PHZ 6607 Special and General Relativity I
Handout #1: The case for General Relativity

General Relativity is an exceptionally powerful physical theory. Ultimately, it is also a theory about geometry, the geometry of a four dimensional spacetime. In that theory, (timelike) geodesics represent the possible paths of infinitesimal test particles, whereby curvature in the spacetime reflects the presence of matter, subjecting the test particle to an influence which we have familiarly come to know as the gravitation force due to that matter. But the theory purports more: not only is curvature in the spacetime geometry meant to be due to the presence of physical matter, but also that curvature in turn determines the motion of all matter present, corresponding entirely to its gravitational interaction, all other interaction between the matter being determined by the other known forces in physics in the usual way. From this point of view, geodesic motion for a point, infinitesimal, test particle, which is presumed to have no influence back on the spacetime itself, represents one of the simplest cases of matter motion which might be taken for consideration.

Why is a theory such as this necessary? Firstly, Newtonian gravity is not a relativistic theory. Secondly, Special Relativity, which serves perfectly well to account for the behaviour of electromagnetic fields in a flat space, does not deal in any way with the consequences of gravitational interactions. Thirdly, Quantum Mechanics makes us acutely aware that particles, even photons identified as particles, must lose energy when climbing out of a gravitational potential. Presumably then, local electromagnetic energy density must depend on exactly where a local region lies within a gravitational potential. We have seen that the scalar potential of electrostatics must in general become a vector potential in the full theory of classical electromagnetism. We are then forced to ask what generalization becomes necessary for the static Newtonian potential, especially to describe situations characterized by very high relative velocities? General Relativity is one of the simplest attempts to answer that question, which it does in a fully relativistic way and which, so far, is not in conflict with experiment, nor with other direct or indirect physical observation.

General Relativity is not the only current theory of gravitation. Moreover, whenever an attempt has been made to construct the corresponding quantum theory, serious difficulties have always arisen which became insurmountable, suggesting to some that if General Relativity is a useful theory in physics, it is only as the (classical) low energy limit of some more fundamental, and very different, quantum theory. There is so little hard experimental evidence to support this view at the present time, but overwhelming evidence within specific theoretical frameworks. Nevertheless, as the difficulty of understanding General Relativity fully is currently so great, we shall entirely content ourselves with its direct investigation in this course, leaving any more complicated approach to a different study.

One further observational aside may well be in order given the current climate. That part of quantum mechanics which showed how a system could have a ground state, and a discrete spectrum of excited states, was far easier to accomplish than was solving all details of the Hydrogen atom. Even today, in the eyes of some Mathematical purists, the existence of the photon as a quantum particle state in electromagnetism is still not completely proven, though there are probably few Physicists who seriously doubt that proposition. The first part of our observation is that the same problem for the propagating degrees of freedom in General Relativity is basically no harder in flat spacetime than it is for electromagnetism. It is the nonlinearities in strong field regions, and the non propagating degrees of freedom which present the most serious problem for a quantum theory of General Relativity. On the other hand, and this is the second part of our observation, the non propagating degrees of freedom are much closer in nature to the particle aspects of the Hydrogen atom problem than they are to field theory aspects of electromagnetism.

PHZ 6607 Special and General Relativity I
Handout #2: Differential Geometry as a Spacetime Tool

As has been suggested, general relativity is a theory of the geometry of spacetime and of how it responds to the presence of matter. Thus, our first task will be to come to an understanding of geometry suitable for use in the theory of general relativity.

Geometry may mean slightly different things in different contexts. The primary elements we will have to deal with are distances and shape. We will need invariant ways to refer to distances, and to characterize shape we will most frequently use curvature. As in real life, the distance between two points will generally depend on the path taken between them. We will thus actually compute distance by integration along each particular path. The basic construct we will use for each infinitessimal element of path length will be a generalization of Pythagoras's theorem and of the familiar cosine rule. This construct will be embodied in the introduction of a metric, simply a convenient way of specifying infinitessmal distances.

