\alpha'}. \tag{122}$$

Using $V^{\alpha'} = \Lambda^{\alpha'}{}_{\alpha}V^{\alpha}$ and $\partial_{\beta'} = \Lambda^{\beta}{}_{\beta'}\partial_{\beta}$, we have

$$V^{\alpha'}{}_{;\beta'} = \Lambda^{\beta}{}_{\beta'}\partial_{\beta}(\Lambda^{\alpha'}{}_{\alpha}V^{\alpha}) = \Lambda^{\beta}{}_{\beta'}\Lambda^{\alpha'}{}_{\alpha}V^{\alpha}{}_{,\beta} + \Lambda^{\beta}{}_{\beta'}\Lambda^{\alpha'}{}_{\alpha,\beta}V^{\alpha} \tag{123}$$

but the connection, in terms of the transformation matrices is $\Gamma^{\mu}{}_{\alpha\beta} = \Lambda^{\mu}{}_{\alpha'}\Lambda^{\alpha'}{}_{\alpha,\beta}$, so $\Lambda^{\alpha'}{}_{\alpha,\beta} = (\Lambda^{\mu}{}_{\alpha'})^{-1}\Gamma^{\mu}{}_{\alpha\beta} = \Lambda^{\alpha'}{}_{\mu}\Gamma^{\mu}{}_{\alpha\beta}$ so the second term above is $\Lambda^{\beta}{}_{\beta'}\Lambda^{\alpha'}{}_{\alpha,\beta}V^{\alpha} = \Lambda^{\beta}{}_{\beta'}\Lambda^{\alpha'}{}_{\mu}\Gamma^{\mu}{}_{\alpha\beta}V^{\alpha} = \Lambda^{\beta}{}_{\beta'}\Lambda^{\alpha'}{}_{\alpha}\Gamma^{\alpha}{}_{\mu\beta}V^{\mu}$ where we have switched the dummy indices $\alpha \Leftrightarrow \mu$, so

$$V^{\alpha'}{}_{;\beta'} = \Lambda^{\alpha'}{}_{\alpha}\Lambda^{\beta}{}_{\beta'}(V^{\alpha}{}_{,\beta} + \Gamma^{\alpha}{}_{\mu\beta}V^{\mu}) = \Lambda^{\alpha'}{}_{\alpha}\Lambda^{\beta}{}_{\beta'}V^{\alpha}{}_{;\beta} \tag{124}$$

which is indeed the transformation law for a $\binom{1}{1}$ tensor.

C Covariant derivative vs. gauge covariant derivative

Recall that, in electromagnetism, we can introduce the coupling of a charged particle to the EM field simply by changing the partial derivatives in the Schrödinger equation (or the KG equation for a complex classical scalar field) to *gauge covariant derivatives:*

$$\partial_\mu\psi \Rightarrow D_\mu\psi = \left(\partial_\mu - \frac{iq}{\hbar}A_\mu\right)\psi. \tag{125}$$

This bears a close resemblance to the covariant derivative of a vector:

$$V^{\nu}{}_{;\mu} = V^{\nu}{}_{,\mu} + \Gamma^{\nu}{}_{\gamma\mu}V^{\gamma}. \tag{126}$$

The similarity can be enhanced if we write the complex wave-function or field as a 2-component 'vector' living is the Argand plane.

People often say that the coupling of charged matter to the EM gauge field $\vec{A}$ exists in order that the Schrödinger (or KG) equation be invariant under an arbitrary local phase shift $\psi \Rightarrow \psi' = e^{i\theta(\vec{x})}\psi$. Whereas the appearance of the connection in the covariant derivative is there in order to make keep the covariant derivative of a vector field invariant under an arbitrary choice of coordinates.

