Schwarzschild Solution and Black Holes

Asaf Pe'er[1]

June 7, 2020

This part of the course is based on Refs. [1], [2], [3] and [4]. Figures 1, 10 - 14 are taken from Sean Carroll's notes in "Level 5: A Knowledgebase for Extragalactic Astronomy and Cosmology" (http://ned.ipac.caltech.edu/level5/March01/Carroll3/frames.html). Figures 4, 5, 8 and 9 are taken from J. Hartle, *Gravity: An Introduction to Einstein's General Relativity* (Ref. [2]).

1. Background

We now turn to look for solutions to Einstein's equation. We first discuss solutions in vacuum - as is the case outside stars (e.g., astronauts orbiting earth, etc). In vacuum, Einstein's equations become $R_{\mu\nu} = 0$. With the possible exception of Minkowski space, by far the most important solution in this case is that discovered by **Karl Schwarzschild**, which describes spherically symmetric vacuum spacetimes. At a later part of the course, we will discuss a different type of solution, for a non-empty space. These solutions will be relevant for the study of the entire universe.

As Einstein's equations are non-linear, the difficult part is to obtain **a** solution. Once we think we found one, we can simply plug it into the equations to verify it. Fortunately, Schwarzschild did the hard work by finding a solution. In fact, there is a theorem, known as **Birkhoff's theorem** that states that the Schwarzschild solution is the *unique* spherically symmetric solution to Einstein's equations in vacuum (I don't really see time for proving it in class).

2. Obtaining a Solution: Derivation of the Schwarzschild Metric

We are looking for a metric tensor representing a **static and isotropic** gravitational field. Being *isotropic*, it is best described in polar coordinates, in which

$$d\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta \; \mathrm{d}\phi^2 \; . \qquad (1)$$

[1]Department of Physics, Bar Ilan University

Being in four-dimensional spacetime, we need two more coordinates. The most general form is to choose arbitrary coordinates $t' = a$ and $r' = b$, and write the metric as

$$ds^2 = g_{aa}(a,b)\mathrm{d}a^2 + g_{ab}(a,b)(\mathrm{d}a\mathrm{d}b + \mathrm{d}b\mathrm{d}a) + g_{bb}(a,b)\mathrm{d}b^2 + r^2(a,b)d\Omega^2 \ . \tag{2}$$

Note the following: (1) we preferred to use $a = t'$ and $b = r'$, since there are cross terms $(\mathrm{d}a\mathrm{d}b + \mathrm{d}b\mathrm{d}a)$ which we would like to get rid of later by changing coordinates; (2) At this point $r(a,b)$ is some as-yet-undetermined function, to which we have merely given a suggestive label. It will be identified with the usual definition of r later.

We can change coordinates from (a,b) to (a,r), by inverting $r(a,b)$. The metric is then

$$ds^2 = g_{aa}(a,r)\mathrm{d}a^2 + g_{ar}(a,r)(\mathrm{d}a\mathrm{d}r + \mathrm{d}r\mathrm{d}a) + g_{rr}(a,r)\mathrm{d}r^2 + r^2 d\Omega^2 \ . \tag{3}$$

Now we want to find a function $t(a,r)$ such that, in the (t,r) coordinate system, there are no cross terms $\mathrm{d}t\mathrm{d}r + \mathrm{d}r\mathrm{d}t$ in the metric. Since

$$\mathrm{d}t = \frac{\partial t}{\partial a}\mathrm{d}a + \frac{\partial t}{\partial r}\mathrm{d}r \ , \tag{4}$$

we find

$$\mathrm{d}t^2 = \left(\frac{\partial t}{\partial a}\right)^2 \mathrm{d}a^2 + \left(\frac{\partial t}{\partial a}\right)\left(\frac{\partial t}{\partial r}\right)(\mathrm{d}a\mathrm{d}r + \mathrm{d}r\mathrm{d}a) + \left(\frac{\partial t}{\partial r}\right)^2 \mathrm{d}r^2 \ . \tag{5}$$