An orange has a different shape from a sheet of paper. There are both topological and geometrical differences between these two objects and, for the most part, we will concentrate on the geometrical differences, as determined by the curvature. We will find that some measures of curvature depend on how we choose to refer distinct points on the objects, while othere do not, and invariance will be an important part of out investigation. Curvature will turn out to be very important for us, since it is the quantity which relates the presence of matter to its influence on the geometry of spacetime.

So far, I have been referring to points and elements of path length as though they were well understood, familiar entities. While that is true, our use of them will eventually go beyond the currently familiar, and we will need to introduce one other, very important geometrical notion, that of a manifold. This will be necessary, not only from a purely geometric point of view, but also to recreate the familiar notion of vector in a sufficiently broad, and more geometrical, context. Vectors, and tensors - their generalization, will become our bread and butter.

This course is not just about geometry. Matter curves geomerty in particular ways, and we will especially want to understand and investigate those ways. But matter moves very differently in a curved space, and we will also need to become familiar with those new and peculiar ways. A good springboard for launching into a full awareness of general relativistic effects will be to look at the relatively minor changes which occur when curvature is everywhere small. This, too, will require a particular language, but, by then, we will be able to use the language already learnt in the discussion of curved lines in a curved manifold.

Much of this course is concerned merely with giving a new mathematical formulation to ideas which are already familiar from physics. The introductory material is necessary to be able to write down and make sense of the non-linear field equations of General Relativity. It also includes an extensive treatment of the familiar topic of special relativity, which we will use heavily in a generalization of existing knowledge to the wider context needed for analysis on manifolds, and in particular for General Relativity. Topics such as vectors, the inertia tensor, inner products, special relativity, spherical harmonics, the rotation group (and angular momentum operators), conservation laws, the theory of intergration, gradient, divergence and curl, Gauss' and Stokes' theorems of vector calculus, Maxwell's equations and other gauge theories of physics will all take on a new light, based on a deeper and unified understanding founded in a geometrical approach to these topics. On the other hand, a topic such as fiber bundles is not essential to General Relativity, and we will not be concerned with it in this course. However the later phases will include many of the concepts and methods which are most heavily used in General Relativity, so we must deal with them in some detail.

PHZ 6607 Special and General Relativity I
Handout #3: Introduction on Differential Geometry

A manifold can be described simply as a mapping of a set of points into $\mathbf{R^n}$. Typical well known examples are the solid cube, the solid sphere and the solid donut, as three dimensional objects with their obvious embedding in $\mathbf{R^3}$. What might not be at all clear is whether the surfaces of each of these objects can be embedded as two dimensional manifolds into $\mathbf{R^2}$. In fact, none of them can be embedded globally, but a refinement of our definition allows us to continue to regard all these surfaces as two dimensional manifolds. It turns out to be sufficient (practically) to require that the embedding can be done locally, in open 'balls' which overlap to cover entirely the complete set of all points to be embedded.

We will find it necessary to define geometric structures on the manifolds with which we shall working. We will also find it very convenient to refer to other objects defined on manifolds, and the most common of these will be tensors, of which vectors are a simple and familiar example. However, perhaps in contrast with your current practice, we will use index position to have a meaning, and will (for a start) consider V^i and V_i as inequivalent, allowing them to refer to different tensor types. The most distinguishing property of tensors which we need to comprehend is their transformation law under changes of coordinates on the underlying manifold. These changes we can recognize from the way partial derivatives change under coordinate transformations. Thus:

$$V'^i = \frac{\partial x'^i}{\partial x^j} V^j,$$

for vectors with their index upstairs, which we will refer to as *contravariant* vectors, and

$$V'_i = \frac{\partial x^j}{\partial x'^i} V_j,$$

for vectors with their index downstairs, which we will refer to as *covariant* vectors. For tensors of more general type, with both upstairs and downstairs indices, the transformation law is simply a generalization of the above, with one factor for each index, the type of factor depending on the type of index - upstairs or downstairs.

