Physics 407-07
Vectors

Let us consider a space of raw points, designated usually by the symbol p. This space has nothing else defined on it– no metric, no topology. These points are supposed to represent the points in space or in spacetime.

On this space, we can define two things, curves (which are maps from the space of the real numbers into the space of points) and functions (which are maps from the space of points into the real numbers). However, we are not going to be interesed in just any such curves or functions, but in a subclass of them, which I will call nice curves of functions. At present, it is unclear what the definition of nice is, however the first condition we will put on them is that if we take a nice curve, designated by $\gamma(\lambda)$, and a nice function, $f(p)$, then the composite function from real numbers to real numbers, namely $f(\gamma(\lambda))$, is differentiable. Ie, for all λ for which the the function is defined, $\frac{df(\gamma(\lambda))}{d\lambda}$ is defined. Of course this is not yet sufficient to define what these nice function or curves are. For example, we could take the nice functions to be all functions, and then the nice curves would have to be constants (ie $\gamma(\lambda) = p_0$ a constant point for all λ). We will soon find that we need or definition of "nice" functions to be a bit more restrictive than that.

However, even with this bare bones structure, we can already define two different kinds of vectors, tangent vectors to a curve $\gamma(\lambda)$ and cotangent (essentailly gradient) vectors to the functions $f(p)$, as little pieces of the curve, and little pieces of the funtions.

We will say that two curves$\gamma(\lambda)$ and $\gamma'(\lambda)$, both going through a point p have the same tangent vector if for **all** nice functions $f(p)$, the derivatives are the same. Ie,

$$\frac{df(\gamma(\lambda))}{d\lambda}|_{\lambda=\lambda_0} = \frac{df(\gamma'(\lambda))}{d\lambda}|_{\lambda=\lambda'_0} \tag{1}$$

where $\gamma(\lambda_0) = \gamma'(\lambda'_0) = p$ the point through which both curves run.

The tangent vector to the curve I will designate by $\left(\frac{\partial}{\partial\gamma}\right)^A$.

Similarly we can define the cotangent vector corresponding to the function f with df_A and define it such that two functions f and f' have the same cotangent vector at the point p if and only if for all nice curves $\gamma(\lambda)$ going through the point p,

$$\frac{df(\gamma(\lambda))}{d\lambda}|_{\lambda=\lambda_0} = \frac{df'(\gamma(\lambda))}{d\lambda}|_{\lambda=\lambda_0} \tag{2}$$

We can represent the tangent vector by a little arrow, which is the curve $\gamma(\lambda)$ from λ_0 to λ_0+1 with the arrow head at λ_0+1. Usually we use "straight lines" but we have no idea yet what a "straight line" means. That takes far more structure on the space. At present, we can take any one of the curves which have the same tangent vector to represent that tangent vector.

Similarly, we can represent the cotangent vectors by "little pieces of a representative function". How do we represent that? By its level surfaces. Ie, chose all points p such that $f(p) = f(p_0)$ and all points p such that $f(p) = f(p_0) + 1$ as representing the functions with the same cotangent vector at p_0.

On cotangent vectors, one has concept of addition. Namely if $f(p)$ and $g(p)$ are two functions, then the cotangent vector to $f(p) + g(p)$ (which is clearly also a nice function) is defined as the sum of the cotangent vectors to f and g

$$d(f+g)_A = df_A + dg_A \tag{3}$$

Similarly we can define the product of a real number times the vector as the cotangent vector to the function multiplied by that number. If the number is r, then

$$d(rf)_A = r(df_A) \tag{4}$$

Both of these definitions are consistant in that it does not matter which representative function we take such that its cotangent vector is df_A and for dg_A, the sum is the same cotangent vector.