We would like to remove the cross-terms in the metric (Equation 3), namely replace the first three terms by

$$m\mathrm{d}t^2 + n\mathrm{d}r^2 \ , \tag{6}$$

for some functions m and n. This is equivalent to the requirements

$$m\left(\frac{\partial t}{\partial a}\right)^2 = g_{aa} \ , \tag{7}$$

$$n + m\left(\frac{\partial t}{\partial r}\right)^2 = g_{rr} \ , \tag{8}$$

and

$$m\left(\frac{\partial t}{\partial a}\right)\left(\frac{\partial t}{\partial r}\right) = g_{ar} \ . \tag{9}$$

We therefore have three equations for the three unknowns $t(a,r)$, $m(a,r)$, and $n(a,r)$, just enough to determine them precisely (up to initial conditions for t). (Of course, they are "determined" in terms of the unknown functions g_{aa}, g_{ar}, and g_{rr}, so in this sense they are still undetermined.)

What we have just shown is that we can put the general (static and isotropic) metric in Equation 2 in the form

$$ds^2 = m(t,r)\mathrm{d}t^2 + n(t,r)\mathrm{d}r^2 + r^2 d\Omega^2 \ . \tag{10}$$

Note that we treat t and r in a symmetric way; the only difference is that we have chosen r to be the one which multiplies the metric for the two-sphere. This is of course motivated by what we know about the metric for flat Minkowski space, which can be written $ds^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2 d\Omega^2$.

Asymptotically ($r \to \infty$), we expect the metric to approach Minkowski metric; this, the metric must have one negative and three positive components (this is known as **pseudo-Riemannian**, as opposed to a metric which have only positive components, known as *Euclidean* or *Riemannian*) . Thus, we choose $m(t,r)$ to be negative. With this choice, we can replace the functions m and n with new functions α and β, and write the metric as

$$ds^2 = -e^{2\alpha(t,r)}\mathrm{d}t^2 + e^{2\beta(t,r)}\mathrm{d}r^2 + r^2 d\Omega^2 \ . \tag{11}$$

This is the best we can do for a general metric in a spherically symmetric spacetime. In order to proceed (determine the values of $\alpha(t,r)$ and $\beta(t,r)$), we need to actually solve Einstein's Equation.

Using the labels $(0,1,2,3)$ for (t,r,θ,ϕ) in the usual way, the Christoffel symbols are given by

$$\begin{array}{ccc} \Gamma^0_{00} = \partial_0\alpha & \Gamma^0_{01} = \partial_1\alpha & \Gamma^0_{11} = e^{2(\beta-\alpha)}\partial_0\beta \\ \Gamma^1_{00} = e^{2(\alpha-\beta)}\partial_1\alpha & \Gamma^1_{01} = \partial_0\beta & \Gamma^1_{11} = \partial_1\beta \\ \Gamma^2_{12} = \frac{1}{r} & \Gamma^1_{22} = -re^{-2\beta} & \Gamma^3_{13} = \frac{1}{r} \\ \Gamma^1_{33} = -re^{-2\beta}\sin^2\theta & \Gamma^2_{33} = -\sin\theta\cos\theta & \Gamma^3_{23} = \frac{\cos\theta}{\sin\theta} \ . \end{array} \tag{12}$$