Certain operations on tensors can now be defined without **any** additional geometric structure being given. Perhaps the most obvious are multiplication between tensors to form tensors of higher rank, and contraction of a pair of indices to form a tensor of lower rank. For example, $A^i A^j$, $A^i B_j$, and $B_i B_j$, are all different examples of distinct types of tensors of rank two, while $A^i B_i$, would be an example of contraction to a tensor of rank zero (*i.e.* a scalar). Notice that at present we have no way of defining the sort of contraction you might expect to correspond to $||A||^2$, which requires the additional structure known as a *metric*. A perhaps unexpected, but familiar, example of a second rank tensor might be a square matrix, M_i^j, defined on a manifold. M_i^i is a scalar known as the trace of the matrix. We can also see that a matrix can be contracted with a vector to form another vector of the same type, *e.g.* $A^i M_i^j$ and $M_i^j B_j$, while the contraction of a matrix with another leads to a third matrix, exactly as in the familiar matrix multiplication, *i.e.* $P_i^k = M_i^j N_j^k$. Two other special tensors, at least one of which you may have encountered before, are ε_{ijk} and ε^{ijk}, each defined to be $+1, -1$, or 0 in locally cartesian coordinates, depending on whether $\{ijk\}$ is an even, odd or not a permutation of $\{123\}$. With these we can also define $Det(M) = \varepsilon^{ijk} M_i^l M_j^m M_k^n \varepsilon_{lmn}$, as well as the familiar vector triple product $\varepsilon_{ijk} A^i C^j D^k$. Finally, another operation which can sometines be used to define a new tensor is differentiation. Thus:

$$\partial_i B_j - \partial_j B_i \ (Exterior)\,;\ A^i \partial_i C^j - C^i \partial_i A^j \text{ and } A^i \partial_i B_j + B_i \partial_j A^i \ (Lie)\,, \text{ where } \partial_i = \partial/\partial x^i,$$

are all tensors. More general differentiation requires additional structure (a *connection*).

PHZ 6607 Special and General Relativity I
Handout #4: Introducing Vectors — the Old and the New!

The concept of a metric can be introduced as a fairly simple generalization of Pythagoras's theorem. But the definitions: - a *vector* is ... a *tensor* is ... a *manifold* is ... may at first sight seem a little formal. Occasionally you will find it useful to refer back to their formal definitions, but for the most part you can use minor extensions of concepts that you may have already encountered, but only implicitly. An important part of what we shall do is to make those hidden concepts explicit: Within certain inexactitudes, we will do this by recognizing familiar examples first, rather than by discussing the formal definitions.

A good example with which to start is the idea of a vector, which in a pre-school definition was "a directed line segment." Well, our vectors will be that and so much more. But other related ideas which we might take for granted will not automatically be applicable. For example, the idea of a length for a vector rests upon a scheme for measuring length, and whereas a vector itself does not represent such a scheme, a metric does. Previously, you have probably always implicitly used a metric, without specific reference to it. The apparently fundamental idea of a position vector will be seen to break down in an elementary way with the formal definition we shall work with henceforth.

The first generalization of the notion of vector we shall introduce is the idea that we can have vectors of different types. In fact you will already have used these two types previously without distinction. For us, this distinction will prove very useful, and an advantage which it entails is that the two types can be "dotted" into each other *without* the need for a metric. What will pass for ordinary vectors we will represent with an index upstairs, thus: V^a. The other type we shall denote with its index downstairs, W_a, and, in three dimensions, the ordinary vector cross (or wedge) product will produce such an entity (almost). As you will recall, the vector triple product is a scalar, and is obtained by dotting a vector into a cross product; that is, by something of the form - $V^a W_a$, where we have introduced a convention in which the repeated index is automatically to be summed, as is required for a dot product. No reference to a metric was necessary to construct this scalar. Instead, we have used the two different types of vectors to accomplish that.

It should have occurred to you as curious, by now, that the cross product of two vectors of one kind - each with an index upstairs - should be capable of resulting in a vector of the other type - with its index downstairs. In this case, to understand clearly, we have to both introduce the broader notion of a tensor, and relate it to a particular object of previous encounter, namely the totally antisymmetric ε-symbol.*

The notation above already hides part of the definition we have always used for a vector, since the idea of basis vectors is implied but not stated. The notation above refers simply to components of the vectors in their respective bases, while a further convention about the dot product between different basis vectors is hidden in the summation convention used above. Although, often, we will restrict our use to what are called coordinate bases, you have already encountered both coordinate and non-coordinate bases, for example when you have dealt with the vector potential in cartesian and polar coordinate systems in Electromagnetism. In particular, a coordinate basis may conveniently be used to investigate how coordinate transformations affect the two different types of vectors. As you might guess, the transformation properties of basis vectors and components are complementary so that vectors themselves remain unchanged, but not their representation.

