

How Gravitational Forces arise from Curvature

Edmund Bertschinger, edbert@mit.edu

February 22, 2005

The Principle of Relativity requires that the laws of nature shall be formulated in an ‘invariant’ fashion, i.e. independently of any special frame of reference. The methods of the calculus of variations automatically satisfy this principle, because the minimum of a scalar quantity does not depend on the coordinates in which that quantity is measured.

— Cornelius Lanczos, *The Variational Principles of Mechanics* (Dover Publications, New York, 1986). page xxviii.

1. Introduction: Extremal Aging and the Equivalence Principle

These notes supplement Chapter 3 of EBH (*Exploring Black Holes* by Taylor and Wheeler). They elaborate on the discussion of the Principle of Extremal Aging and the motion of massive bodies in curved spacetime.

In the second set of supplementary notes — *Gravity, Metrics and Coordinates* — we introduced the idea that matter induces spacetime curvature, and we argued that this curvature is fully described by the metric. In the case of a static, spherically symmetric mass distribution, we presented the Einstein Field Equations whose solution gives the metric and therefore the geometry of spacetime. The current notes complete the other half of our introduction to general relativity by deriving the equations of motion of massive bodies falling freely in a gravitational field.

The basic idea was introduced in the first set of notes, *Coordinates and Proper Time*, where it was shown that, in the absence of gravity, the worldline of a massive particle maximizes the total proper (or wristwatch) time $\int_A^B d\tau$ between fixed events A and B on a spacetime diagram. In the flat spacetime of special relativity this leads to a straight-line path in a Minkowski diagram. Thus, a twin who stays at home (or moves at constant velocity in any inertial reference frame) ages more than a twin who leaves home on a rocket flight and returns at a later time. The unaccelerated path has maximal proper time of all paths that begin and end at fixed events.¹

What then is the path of a freely-falling body in the presence of gravity? According to Newtonian physics, the answer is given by solving $\vec{F} = m\vec{a} = md^2\vec{x}/dt^2$ as a differential equation for $\vec{x}(t)$. Gravity causes acceleration, and one might expect therefore that the maximal-aging result breaks down. A body obviously accelerates in a gravitational field, and so an unaccelerated path cannot

¹Maxima or saddle points may also give allowed paths, hence the name Principle of Extremal Aging.

be the correct one. However, according to general relativity it still works out that the correct path is the one that maximizes proper time!

It seems astonishing that a result from special relativity carries over directly to general relativity without modification. The key is that, in the paradigm of general relativity, free-fall motion arises not from acceleration but from the effects of spacetime curvature. As we will see, the appearance of acceleration arises naturally from extremal paths in a curved spacetime.

We say “appearance of acceleration” because ordinary acceleration depends on the motion of one’s reference frame. In an inertial reference frame in Newtonian gravity, a body moves at a constant velocity if no forces act on it. In Newtonian theory, an inertial reference frame can be extended over all of spacetime. But we have already argued in the first set of notes that there are no *global* inertial reference frames in curved spacetime.² Consequently the notion of acceleration is ambiguous! Acceleration depends on frame, and if there are no preferred frames, there is no preferred concept of acceleration.

The ambiguity of acceleration is encapsulated in Einstein’s famous Equivalence Principle:³

There is no way to distinguish between the effects on an observer of a uniform gravitational field and of constant acceleration.

The meaning of this statement is shown by a famous thought experiment proposed by Einstein. Consider a person in an elevator whose cable has broken and is accelerating at 9.8 ms^{-2} toward the center of the earth. The person is weightless and floats in the middle of the elevator with no acceleration relative to the elevator. (This follows from the Newtonian equivalence of gravitational and inertial mass discussed in the notes *Coordinates and Proper Time* — recall that all bodies fall the same way in a gravitational field.) Locally, that is over sufficiently small intervals of time and space, there is no way for the person in the elevator to determine whether she is in an elevator falling in a gravitational field or floating freely in an unaccelerated spaceship far from all sources of gravity.

We know already that, in the absence of gravity, unaccelerated bodies take the path that maximizes proper time. Arguing from the Equivalence Principle, Einstein proposed that the same “extremal aging” principle holds in curved spacetime. Locally there is no difference between free-fall and unaccelerated motion. Locally, bodies must therefore follow paths that maximize proper time. If a long path is divided into many short segments, each of which maximizes proper time, then the total path maximizes proper time. The effects of gravity must be represented not by local acceleration but rather by global curvature.