(All the other coefficients are either zero, or related to the above by symmetries.) From these we get the following non-vanishing components of the Riemann tensor:

$$\begin{array}{rcl} R^0{}_{101} & = & e^{2(\beta-\alpha)}[\partial_0^2\beta + (\partial_0\beta)^2 - \partial_0\alpha\partial_0\beta] + [\partial_1\alpha\partial_1\beta - \partial_1^2\alpha - (\partial_1\alpha)^2] \\ R^0{}_{202} & = & -re^{-2\beta}\partial_1\alpha \\ R^0{}_{303} & = & -re^{-2\beta}\sin^2\theta\ \partial_1\alpha \\ R^0{}_{212} & = & -re^{-2\alpha}\partial_0\beta \\ R^0{}_{313} & = & -re^{-2\alpha}\sin^2\theta\ \partial_0\beta \\ R^1{}_{212} & = & re^{-2\beta}\partial_1\beta \\ R^1{}_{313} & = & re^{-2\beta}\sin^2\theta\ \partial_1\beta \\ R^2{}_{323} & = & (1-e^{-2\beta})\sin^2\theta \ . \end{array} \tag{13}$$

Taking the contraction as usual yields the Ricci tensor:

$$\begin{aligned}
R_{00} &= [\partial_0^2\beta + (\partial_0\beta)^2 - \partial_0\alpha\partial_0\beta] + e^{2(\alpha-\beta)}[\partial_1^2\alpha + (\partial_1\alpha)^2 - \partial_1\alpha\partial_1\beta + \tfrac{2}{r}\partial_1\alpha] \\
R_{11} &= -[\partial_1^2\alpha + (\partial_1\alpha)^2 - \partial_1\alpha\partial_1\beta - \tfrac{2}{r}\partial_1\beta] + e^{2(\beta-\alpha)}[\partial_0^2\beta + (\partial_0\beta)^2 - \partial_0\alpha\partial_0\beta] \\
R_{01} &= \tfrac{2}{r}\partial_0\beta \\
R_{22} &= e^{-2\beta}[r(\partial_1\beta - \partial_1\alpha) - 1] + 1 \\
R_{33} &= R_{22}\sin^2\theta \ .
\end{aligned} \tag{14}$$

In vacuum, $R_{\mu\nu} = 0$. From $R_{01} = 0$ we get

$$\partial_0\beta = 0 \ . \tag{15}$$

We proceed by taking the time derivative of $R_{22} = 0$ and using $\partial_0\beta = 0$, we get

$$\partial_0\partial_1\alpha = 0 \ . \tag{16}$$

We can therefore write

$$\begin{aligned}
\beta &= \beta(r) \\
\alpha &= f(r) + g(t) \ .
\end{aligned} \tag{17}$$

The first term in the metric (Equation 11) is therefore $-e^{2f(r)}e^{2g(t)}\mathrm{d}t^2$. But we could always simply redefine our time coordinate by replacing $\mathrm{d}t \to e^{-g(t)}\mathrm{d}t$; in other words, we are free to choose t such that $g(t) = 0$, whence $\alpha(t,r) = f(r)$. We therefore have

$$ds^2 = -e^{2\alpha(r)}\mathrm{d}t^2 + e^{2\beta(r)}\mathrm{d}r^2 + r^2 d\Omega^2 \ . \tag{18}$$

All of the metric components are independent of the coordinate t. We have therefore proven a key result: **any spherically symmetric vacuum metric possesses a timelike Killing vector.** Equation 18 is the most general metric that describes static, isotropic gravitational field.

This property is so interesting that it gets its own name: a metric which possesses a timelike Killing vector is called **stationary**. There is also a more restrictive property: a metric is called **static** if it possesses a timelike Killing vector which is orthogonal to a family of hypersurfaces. (A hypersurface in an n-dimensional manifold is simply an $(n-1)$-dimensional submanifold.)

The metric (Equation 18) is not only stationary, but also static; the Killing vector field ∂_0 is orthogonal to the surfaces $t = const$ (since there are no cross terms such as $\mathrm{d}t\mathrm{d}r$ and so on). Roughly speaking, a static metric is one in which nothing is moving, while a stationary metric allows things to move but only in a symmetric way. For example, the static spherically symmetric metric (Equation 18) will describe non-rotating stars or black holes, while rotating systems (which keep rotating in the same way at all times) will be described by stationary metrics. It's hard to remember which word goes with which concept, but the distinction between the two concepts should be understandable.