For tangent vectors we have problems. We can define the a tangent vector r times as big by defining the curve

$$\gamma_r(\lambda) = \gamma(\lambda_0 + 5(\lambda - \lambda_0)) \tag{5}$$

where $\gamma(\lambda_0) = p$ the point of interest, but there is no way of adding two such curves. If one had two curves $\gamma(\lambda)$ and $\gamma'(\lambda)$, both going through the point p (and let us assume, without loss of generality that both curves went through p at $\lambda = 0$) that if one defined a curve $\Gamma(\lambda)$ which was supposed to be a representative sum of the two curves, that

$$\frac{df(\Gamma(\lambda))}{d\lambda} = \frac{df(\gamma(\lambda))}{d\lambda} + \frac{df(\gamma'(\lambda))}{d\lambda} \tag{6}$$

for all nice f, but it is not clear that any such Γ exists. In fact with a perverse enough definition of "nice" functions, it need not exist.

In some sense, the cotangent vectors are more general than the tangent vectors.

We see that there also exists a definition of the product of a tangent vector and a cotangent vector. Given a tangent vector at a point V^A and a cotangent vector W_A, we can define the product $V^A W_A$ as

$$V^A W_A = df(\gamma(\lambda))d\lambda \tag{7}$$

where f and γ are chosen such that $V^A = \frac{\partial}{\partial\gamma}^A$ and $W_B = df_B$. (Note that the value of the subscript or superscript letter does not matter, it is its existence that matters. However, since we will be defining other products of vectors, in defining this product, we make sure that the subscript on the cotangent vector is the same as the superscript on the tangent vector.)

0.1 Coordinates

Let us now restrict our space of discourse still futher. It is not clear what this restriction really means, but historically it has proven to be very useful.

Let us assume that we can choose, at least over a subset of the "Manifold" of points we are interested in looking at, N nice functions, which I will designate by $\{x^i\}$, where i goes from 1 to N, or sometimes from 0 to N-1. These functions are to be such that if we look at the set of points p which obey the equation

$$x^i(p) = x^i_0 \tag{8}$$

for all i, where x^i_0 are a set of N real numbers, then there is at most one point p which satisfies this in the subset of all points we are interested in. Ie, the x^i can be regarded as labels for the points, with a unique label for each point p in the subset. These N functions are the coordinates for the point

p. Note that they are arbitary, in that I have said nothing about what these functions are or how they are chosen, except that they are unique for each point p.

Now we place a further restriction on our set of nice functions. I will assume that the curve through the point p_0 defined by

$$\begin{aligned} x^i(p) &= x^i(p_0) \quad \text{for all } i \neq j &(9)\\ x^j(p) &= x^j(p_0) + \lambda &(10) \end{aligned}$$

for specific choice of j, is also a nice curve. This curve, which I could designate by $\gamma_j(\lambda)$ is the j^th coordinate axis. The tangent vector to this curve (and since it is **by assumption** a nice curve, it has a tangent vector) is designated either by $\frac{\partial}{\partial \gamma_j}^A$ or more generally by $\frac{\partial}{\partial x^j}^A$. Note that while it looks like a partial derivative, this is simply a symbol designating the tangent vector to the curve which is the j^{th} coordinate axis.

However, if we have a function $f(p)$ expressed as a function of the coordinate $F(x) = f(p(x))$, then the partial derivative of F is exactly the derivative of f along the coordinate axis. Ie, $\frac{\partial F}{\partial x^i} = \frac{df(\gamma^i(\lambda))}{d\lambda}$. This is where the notation comes from.

We will always assume that all of the spaces we study have such coordinates, and that the definition of nice functions are such that such coordinates exist. If a space has such sets of nice functions, mathematicians call such spaces differentiable manifolds.

We can now define the sum of two tangent vectors. Consider two curves γ and γ' going through the point p_0. Let us assume that for both curves $\gamma(0) = \gamma'(0) = p_0$. Then define a new curve

$$\Gamma(\lambda) = P(\{x^i(\gamma(\lambda)) + x^i(\gamma'(\lambda)) - x^i(p_0)\}) \tag{11}$$

where the function $P(\{x^i\})$ is the point desigated by the set of coordinate values $\{x^i\}$. Ie, the sum curve is defined via the sum of the coordinates of curves γ and γ'. Then the tangent vector to Γ is defined as the sum of the tangent vectors to γ and γ'.