²One might ask whether it is possible to have a relativistic theory of gravitation in *flat* spacetime, where global inertial frames exist. No successful theory of this type has been proposed. General relativity requires *curved* spacetime.

³This is the Weak Equivalence Principle. There is a more general Strong Equivalence Principle that we will not discuss here.

This story seems too simple to be true. However, the Equivalence Principle is one of the best tested laws of physics. As far as we can tell, Nature obeys Einstein’s Equivalence Principle and the Principle of Extremal Aging that follows from it.

The rest of these notes present the mathematics of Extremal Aging in curved spacetime. The starting point is the metric, which we can write in general as

$$d\tau^2 = \sum_{\mu=0}^3 \sum_{\nu=0}^3 g_{\mu\nu}(x) dx^\mu dx^\nu . \quad (1)$$

We group together the four coordinates into the set $\{x^\mu\}$ where the superscript μ is an index that labels the coordinate, and not an exponent. For example, with spherical coordinates we may write $x^\mu \in \{t, r, \theta, \phi\}$. The notation is really a detail, however. What is important is that the proper time depends quadratically on the coordinate differentials. The 4×4 matrix $[g_{\mu\nu}]$ is called the matrix of metric coefficients, often abbreviated to the metric. In special relativity, all the $g_{\mu\nu}$ vanish except $g_{00} = -g_{11} = -g_{22} = -g_{33} = 1$. In general, however, any or all of the $g_{\mu\nu}$ can be nonzero and can depend on the coordinates, so we write $g_{\mu\nu} = g_{\mu\nu}(x)$. For examples, see equations (4), (7)–(9), (17), and (21) of the notes *Gravity, Metrics and Coordinates*.

The Principle of Extremal Aging may now be stated mathematically. Let $x^\mu(\lambda)$ be a path in spacetime where λ is a parameter which uniquely labels each point along the curve. Of all paths that begin at x_A^μ and end at x_B^μ , a freely falling particle takes the path that extremizes the proper time

$$\tau_{AB} = \int_A^B d\tau = \int_A^B \left(\sum_{\mu=0}^3 \sum_{\nu=0}^3 g_{\mu\nu}(x) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \right)^{1/2} d\lambda . \quad (2)$$

The events A and B are fixed. The physical path is the one passing through these events that extremizes the integral τ_{AB} .

Although the concept of extremal path seems abstract when presented in equation form, by reviewing Section 5 of the notes *Coordinates and Proper Time* you will see that it has a simple geometric meaning. In curved spacetime, unfortunately, that meaning cannot be represented so easily as in flat spacetime. Here we have to do a little more work to extremize τ_{AB} than we did in the first set of notes. The following sections are devoted to this topic.

2. Calculus of Variations

The mathematical problem we have stated is a classical problem that was addressed in the eighteenth century by the mathematicians Euler, Lagrange, and others. The presentation here follows Chapter 6 of *Classical Dynamics of Particles and Systems* by Marion and Thornton (4th edition, Harcourt Brace & Co.). We first present the mathematics for one function of a single variable, which would apply only to a trajectory in one dimension. We then extend the treatment

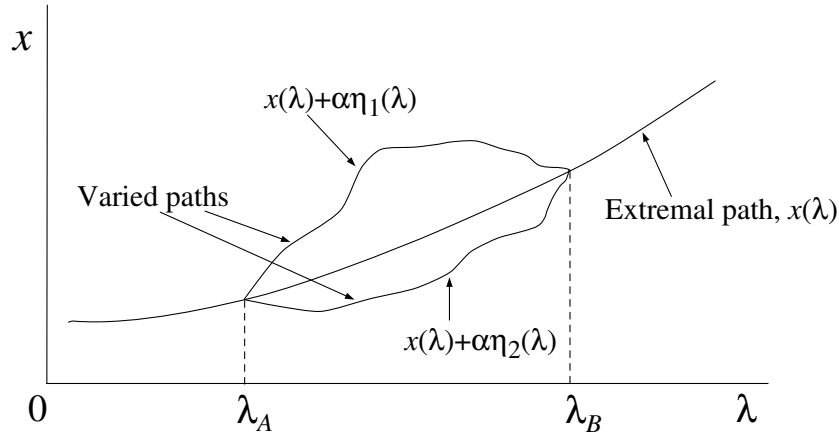


Fig. 1.— The extremal path and variations around it.

to multiple dimensions, as needed for the application to general relativity. The treatment in this section is quite formal; you are not expected to understand everything at once. The examples will make clear how the method works in practice.