This is rather pleasing, but the analogy is actually rather weak. It is true that, for *particles*, there is a great similarity between the geodesic equation for a massive particle – where, as we saw, in the weak field limit the spatial components are $\ddot{\mathbf{x}} = -\boldsymbol{\nabla}\phi_{\rm N}$ (with $\phi_{\rm N}$ the Newtonian gravitational potential) – and the equation of motion for a charge particle $\dot{u}_\alpha = qA_{[\mu,\nu]}u^\nu$, which gives, in the non-relativistic limit $\ddot{\mathbf{x}} = -q\boldsymbol{\nabla}\varphi = -cq\boldsymbol{\nabla}A_0$. So the result of a gravitational potential $\phi(\mathbf{x})$ on a mass is similar to that of an EM potential on a charge. But the relevant Christoffel symbol that appears in the covariant derivative in the weak field limit is, as we saw, $\Gamma^i{}_{00} = \partial_i\phi/c^2$, which is the *gradient* of the gravitational potential whereas it is the potential $\vec{A}$ itself that appears in the gauge covariant derivative, and not its derivative. So the respective potentials enter the two types of derivative above in different ways.

This difference is apparent if we couple the Schrödinger equation to gravity with the comma $\Rightarrow$ semicolon rule. Without gravity (or acceleration) the Schrödinger/KG equation is

$$\psi^{,\mu}{}_{,\mu} = k_{\rm C}^2\psi. \tag{127}$$

where $k_{\rm C} = mc/\hbar$ is the Compton wave number.

Now ψ is a scalar, so $\psi^{;\mu} = \psi^{,\mu}$ so $\psi^{;\mu}{}_{;\mu} = \psi^{,\mu}{}_{;\mu} = \psi^{,\mu}{}_{,\mu} + \Gamma^\mu{}_{\beta\mu}\psi^{,\beta}$ and the equation of motion is:

$$\psi^{,\mu}{}_{,\mu} + \Gamma^\mu{}_{\beta\mu}\psi^{,\beta} = k_{\rm C}^2\psi. \tag{128}$$

If we consider the gravitational influence as a small perturbation to a zeroth order (in $\phi_{\rm N}$) solution $\psi \simeq \psi_0\cos(ck_{\rm C}t)$ the gravitational influence involves $\Gamma^\mu{}_{0\mu}$ as we can neglect spatial gradients. But from the formula (108) for the Christoffel symbols this is

$$\Gamma^\mu{}_{0\mu} = \tfrac{1}{2}g^{\mu\gamma}(g_{\gamma 0,\mu} + g_{\gamma\mu,0} - g_{0\mu,\gamma}) \simeq \tfrac{1}{2}\partial_0\phi_{\rm N}/c^2 \tag{129}$$

so, perhaps surprisingly, the *spatial* derivative of the gravitational potential does not actually appear. For e.g. gravitational fields we encounter in Nature the potential changes on the *dynamical timescale* (the orbital timescale) and time derivatives of the potential are smaller than spatial derivatives by on the order of the velocity and so the gravitational coupling via the connecting here is negligibly small.

The way that gravity actually couples to such a field is not via the connection at all. It comes about because $\psi^{,\mu}{}_{,\mu} = g^{\mu\alpha}\psi_{,\alpha\mu} \simeq -(1+2\phi_{\rm N}/c^2)\frac{1}{c}^2\ddot{\psi} + \nabla^2\psi$. For non-relativistic fields $\ddot{\psi} \simeq -c^2k_{\rm C}^2\psi$ so the equation of motion becomes

$$\ddot{\psi} - c^2\nabla^2\psi = \omega_{\rm C}^2(\mathbf{x})\psi \tag{130}$$

where the Compton frequency is $\omega_{\rm C}(\mathbf{x}) = m(\mathbf{x})c^2/\hbar$ with $m(\mathbf{x}) \simeq m(1 + \phi_{{\rm N}(\mathbf{x})}/c^2)$. So the this behaves like a massive field where the mass is position dependent. This is preciely the same equation as obeyed by EM waves in a plasma, if we replace $\omega_{\rm C}(\mathbf{x})$ by the plasma frequency. This allows one to understand how matter waves (as dark matter) could be trapped in potential wells of galaxies etc. much as EM waves get trapped by the ionosphere.