The curve Γ clearly depends not only on the curves γ and γ' but also on the coordinates which we have chosen. However it is possible to prove, because of the nice properties of the coordinates, that the tangent vector to Γ depends only on the tangent vectors to γ and γ'.

We now have two different kinds of vectors, tangent and cotangent, defined. (in the older literature these are called contravariant and covariant vectors). Note that while they are both vectors, they really have nothing to do with each other. They are simply two different kinds of mathematical and physical things that we can define. Just as curves and functions are two different kinds of things, which really have little to do with each other, so are tangent vectors (little pieces of curves) and cotangent vectors (little pieces of functions.)

0.2 Components

Given our coordinates, and out definitions of vectors, we can express the vectors in terms of each other.

Consider the function f(p) defined near the point p_0. Let us define another function F(p) by

$$F(p) = f(p_0) + \sum_i \frac{\partial f(P(\{x^i\}))}{\partial x^i}|_{\{x^i = x^i(p_0)\}} (x^i(p) - x^i(p_0)) \tag{12}$$

It is possible to show that F(p) has the same cotangent vector as $f(p)$ has at the point p_0. Ie for all curves $\gamma(\lambda)$, the derivative along the curve of these two functions is the same at the point p_0. This means that

$$df_A = dF_A \tag{13}$$

But

$$dF_A = \sum_i \frac{\partial f(P(\{x^i\}))}{\partial x^i}|_{\{x^i = x^i(p_0)\}} dx^i_A \tag{14}$$

since it is a sum with constant coefficients of the coordinate functions $x^i(p)$. Thus we can write

$$df_A = \sum_i \frac{\partial f(P(\{x^i\}))}{\partial x^i}|_{\{x^i = x^i(p_0)\}} dx^i_A \tag{15}$$

The coefficients, $\frac{\partial f(P(\{x^i\}))}{\partial x^i}|_{\{x^i = x^i(p_0)\}}$ are called the components of df_A in the coordinate system $\{x^i\}$.

Similarly we can for any curve γ write

$$(\frac{\partial}{\partial\gamma})^A = \sum_i \frac{dx^i(\gamma(\lambda))}{d\lambda} \left(\frac{\partial}{\partial x^i}\right)^A \tag{16}$$

Then the $\frac{dx^i(\gamma(\lambda))}{d\lambda}$ are the components of $(\frac{\partial}{\partial\gamma})^A$ in the coordinate system $\{x^i\}$.

Finally, we can see that

$$(\frac{\partial}{\partial x^i})^A dx^j_A = \delta^j_i \tag{17}$$

and thus

$$V^A W_A = \sum_i V^i W_i \tag{18}$$

This also shows that this product of the sums of components is independent of which coordinate system one happens to have chosen, because the left had side was defined without any reference to coordinates.

0.3 metric

While the above structures are useful, in almost all of physics, another structure plays a crucial role, namely a metric. This is something which determines the size of things. The metric is defined as the generalisation of the dot product of two tangent vectors. In particular, given two tangent vectors V^A and W^B (again the value of the superscript does not matter). We thus define a function g of the two vectors

$$g(V^A, W^B) \tag{19}$$

to the real numbers as the "dot product" of two tangent vectors. We demand, primarily be analogy with the dot product, that this metric be linear in both arguments.

$$g(V^A, W^B + Z^B) = g(V^A, W^B) + g(V^A, Z^B) \tag{20}$$

and that g be symmetric

$$g(V^A, W^B) = g(W^B, V^A) \tag{21}$$

Now we define the length squared of a vector to be given by $g(V^A, V^A)$. This allows us also to define the dot product in terms of lenths

$$g(V^A, W^B) = \frac{1}{2}\left(g(V^A + W^A, V^B + W^B) - g(V^A, V^B) - g(W^A, W^B)\right) \quad (22)$$