We continue as follows. Since both R_{00} and R_{11} vanish, we can write

$$0 = e^{2(\beta-\alpha)} R_{00} + R_{11} = \frac{2}{r}(\partial_1 \alpha + \partial_1 \beta) \ , \tag{19}$$

which implies $\alpha = -\beta +$ constant. Once again, we can get rid of the constant by scaling our coordinates, so we have

$$\alpha = -\beta \ . \tag{20}$$

Next we use $R_{22} = 0$, which now reads

$$e^{2\alpha}(2r\partial_1 \alpha + 1) = 1 \ . \tag{21}$$

This is completely equivalent to

$$\partial_1(re^{2\alpha}) = 1 \ . \tag{22}$$

We can solve this to obtain

$$e^{2\alpha} = 1 + \frac{\mu}{r} \ , \tag{23}$$

where μ is some undetermined constant. With Equations 20 and 23, the metric becomes

$$ds^2 = -\left(1 + \frac{\mu}{r}\right) \mathrm{d}t^2 + \left(1 + \frac{\mu}{r}\right)^{-1} \mathrm{d}r^2 + r^2 d\Omega^2 \ . \tag{24}$$

There is no more freedom left except for the single constant μ. However, fortunately, it is straightforward to check that for any value of μ this metric solves the two equations, $R_{00} = 0$ and $R_{11} = 0$.

The last thing to do is to assign a physical interpretation to the constant μ. The most important use of a spherically symmetric vacuum solution is to represent the spacetime outside a star or planet or whatnot. In that case we would expect to recover the weak field limit as $r \to \infty$. In this limit, Equation 24 implies

$$\begin{aligned} g_{00}(r \to \infty) &= -\left(1 + \tfrac{\mu}{r}\right) \ , \\ g_{rr}(r \to \infty) &= \left(1 - \tfrac{\mu}{r}\right) \ . \end{aligned} \tag{25}$$

The weak field limit, on the other hand, has

$$\begin{aligned} g_{00} &= -(1 + 2\Phi) \ , \\ g_{rr} &= (1 - 2\Phi) \ , \end{aligned} \tag{26}$$

with the potential $\Phi = -GM/r$. Therefore the metrics do agree in this limit, if we set $\mu = -2GM$.

Our final result is the celebrated **Schwarzschild metric**,

$$\boxed{ds^2 = -\left(1 - \frac{2GM}{c^2 r}\right)(c\mathrm{d}t)^2 + \left(1 - \frac{2GM}{c^2 r}\right)^{-1} \mathrm{d}r^2 + r^2 d\Omega^2 \ .} \tag{27}$$

(where we have added explicitly c^2, and of course, $d\Omega^2$ is given in Equation 1).

The Schwarzschild metric is the correct metric describing spherically symmetric vacuum solution to Einstein's equations. In principle, M here is just a parameter, which we happen to know can be interpreted as the conventional Newtonian mass that we would measure by studying orbits at large distances from the gravitating source. In the limit $M \to 0$ we recover Minkowski space, which is to be expected. Furthermore, note that the metric becomes progressively Minkowskian as we go to $r \to \infty$; this property is known as **asymptotic flatness**.

The fact that the Schwarzschild metric is not just a good solution, but is the unique spherically symmetric vacuum solution, is known as **Birkhoff's theorem**. It is interesting to note that the result is a static metric. We did not say anything about the source except that it be spherically symmetric. Specifically, we did not demand that the source itself be static; it could be, e.g., a collapsing star, as long as the collapse were symmetric.

3. Singularities of Schwarzschild Metric

Let us look at the Schwarzschild metric (Equation 27). From the form of the metric, we immediately see **two** interesting points: at $r = 0$ and $r = 2GM/c^2$, the metric coefficients become infinite. Thus, something may go wrong here.