The calculus of variations is a method for determining the function $x(\lambda)$ such that the integral

$$S = \int_{\lambda_A}^{\lambda_B} f[x(\lambda), \dot{x}(\lambda); \lambda] d\lambda \quad (3)$$

is a maximum or a minimum (i.e. an extremum). Here, x is the dependent variable, λ the independent variable (a dummy variable of integration often called t), and $\dot{x} \equiv dx/d\lambda$. The function $f(x, \dot{x}; \lambda)$ is assumed to be given — it is an input that the user of calculus of variations must provide. (Our goal will be to use the integrand of eq. 2 but for now we do not specify f .) The semicolon is purely a matter of convention; f can be a function of all three variables. Most importantly, both λ and $x(\lambda)$ are fixed at the endpoints. Thus, the solution $x(\lambda)$ is required to satisfy $x(\lambda_A) = x_A$ and $x(\lambda_B) = x_B$ where (λ_A, λ_B) and (x_A, x_B) are all specified. The problem is simply to find the function $x(\lambda)$ at all intermediate points $\lambda_A < \lambda < \lambda_B$ such that S is extremal.

Figure 1 shows the idea of the calculus of variations. The extremal path is $x(\lambda)$. Two different variations are shown, using different choices for the function $\eta(\lambda)$, $\eta_1(\lambda)$ and $\eta_2(\lambda)$. Notice that the endpoints are fixed.

What do we mean by extremal? To answer this, let us consider the following function:

$$x(\alpha, \lambda) = x(0, \lambda) + \alpha\eta(\lambda) \quad (4)$$

where $\eta(\lambda)$ is some function that has a continuous first derivative and that vanishes at λ_A and λ_B . With these boundary conditions $x(\alpha, \lambda)$ will satisfy the correct endpoint conditions if $x(0, \lambda_A) = x_A$ and $x(0, \lambda_B) = x_B$.

By varying α , we change x and \dot{x} for a given λ and therefore we change the value of the integral S in equation (3). Therefore $S = S(\alpha)$. We have converted S into an ordinary function⁴ of α . The extremal condition is now the usual condition on the derivative of a function: $dS/d\alpha = 0$. This ensures that S is an extremum for a given function $\eta(\lambda)$, but that is not enough. We require $S(\alpha)$ to be an extremum at the same value of α for *any function* $\eta(\lambda)$ with $\eta(\lambda_A) = \eta(\lambda_B) = 0$.

Now we proceed to impose the extremal condition $dS/d\alpha = 0$, using equations (3) and (4):

$$\frac{dS}{d\alpha} = \frac{d}{d\alpha} \int_{\lambda_A}^{\lambda_B} f(x, \dot{x}; \lambda) d\lambda . \quad (5)$$

The limits of integration are fixed; α appears only inside f through equation (4). Using the chain rule gives

$$\frac{dS}{d\alpha} = \int_{\lambda_A}^{\lambda_B} \left(\frac{\partial f}{\partial x} \frac{\partial x}{\partial \alpha} + \frac{\partial f}{\partial \dot{x}} \frac{\partial \dot{x}}{\partial \alpha} \right) d\lambda . \quad (6)$$

From equation (4),

$$\frac{\partial x}{\partial \alpha} = \eta(\lambda) , \quad \frac{\partial \dot{x}}{\partial \alpha} = \frac{d\eta}{d\lambda} . \quad (7)$$

Substituting into equation (6) gives

$$\frac{dS}{d\alpha} = \int_{\lambda_A}^{\lambda_B} \left[\frac{\partial f}{\partial x} \eta(\lambda) + \frac{\partial f}{\partial \dot{x}} \frac{d\eta}{d\lambda} \right] d\lambda . \quad (8)$$

The second term in equation (8) can be integrated by parts using

$$\left(\frac{\partial f}{\partial \dot{x}} \right) \frac{d\eta}{d\lambda} = \frac{d}{d\lambda} \left[\left(\frac{\partial f}{\partial \dot{x}} \right) \eta(\lambda) \right] - \eta(\lambda) \frac{d}{d\lambda} \left(\frac{\partial f}{\partial \dot{x}} \right) ,$$

giving

$$\frac{dS}{d\alpha} = \int_{\lambda_A}^{\lambda_B} \left[\frac{\partial f}{\partial x} \eta(\lambda) - \eta(\lambda) \frac{d}{d\lambda} \left(\frac{\partial f}{\partial \dot{x}} \right) \right] d\lambda + \left[\frac{\partial f}{\partial \dot{x}} \eta(\lambda) \right]_{\lambda_A}^{\lambda_B} . \quad (9)$$