Writing V^A and W^B in terms of coordinates components, we get

$$g(V^A, W^B) = \sum_{ij} V^i W^j g(\left(\frac{\partial}{\partial x^i}\right)^A, \left(\frac{\partial}{\partial x^j}\right)^B) \equiv \sum_{ij} V^i W^j g_{ij} \quad (23)$$

The numbers $g_{ij} = g(\left(\frac{\partial}{\partial x^i}\right)^A, \left(\frac{\partial}{\partial x^j}\right)^B)$ are called the components of the metric in the coordinate system x^j. Note again that the metric was defined independent of coordinates, and thus since the left had side is independent of coordinates, so must the sum of the right hand side be, even though the values of the coefficients clearly do depend on the coordinates.

0.4 Length of a curve

Given a curve $\gamma(\lambda)$, the lenght of the curve from the point $p_1 = \gamma(\lambda_1)$ to $p_2 = \gamma(\lambda_2)$ is defined to be

$$\begin{aligned} &\int_{\lambda_1}^{\lambda_2} \sqrt{g(\left(\frac{\partial}{\partial\gamma}\right)^A, \left(\frac{\partial}{\partial\gamma}\right)^B)} d\lambda && (24) \\ &= \int_{\lambda_1}^{\lambda_2} \sqrt{\sum_{ij} g_{ij} \frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}} d\lambda && (25) \end{aligned}$$

Note that because of the square root, the right side of this equation is independent of the parameterisation λ we choose for the curve. Ie, the length is function only of the curve between the two points and not of the parameterisation one uses along the curve.

Note that we will run into trouble if the expression for the length of the tangent vector is negative, since then the square root would be imaginary. In this case one must fudge things. One usually defines the length of a curve by taking the absolute value inside the square root, but this can run into trouble if the argument alters in sign along the curve. One almost always ignores such possibilities. One distinguishes the curves by the sign of the argument of the square root, and keeps curves with different signs separate.

0.5 Straight Lines

Now that we have a notion of length, we can discuss what we mean by a straight line. Euclid had the same problem, and he defined a straight line as the shortest distance between two points. While this often works, in special relativity, we know that for some straight lines (eg timelike curves) the straight line is the longest distance between two points.

Let us define a family of curves, $\gamma(\epsilon, \lambda)$ where the ϵ designates different curves between two points, which I will assume are always located at λ_1 and λ_2. Let the function $D(\epsilon)$ designate the distance between these two points along the various curves. We will say that the curve $\gamma(0, \lambda)$ is a straight line between the two points if for all sets of curves $\gamma(\epsilon, \lambda)$ such that $\gamma(0, \lambda)$ is that same curve, that $\frac{dD(\epsilon)}{d\epsilon}$ is zero. Ie, for all sets of curves, the given curve is at a relative minimum, maximum, or inflection point. Note we will always demand that the curves be nice curves.

Writing this in terms of coordinates, we have the expression

$$\frac{dD}{d\epsilon} = \int_{\lambda_1}^{\lambda_2} \frac{d}{d\epsilon}\sqrt{\sum_{ij} g_{ij}(x^k(\gamma(\epsilon,\lambda)))\frac{dx^i(\gamma(\epsilon,\lambda))}{d\lambda}\frac{dx^j(\gamma(\epsilon,\lambda))}{d\lambda}}d\lambda \tag{26}$$

Defining

$$S(\epsilon,\lambda) = \sqrt{\sum_{ij} g_{ij}(x^k(\gamma(\epsilon,\lambda)))\frac{dx^i(\gamma(\epsilon,\lambda))}{d\lambda}\frac{dx^j(\gamma(\epsilon,\lambda))}{d\lambda}} \tag{27}$$

we have

$$\int \frac{1}{2S} \qquad \sum_{ij}\left(\sum_k \frac{\partial g_{ij}}{\partial x^k}\frac{dx^k}{d\epsilon}\frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}\right. \tag{28}$$

$$\left. g_{ij}\frac{d^2x^i}{d\epsilon d\lambda}\frac{dx^j}{d\lambda} g_{ij}\frac{dx^i}{d\lambda}\frac{d^2x^j}{d\epsilon d\lambda}\right) d\lambda \tag{29}$$