However, the metric coefficients are coordinate-dependent quantities: we may be able to choose different coordinates, in which these coefficients would be nicely behaved. A good analogue is the description of a plane using polar coordinates: $ds^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$. At $r = 0$ the metric becomes degenerated, and the component $g^{\theta\theta} = r^{-2}$ of the inverse metric blows up – even though there is nothing unusual in the point $r = 0$ on the plane, it is really not different than any other point.

Thus, we really need a coordinate-independent way of telling whether there is something "bad" happening to the geometry. To make a long story short (this is a very long story), we are looking for scalars constructed from the curvature, such as the Ricci scalar, but we could think of other as well. If any of these scalars (not necessarily all of them !) go to infinity as we approach a point of interest, we will regard that point as a **singularity of the curvature**. (We should also check that the point is not "infinitely far away"; that is, that it can be reached by traveling a finite distance along a curve.)

In the case of the Schwarzschild metric, a direct calculation gives

$$R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma} = \frac{12G^2M^2}{r^6} \ . \tag{28}$$

This is enough to convince us that $r = 0$ **represents a real singularity**. At the other trouble spot, $r = 2GM/c^2$, you could check and see that none of the curvature invariants blows up. Thus, it may actually not represent a real singularity, but simply a bad choice of coordinate system. This could be verified if we could transform the coordinates into a different set, in which this surface is well behaved. This is indeed possible (such coordinates are known as **Kruskal coordinates**, and whether or not we discuss them depends on the time we have...).

Nonetheless, the surface $r = 2GM/c^2$ is of interest, and is known as **Schwarzschild radius**. But we should be careful about how interest: recall that the Schwarzschild solution we derived is valid *only in vacuum*, and thus it holds **outside** the surface of a star. However, if we look at the sun, for example, $M_\odot \approx 2 \times 10^{33}$ gr, and $2GM_\odot/c^2 \approx 3 \times 10^5$ cm, while the radius of the sun is $R_\odot \approx 7 \times 10^{10}$ cm, or $R_\odot \approx 5 \times 10^5 GM_\odot/c^2$. Thus, $r = 2GM_\odot$ is far inside the solar interior, where we do not expect the Schwarzschild metric to imply.

Thus, when dealing with the gravitational field of the sun, there are no singularities to deal with at all. Nevertheless, there are objects for which the full Schwarzschild metric is required. You have heard of them: these are **black holes**, to be discussed later. Thus, we will study the full properties of the Schwarzschild metric, but please have in mind that in the solar neighborhood, we need not worry about singularities.

4. Motion of Particles in Schwarzschild Geometry

4.1. The Geodesic Equations

Perhaps the easiest way to explore the consequences of the Schwarzschild metric, is to explore particle trajectories - along geodesics - in this metric.

Let us do that. The first step is to find the non-zero Christoffel symbols for the

Schwarzschild metric:

$$\begin{array}{ccc} \Gamma^1_{00} = \frac{GM}{r^3}(r-2GM) & \Gamma^1_{11} = \frac{-GM}{r(r-2GM)} & \Gamma^0_{01} = \frac{GM}{r(r-2GM)} \\ \Gamma^2_{12} = \frac{1}{r} & \Gamma^1_{22} = -(r-2GM) & \Gamma^3_{13} = \frac{1}{r} \\ \Gamma^1_{33} = -(r-2GM)\sin^2\theta & \Gamma^2_{33} = -\sin\theta\cos\theta & \Gamma^3_{23} = \frac{\cos\theta}{\sin\theta} \ . \end{array} \tag{29}$$