Now we impose the condition of fixed endpoints, $\eta(\lambda_A) = \eta(\lambda_B) = 0$ and get

$$\frac{dS}{d\alpha} = \int_{\lambda_A}^{\lambda_B} \left[\frac{\partial f}{\partial x} - \frac{d}{d\lambda} \left(\frac{\partial f}{\partial \dot{x}} \right) \right] \eta(\lambda) d\lambda . \quad (10)$$

We require that $dS/d\alpha = 0$ for *arbitrary* $\eta(\lambda)$ as long as $\eta(\lambda_A) = \eta(\lambda_B) = 0$. In particular, we could take $\eta(\lambda)$ to be zero everywhere except in the immediate neighborhood of some point λ_0 ,

⁴This function also depends on the functional form of $\eta(\lambda)$. It is assumed that the function $\eta(\lambda)$ does not change when α is varied.

where it is made so large that the integral would be nonzero unless the quantity in square brackets vanishes at λ_0 . But λ_0 could be any point in the range $[\lambda_A, \lambda_B]$. We conclude that the only way for the path to be extremal under all circumstances is that the path obey the Euler-Lagrange equation

$$\frac{d}{d\lambda} \left(\frac{\partial f}{\partial \dot{x}} \right) - \frac{\partial f}{\partial x} = 0 . \quad (11)$$

This is a very important result. Having derived it, we no longer need to bother about α and $\eta(\lambda)$ — equation (11) is now simply a second-order differential equation for $x(\lambda)$. A geometric principle — requiring that an integral be extremized — is mathematically equivalent to a differential equation. Ah, the beauty of the calculus of variations!

Before applying the Euler-Lagrange equation we need to generalize it to include more than one dependent variable because spacetime has more than one dimension. Suppose that our integral depends on more than one dependent variable, e.g. $f = f[t, r, \theta, \phi, dt/d\lambda, dr/d\lambda, d\theta/d\lambda, d\phi/d\lambda; \lambda]$ or, in the brief notation introduced at the end of Section 2, $f = f[x^\mu(\lambda), dx^\mu/d\lambda; \lambda]$. Now the variation of the trajectory requires several different functions $\eta(\lambda)$, one for each dependent variable: $x^\mu(\alpha, \lambda) = x^\mu(0, \lambda) + \alpha\eta^\mu(\lambda)$.⁵ Repeating the above derivation leads to a very simple generalization of equation (10):

$$\frac{dS}{d\alpha} = \int_{\lambda_A}^{\lambda_B} \sum_{\mu} \left\{ \frac{\partial f}{\partial x^\mu} - \frac{d}{d\lambda} \left[\frac{\partial f}{\partial (dx^\mu/d\lambda)} \right] \right\} \eta^\mu(\lambda) d\lambda . \quad (12)$$

(To avoid any confusion, we wrote out $\dot{x}^\mu = dx^\mu/d\lambda$ explicitly.) Now all of the variations $\eta^\mu(\lambda)$ are arbitrary, requiring that each term in the braces vanish:

$$\frac{d}{d\lambda} \left[\frac{\partial f}{\partial (dx^\mu/d\lambda)} \right] - \frac{\partial f}{\partial x^\mu} = 0 . \quad (13)$$

We have derived a powerful tool. Applying it requires only that we specify the function f giving the integral we wish to extremize. We illustrate this with a geometric example before applying it to general relativity.

3. Example: Geodesics in the Plane

A straight line is the shortest distance between two points. Can you prove that commonplace of Euclidean space?

The proof is easy using the calculus of variations. The distance measured along a curve between two points is simply

$$L = \int_A^B \sqrt{dx^2 + dy^2} = \int_A^B \left[\left(\frac{dx}{d\lambda} \right)^2 + \left(\frac{dy}{d\lambda} \right)^2 \right]^{1/2} d\lambda . \quad (14)$$

⁵Recall that μ in x^μ and η^μ is an index labelling one of a set of values, not an exponent.

In the second form we note explicitly that the path between A and B is a parameterized curve $[x(\lambda), y(\lambda)]$. This is exactly in the form needed for the calculus of variations, with

$$f \left[x, y, \frac{dx}{d\lambda}, \frac{dy}{d\lambda}; \lambda \right] = \left[\left(\frac{dx}{d\lambda} \right)^2 + \left(\frac{dy}{d\lambda} \right)^2 \right]^{1/2}. \quad (15)$$

Note that f depends on only two of its possible arguments, $dx/d\lambda$ and $dy/d\lambda$.