Since g_{ij} is symmetric, and since i, j, k are just "dummy" summation variables, we can rename them in the various terms to get

$$\int \frac{1}{2S}\sum_k\left(\sum_{ij}\frac{\partial g_{ij}}{\partial x^k}\frac{dx^k}{d\epsilon}\frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}\right. \tag{30}$$

$$\left.\sum_j 2g_{kj}\frac{d^2x^k}{d\epsilon d\lambda}\frac{dx^j}{d\lambda}\right) d\lambda \tag{31}$$

Integrating the second term by parts, and recalling that $\frac{dx^i}{d\epsilon}$ is zero at λ_1 and λ_2 (all the curves we are comparing are supposed to go through the same points at their endpoints), this expression becomes

$$\int \sum_k \frac{dx^k}{d\epsilon} \left(\frac{1}{2S} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^k}{d\epsilon} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} \right. \tag{32}$$

$$\left. - \frac{d}{d\lambda} \left[\frac{1}{S} g_{kj} \frac{dx^j}{d\lambda} \right] \right) d\lambda \tag{33}$$

Since we said that we wanted this to be zero for all sets of curves, the only way we can do this is if each term multiplying any $\frac{dx^k}{d\epsilon}$ for each value of k and for each point along the curve must be zero. Otherwise one can always choose a set of curves such that the integral is not zero. Ie, we get the second order differential equation

$$\frac{d}{d\lambda} \left(\frac{1}{S} \sum_j g_{kj} \frac{dx^j}{d\lambda} \right) = \frac{1}{2S} \sum_{ij} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda} \tag{34}$$

This is the geodesic equation.

This equation and the derivation can be simplified if we make a special choice for the parameter λ. Namely, if we choose λ to be such that $S = 1$, the S disappears from the above equation. This parameter is usually designated by s. Furthermore, if we choose this parameterisation, then we have that

$$\left(\frac{d}{d\epsilon} \int S^n ds \right) |_{\epsilon=0} = \left(n \int S^{n-1} \frac{dS}{d\epsilon} ds \right) |_{\epsilon=0} = n \left(\int \frac{dS}{d\epsilon} ds \right) |_{\epsilon=0} = 0 \tag{35}$$

since along the solution curve $S = 1$. Ie, if we choose our parameter λ correctly (ie, to be equal to the path length), we can place an arbitrary power of S, and in particular we can use n=2 to get rid of the horrible square root, in the variation.

Thus, if we choose this parameterization, the geodesic equation becomes

$$\frac{d}{ds} \left(\sum_j g_{kj} \frac{dx^j}{ds} \right) = \sum_{ij} \frac{\partial g_{ij}}{\partial x^k} \frac{dx^i}{ds} \frac{dx^j}{ds} \tag{36}$$

with the additional requirement that

$$\sum_{ij} g_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds} = 1 \tag{37}$$

It is easy to show that this second constraint equation is consistant with the second order equations above.

An example: Consider the metric

$$ds^2 = dr^2 + r^2 d\theta^2 \tag{38}$$

Using the above we need to extremize the integral

$$\frac{d}{d\epsilon}\int (\frac{dr}{ds})^2 + r^2(\frac{d\theta}{ds})^2 ds = 0 \tag{39}$$