---

* In fact, I shall later have to point out that there is a deception in this statement.

PHZ 6607 Special and General Relativity I
Handout #5: Introducing Tensors — the New and the Newer!

One difficult result we will have to deal with at this point is the fact that, in three space, the so-called position vector is **not** a vector under general coordinate transformations, although it does behave as a vector under the restricted class of transformations known as rotations. However, the components of infinitesimal displacement really do form a (contravariant) vector since they satisfy the appropriate transformation law:

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j.$$

Thus some of the most common vectors we use to refer to the motion of a particle in Newtonian physics, namely velocity and momentum, both behave as contravariant three-vectors; in that context, time is merely a parameter with which we can label points along the path of the particle. (Before considering acceleration, we will need to know how to deal more fully with the differentiation of a vector.) The fact that proper four-vectors can be formed in a similar way also plays a crucial rôle in the invariant formulation of special relativity. This will be a particularly useful perspective to keep in mind when we come to deal with special relativity again shortly. At that time we will also expand on the usual understanding of electromagnetism within the framework of special relativity.

The velocity vector along a curve is a familiar example of a (contravariant) vector. The gradiant operator acting on a function defined everywhere on a manifold is also familiar, and it creates a (covariant) vector field, which is itself defined at every point on a manifold. Their dot product can be defined without any additional structure on the manifold, and it results in something we can actually identify, namely the total change of the function along the curve, which is a scalar quantity:

$$\frac{df(x^i)}{dt} = \frac{dx^i}{dt} \frac{\partial f(x^i)}{\partial x^i} = v^i \cdot \nabla_i(f).$$

One might perhaps use an anology from quantum mechanics to clarify the distinction between a covariant and a contravariant vector. In quantum mechanics one has both bra states, $\langle \phi |$, and ket states, $| \psi \rangle$. We also normally think that we can map from one of these to the other, but technically this requires a structure called complex conjugation to be defined with respect to Hilbert space. The advantage of having both types of state vectors on Hilbert space is that we can form their their (dot or inner) product. Similarly, without a metric, covariant and contravariant vectors are completely distinct, but we can still form their dot product. The advantage of a metric as a defined structure on a manifold (anologously to the complex structure on Hilbert space) is that it allows to construct a map from contravariant vectors to covariant vectors.

The notion of a tensor itself is fairly easy to introduce. Examples of tensors can be obtained from a particular product of vectors, in which the indices are all independent, and not in general summed; for example $T^{ab} = X^a Y^b$, would be a tensor of rank two, as also would $V^a W_b$ and $W_a W_b$, in particular, though again of different types.

Consider the two dimensional metric for flat space in cartesian coordinates:

$$ds^2 = dx^2 + dy^2,$$

and impose the coordinate transformation:

$$x' = x\cos\theta + y\sin\theta \text{ and}$$
$$y' = -x\sin\theta + y\cos\theta,$$

for which the new metric becomes:
$$ds^2 = dx'^2 + dy'^2.$$

As you know, this is a simple example of a rotation which preserves angles and distances, and it clearly leaves the form of the flat space metric unchanged. Now consider the coordinate transformation:

$$x' = x\cosh\theta + t\sinh\theta \text{ and}$$
$$t' = x\sinh\theta + t\cosh\theta,$$

for the two dimensional flat spacetime metric:

$$ds^2 = -dt^2 + dx^2.$$

You will find it straightforward (and to your advantage) to show that the form of this metric also remains unchanged, and you will recognize this coordinate transformation as a simple example of a Lorentz transformation, for which $v/c = \tanh\theta$ and $\gamma = \cosh\theta$. Thus, boosts are to a Lorentzian (or pseudo-Reimannian) geometry, with indefinite metric, what rotations are to an Euclidean (or Reimannian) geometry, with definite metric. In accord with the similarity we have just established between these two operations, there is also a Lorentzian analogue of the transformation to polar coordinates, *i.e.*:

$$x = \rho\cosh\tau, \text{ and } t = \rho\sinh\tau,$$

for which the metric becomes:

$$ds^2 = -\rho^2 d\tau^2 + d\rho^2.$$

Although you will recognize that this coordinate transformation, as given, is not defined over all of Minkowski space, a similar transformation turns out to be extremely important in the study of black hole spacetimes, and we will be making further use of it in that context.