The condition that the path be extremal is simply equation (13). We need the following derivatives of f from equation (15):

$$\frac{\partial f}{\partial(dx/d\lambda)} = \frac{1}{f} \frac{dx}{d\lambda}, \quad \frac{\partial f}{\partial(dy/d\lambda)} = \frac{1}{f} \frac{dy}{d\lambda}. \quad (16)$$

Substituting into equation (13), we get

$$\frac{d}{d\lambda} \left(\frac{1}{f} \frac{dx}{d\lambda} \right) = 0, \quad \frac{d}{d\lambda} \left(\frac{1}{f} \frac{dy}{d\lambda} \right) = 0, \quad (17)$$

which may be immediately integrated to give

$$\frac{1}{f} \frac{dx}{d\lambda} = c_x, \quad \frac{1}{f} \frac{dy}{d\lambda} = c_y, \quad (18)$$

where c_x and c_y are constants. Substituting back into equation (15), we find that c_x and c_y are not independent but must satisfy $c_x^2 + c_y^2 = 1$.

Equations (18) are still differential equations for the path $\{x(\lambda), y(\lambda)\}$, and we can integrate them by taking the ratio:

$$\frac{dy/d\lambda}{dx/d\lambda} = \frac{dy}{dx} = \frac{c_y}{c_x}. \quad (19)$$

Because this ratio is a constant, we can now regard y as being a function of x and integrate at once to arrive at the familiar form for the equation of a line,

$$y = ax + b, \quad a \equiv \frac{c_y}{c_x}. \quad (20)$$

Here, b is another constant of integration. Note that the condition $c_x^2 + c_y^2 = 1$ implies (if c_x and c_y are real) that we can always write $c_x = \cos \alpha$ and $c_y = \sin \alpha$ for some real angle α , implying $c_y/c_x = \tan \alpha$. This result allows the slope a to take any real value, $-\infty < a < \infty$.

Notice that λ does not appear explicitly in our solution. In fact, the coordinates may *any* functions of λ , as long as $y(\lambda) = ax(\lambda) + b$. We introduced λ as a parameter along the curve in order to write the path length as an integral over λ , but once we have the general solution for $y(x)$, the parameterization $\{x(\lambda), y(\lambda)\}$ is no longer necessary.

Thus, the calculus of variations has allowed us to derive the equation of a line from the geometric condition of the path of shortest (or extremal) distance between two points. Extremal paths are also called geodesics. Geodesics in the plane are straight lines.

4. Example: Schwarzschild Spacetime

Our main interest lies in the metric of a non-rotating black hole, which in Schwarzschild coordinates in the equatorial plane takes the form

$$d\tau^2 = A(r)dt^2 - A(r)^{-1}dr^2 - r^2d\phi^2, \quad A(r) \equiv 1 - \frac{2GM}{r}. \quad (21)$$

Now the integrand of the proper time is

$$f(t, r, \phi, dt/d\lambda, dr/d\lambda, d\phi/d\lambda; \lambda) = \left[A(r) \left(\frac{dt}{d\lambda} \right)^2 - A(r)^{-1} \left(\frac{dr}{d\lambda} \right)^2 - r^2 \left(\frac{d\phi}{d\lambda} \right)^2 \right]^{1/2}. \quad (22)$$

The procedure now is the same as in the preceding section. First we need the derivatives of f :

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial \phi} = 0, \quad \frac{\partial f}{\partial r} = \frac{1}{f} \left(\frac{1}{2} \frac{dA}{dr} \right) \left[\left(\frac{dt}{d\lambda} \right)^2 + \frac{1}{A^2} \left(\frac{dr}{d\lambda} \right)^2 \right] - \frac{r}{f} \left(\frac{d\phi}{d\lambda} \right)^2, \quad (23)$$

and

$$\frac{\partial f}{\partial(dt/d\lambda)} = \frac{A}{f} \frac{dt}{d\lambda}, \quad \frac{\partial f}{\partial(dr/d\lambda)} = -\frac{1}{Af} \frac{dr}{d\lambda}, \quad \frac{\partial f}{\partial(d\phi/d\lambda)} = -\frac{r^2}{f} \frac{d\phi}{d\lambda}. \quad (24)$$