Let us first take the set of curves to be such that $\frac{d\theta}{d\epsilon}$=0. Then we have

$$\begin{aligned} 0 &= \int 2(\frac{dr}{ds})\frac{d^2r}{d\epsilon ds} + 2r\frac{dr}{d\epsilon}(\frac{d\theta}{ds})^2 ds & (40)\\ &= 2\int \frac{dr}{d\epsilon}(-\frac{d^2r}{ds^2} + r(\frac{d\theta}{ds})^2)ds & (41)\end{aligned}$$

and since $\frac{dr}{d\epsilon}$ is arbitrary (except at the end points which was why the endpoint contributions in the integration by parts disappeared), we must have as our first equation that

$$-\frac{d^2r}{ds^2} + r(\frac{d\theta}{ds})^2 = 0 \tag{42}$$

Now choosing the set of paths so that $\frac{dr}{d\epsilon}$ is zero, we get

$$\begin{aligned} 0 &= \int 2r^2\frac{d\theta}{ds}\frac{d^2\theta}{ds^2}ds & (43)\\ &= -2\int \frac{d\theta}{d\epsilon}(\frac{d}{ds}\left(r^2\frac{d\theta}{ds}\right) & (44)\end{aligned}$$

which by the same reasoning on the arbitrariness of $\frac{d\theta}{d\epsilon}$ gives

$$\frac{d}{ds}\left(r^2\frac{d\theta}{ds}\right) = 0 \tag{45}$$

The third equation is

$$(\frac{dr}{ds})^2 + r^2(\frac{d\theta}{ds})^2 = 1 \tag{46}$$

The solution is

$$\frac{d\theta}{ds} = \frac{L}{r^2} \tag{47}$$

for some constant L and then

$$(\frac{dr}{ds})^2 = 1 - \frac{L^2}{r^2} \tag{48}$$

which gives

$$\begin{aligned} s &= \int \frac{dr}{\sqrt{1 - \frac{L^2}{r^2}}} && (49) \\ s - s0 &= \sqrt{r^2 - L^2} && (50) \end{aligned}$$

or

$$r = \sqrt{(s - s_0)^2 + L^2} \tag{51}$$

Substituting into theequation for θ we have

$$\theta - \theta_0 = \int \frac{L}{(s - s_0)^2 + L^2} ds = atan(\frac{s - s_0}{L}) \tag{52}$$

Note that if we choose s_0 such that $\theta_0 = 0$, and define

$$x = r\cos(\theta), \qquad y = r\sin(\theta) \tag{53}$$

we have

$$x = L, \qquad y = s - s_0 \tag{54}$$

0.6 Inverse metric

We now have two different functions of a tangent vector which give a number. For any tangent vector V^A, the function of Z^B given by $f_{V^A}(Z^B) = g(V^A, Z^B)$. Similarly for any cotangent vector U_A we have the function

$h_{U_A}(Z^B) = Z^A U_A$ is also a function from the set of vectors to the real numbers. Now, given a vector V^A one can always find a cotangent vector U_B such that $h_{U_A}() = f_{V^A}$. To see this is most easily done using the components.

$$h_{U_A}(Z^B) = \sum_i U_i Z^i \tag{55}$$

$$f_{V^A}(Z^B) = \sum_i (\sum_j g_{ji} V^j) Z^i \tag{56}$$

Ie, if we choose

$$U_i = \sum_j g_{ij} V^j \tag{57}$$

we see that both functions f and h give the same value for all values of Z^A. Ie, the metric allows us to associate a unique cotangent vector for each tangent vector. Furthermore, it may allow us to associate a length to cotangent vectors. Ie, if U_A is assciated with V^A an W_B is associated with Z^B we can define a dot product $\tilde{g}(U_A, W_B) = g(V^A, Z^B)$ However, it is not true that this defines a metric for all cotangent vectors necessarily. If the U_A associated with V^A is the null vector, when V^A is not a null vector, then there will be cotangent vectors which have no tangent vector as their image. The easiest case to see is if the metric g is zero for all arguments. Then clearly each tangent vector has the zero cotangent vector associted with it and for no non-zero cotangent vector is there any tangent vector.