    Another exercise which is very helpful in developing an understanding of black hole spacetimes, is to solve the equations of motion for a particle whose Lagrangian is given by:

$$\mathcal{L} = \frac{1}{2}mg_{\mu\nu}\dot{x}^\mu\dot{x}^\nu, \text{ in which } \dot{} = d/d\lambda,$$

where now the metric is that of four dimensional flat spacetime, but again expressed in (spatial) spherical polar coordinates. To obtain the full advantage of this exercise it is important to have explicit expressions for the conserved components of linear and angular momentum in terms of the canonical momenta defined from this Lagrangian (namely, $\partial\mathcal{L}/d(\dot{x}^\mu)$). You should give this exercise some further thought, since it will be appearing as a homework problem shortly.

You should now be able to realize that the invariant line element:

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \equiv g_{\mu\nu}dx^\mu dx^\nu,$$

is preserved under a restricted Lorentz transformation. In fact the complete Lorentz group of transformations can be defined as that which leaves this invariant element unchanged. This might appear to be the reverse of the order in which things have been defined for you in Special Relativity, but actually we also define other groups of transformations in exactly this way. Two examples which we shall encounter again are the Rotations in three space, which leave angles and distances unchanged, and Conformal transformations which leave angles unchanged, but not distances.

This line element also introduces one of the structures on manifolds which we shall be working with continually, namely the metric tensor, $g_{\mu\nu}$. The metric carries the information which determines distances (and time separations), and is very helpful in establishing a relationship between the manifolds with which we shall be working, and the physics on them which we shall be interested in discussing. Another structure which will also prove to be essential is a thing called a connection, which again can be demonstrated from purely flat space considerations. Unlike the metric, the connection is not a tensor, but it will be required in the definition of the curvature tensor, which we shall also be using repeatedly.

There is a very easy way to witness the emergence of a connection. As you will notice from a course textbook, the equation for a geodesic, as a suitably parameterized curve embedded in a higher dimensional manifold, may be given by:

$$\frac{d^2x^i}{d\lambda^2} + \Gamma^i_{jk}\frac{dx^j}{d\lambda}\frac{dx^k}{d\lambda} = 0,$$

where λ is a parameter along the curve; in three dimensional space it may be taken as time. Now geodesics in Euclidean three space, $\mathbf{R^3}$, are actually straight lines. And you also know that free motion in flat space takes place in straight lines. Suppose, then, we look at trying to derive this information in spherical polar coordinates, instead of the usual cartesian coordinates. Then, for some chosen orientation of coordinates, the non-relativistic Lagrangian for this motion would be:

$$\pounds = \frac{1}{2}mv^2 = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\right).$$

The Euler Lagrange equations become:

$$-\ddot{r} + r\dot{\theta}^2 + r\sin^2\theta\,\dot{\phi}^2 = 0,$$
$$-(r^2\dot{\theta})\dot{} + r^2\sin\theta\cos\theta\,\dot{\phi}^2 = 0, \text{ and}$$
$$-(r^2\sin^2\theta\,\dot{\phi})\dot{} = 0,$$

where time has indeed assumed the rôle of the parameter λ above. From these we are (almost) immediately able to read off and identify:

$$\Gamma^r_{\theta\theta} = -r, \ \Gamma^r_{\phi\phi} = -r\sin^2\theta, \ \Gamma^\theta_{r\theta} = \frac{1}{r}, \ \Gamma^\phi_{r\phi} = \frac{1}{r}, \ \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta, \text{ and } \Gamma^\phi_{\theta\phi} = \cot\theta,$$

all other possibilities being zero. We shall shortly find the geometrical reason for these results which, in flat space, correspond to a rather special, but physically natural, choice for the structure which has been called the connection, and is represented here by Γ.