Substituting into equation (13), we get three equations of motion:

$$\frac{d}{d\lambda} \left(\frac{A}{f} \frac{dt}{d\lambda} \right) = 0, \quad (25a)$$

$$\frac{d}{d\lambda} \left(-\frac{1}{Af} \frac{dr}{d\lambda} \right) = 0, \quad (25b)$$

$$\frac{d}{d\lambda} \left(-\frac{r^2}{f} \frac{d\phi}{d\lambda} \right) + \frac{1}{f} \left(\frac{1}{2} \frac{dA}{dr} \right) \left[\left(\frac{dt}{d\lambda} \right)^2 + \frac{1}{A^2} \left(\frac{dr}{d\lambda} \right)^2 \right] - \frac{r}{f} \left(\frac{d\phi}{d\lambda} \right)^2 = 0. \quad (25c)$$

Although the third equation is complicated, the first two may be integrated at once to give

$$\frac{A}{f} \frac{dt}{d\lambda} = C_1, \quad -\frac{r^2}{f} \frac{d\phi}{d\lambda} = C_2, \quad (26)$$

where C_1 and C_2 are constants along the path.

5. Constants of Motion

The results obtained for the Schwarzschild metric are more complicated than what we obtained for geodesics in the Euclidean plane. One might think the best way to proceed now is by analogy: take the ratio of the two parts of equations (26) and integrate. However, integration this way cannot succeed because of the factors $A(r)$ and r^2 , which are some unknown functions of λ .

One way to proceed would be simply to regard equations (25) as a set of three second-order differential equations for the trajectory $\{t(\lambda), r(\lambda), \phi(\lambda)\}$. However, we can simplify matters by using a trick that always works for the proper time integral, equation (2). Given any solution to equations (25), we may construct another by changing the parameterization of the curve, writing $\lambda = \lambda(\eta)$ with any monotonic function. For example, consider the equation of a line $y = 2x + 3$. This is obeyed equally well with $x(\lambda) = \lambda^2$ and $y(\lambda) = 2\lambda^2 + 3$, as it is with $x(\lambda) = \exp(\lambda)$ and $y(\lambda) = 2\exp(\lambda) + 3$. Both curves give the same line even though they are different solutions for $\{x(\lambda), y(\lambda)\}$. They are related by the simple transformation $\lambda = \exp(\eta/2)$.

From equation (2), it is easy to see that if the independent variable λ is changed to η using $d\lambda = h(\eta)d\eta$ for some positive function $h(\eta)$, the integral τ_{AB} does not change at all:⁶

$$\int_A^B \left(\sum_{\mu=0}^3 \sum_{\nu=0}^3 g_{\mu\nu}(x) \frac{1}{h^2} \frac{dx^\mu}{d\eta} \frac{dx^\nu}{d\eta} \right)^{1/2} h(\eta) d\eta = \int_A^B \left(\sum_{\mu=0}^3 \sum_{\nu=0}^3 g_{\mu\nu}(x) \frac{dx^\mu}{d\eta} \frac{dx^\nu}{d\eta} \right)^{1/2} d\eta . \quad (27)$$

But η is just a dummy variable of integration and we are free to change its name back to λ without changing the integral. We can take advantage of this freedom to choose λ to be any monotonic parameter along the curve.

For massive bodies, there is an obvious parameter that increases monotonically along the worldline of a body: the proper time τ ! We can therefore choose $\lambda = \tau$. Now, since $\tau_{AB} = \int_A^B f d\lambda$, it follows that we can set $f = 1$ in equations (25), and rename⁷ $\lambda \rightarrow \tau$.

This simplification allows us to rewrite equations (26) in a simpler, more physical form:

$$\left(1 - \frac{2GM}{r} \right) \frac{dt}{d\tau} = \frac{E}{m} , \quad r^2 \frac{d\phi}{d\tau} = \frac{L}{m} . \quad (28)$$

We have changed the names of the constants of integration by introducing the particle rest mass m , the energy at infinity E , and the angular momentum L . Their significance is the following: E/m and L/m are constants along the path of any free-falling body in the Schwarzschild spacetime. We have derived EBH equations [3.12] and [4.2] by using the calculus of variations.

Using the conservation of E and L , we can work out the trajectories of massive bodies. That is the subject of Chapters 3 and 4 of EBH.

⁶This argument relies on the fact that there is no explicit dependence on λ inside the integral; only $d\lambda$ appears.

⁷For massless particles, we cannot set λ to the proper time, because $d\tau = 0$ in this case. However, we can set $f = 1$ for any finite rest mass no matter how small, and take the limit as the rest mass goes to zero. For massless particles we can therefore set $f = 1$ but we call the parameter λ since it is not proper time.