In everything we do we will assume that this is not true, but rather that for each cotangent vector there is a unique, non-zero tangent vector which gives that cotangent vector via the metric. Ie, for each non-sero U^A there exists a unique V^A such that

$$U_i = \sum_j g_{ij} V^j \tag{58}$$

This means that there must be another set of numbers, which I will designate by g^{ij} such that if

$$U_i = \sum_j g_{ij} V^j \tag{59}$$

then

$$V^i = \sum_k g^{ik} U_k \tag{60}$$

This gives that for all vectors V^A, we have

$$V^i = \sum_{kj} g^{ik} g_{kj} V^k \tag{61}$$

or

$$\sum_{kj} g^{ik} g_{kj} = \delta^i_j \tag{62}$$

Ie, the matrix represented by the coefficients g^{ik} is the inverse matrix to the matrix g_{ik}.

Thus the tangent metric and the cotangent metric can be used to map tangent vectors to cotangent vectors or cotangent vectors to tangent vectors.

Consider a function $f(p)$ and a curve $\gamma(\lambda)$ such that $\gamma(\lambda)$ lies entirely withing the level surface of f. Ie, $f(\gamma(\lambda)) = f(\gamma(\lambda'))$ for all λ, λ'. Then $df_A \left(\frac{\partial}{\partial\gamma}\right)^A = 0$ and the tangent vector associated with df_A must be perpendicular (have zero dot product) with all of the tangent vectors which lie within the level surface of f. This is the usual gradient vector as an arrow that you have learned about in previous course. Ie, the gradient, as a cotangent vector is defined even in the most primative structure of the theory, but the association of a tangent vector (an arrow) with the gradient requires the existence of the metric.

We note that this can lead to some very strange situation. We will find that there exist metrics (eg the special relativisitic metric) such that a vector can be perpendicular to itself (ie have zero length). This means that the gradient vector, regarded as an arrow, can be a tangent vector which lies within the level surface itself. Ie, a tangent vector can both be tangent to the surface (ie to a curve which lies in the surface) and at the same time be perpendicular to the surface (have zero dot product with all tangent vectors to the surface).

0.7 Notation:

As with most physicists, I am lazy. I do not want to write additional symbolism when more compact will do. Thus if a have a function of coordinates $f(\{x^i\})$, instead of writing the partial derivative with respect to x^k as $\frac{\partial f}{\partial x^k}$, I will use the more compact notation

$$\partial_k f \equiv \frac{\partial f}{\partial x^k} \tag{63}$$

And sometimes I will use an even more compact notiation

$$f_{,k} \equiv \partial_k f \tag{64}$$

this being even simpler to write. Of course it can also be confusing if you are not used to the notation.

For the metric, if we define the length of a curve by $s(\lambda)$ we know that the length of a tangent vector with components $\frac{dx^i}{d\lambda}$ as

$$\frac{ds}{d\lambda} = \sum_{ij} g_{ij} \frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda} \tag{65}$$

To specify the metric we could write out the above in detail. Since the value of λ is irrelevant, one often simply removes all of the $d\lambda$ and writes the metric as

$$ds^2 = \sum_{ij} g_{ij} dx^i dx^j \tag{66}$$

One thing to be careful of is to remember that since the metric is symmetric, there will be two terms multiplying each $dx^i dx^j$ if i and j are not equal. Thus

$$ds^2 = (1 - \frac{2M}{r})du^2 + 2dudr - r^2(d\theta^2 + \sin(\theta)^2 d\phi^2) \tag{67}$$

has components of the metric given by

$$g_{uu} = (1 - \frac{2M}{r}) \tag{68}$$

$$g_{ur} = g_{ru} = 1 \tag{69}$$

$$g_{\theta\theta} = r^2 \tag{70}$$

$$g_{\phi\phi} = r^2 sin(\theta)^2 \tag{71}$$

and all the rest of the components being 0. Note that g_{ur} is 1, not 2. Einstein himself messed up in one of his notebooks, and confused himself for a year (thinking that the flat spacetime metric in rotating coordinates was not a solution of his equations) because he forgot this.