The geodesic equation therefore turns into the following four equations, where λ is an affine parameter:

$$\frac{d^2t}{d\lambda^2} + \frac{2GM}{r(r-2GM)}\frac{dr}{d\lambda}\frac{dt}{d\lambda} = 0 \ , \tag{30}$$

$$\begin{array}{c} \frac{d^2r}{d\lambda^2} + \frac{GM}{r^3}(r-2GM)\left(\frac{dt}{d\lambda}\right)^2 - \frac{GM}{r(r-2GM)}\left(\frac{dr}{d\lambda}\right)^2 \\ -(r-2GM)\left[\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\lambda}\right)^2\right] = 0 \ , \end{array} \tag{31}$$

$$\frac{d^2\theta}{d\lambda^2} + \frac{2}{r}\frac{d\theta}{d\lambda}\frac{dr}{d\lambda} - \sin\theta\cos\theta\left(\frac{d\phi}{d\lambda}\right)^2 = 0 \ , \tag{32}$$

and

$$\frac{d^2\phi}{d\lambda^2} + \frac{2}{r}\frac{d\phi}{d\lambda}\frac{dr}{d\lambda} + 2\frac{\cos\theta}{\sin\theta}\frac{d\theta}{d\lambda}\frac{d\phi}{d\lambda} = 0 \ . \tag{33}$$

At first sight, there does not seem to be much hope for simply solving this set of 4 coupled equations by inspection. Fortunately our task is greatly simplified by the high degree of symmetry of the Schwarzschild metric. We know that there are four Killing vectors: three for the spherical symmetry, and one for time translations. Each of these will lead to a constant of the motion for a free particle. Recall that if K^μ is a Killing vector, we know that

$$K_\mu \frac{dx^\mu}{d\lambda} = \text{constant} \ . \tag{34}$$

In addition, there is another constant of the motion that we always have for geodesics; metric compatibility implies that along the path the quantity

$$\epsilon = -g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} \tag{35}$$

is constant. (This is simply normalization of the 4-velocity: take $\lambda = \tau$ and get $g_{\mu\nu}U^\mu U^\nu = -\epsilon$, with $\epsilon = 1$ for massive particles and $\epsilon = 0$ for massless particles. We may also consider space-like geodesics, for which $\epsilon = -1$).

4.2. The Gravitational Potential

Instead of trying to solve directly the geodesic equations using the four conserved quantities associated with Killing vectors, let us first analyze the constraints.

In flat space time, the symmetries represented by the Killing vectors lead to very familiar conserved quantities: Invariance under time translations leads to *conservation of energy*, while invariance under spatial rotations leads to conservation of the three components of *angular momentum*.

Essentially the same applies to the Schwarzschild metric. We can think of the angular momentum as a three-vector with a magnitude (one component) and direction (two components). Conservation of the *direction* of angular momentum means that the particle will move in a plane. We can choose this to be the equatorial plane of our coordinate system; if the particle is not in this plane, we can rotate coordinates until it is. Thus, the two Killing vectors which lead to conservation of the direction of angular momentum imply

$$\theta = \frac{\pi}{2} \ . \tag{36}$$

(and of course $d\theta/d\lambda = 0$).

The other two Killing vectors correspond to energy and the magnitude of angular momentum. The timelike Killing vector is $K^\mu = (1, 0, 0, 0)^T$, and thus

$$K_\mu = K^\nu g_{\mu\nu} = \left(-\left(1 - \frac{2GM}{r}\right), 0, 0, 0 \right) \tag{37}$$

which gives rise to conservation of energy, since using Equation 34,

$$K_\mu \frac{dx^\mu}{d\lambda} = \left(1 - \frac{2GM}{r}\right) \frac{dt}{d\lambda} = E \ , \tag{38}$$

where E is constant of motion.

Similarly, The Killing vector whose conserved quantity is the magnitude of the angular momentum is $L = \partial_\phi$ ($L^\mu = (0, 0, 0, 1)^T$), and thus

$$L_\mu = \left(0, 0, 0, r^2 \sin^2\theta\right) \ . \tag{39}$$

Using $\sin\theta = 1$ derived from Equation 36, one finds

$$r^2 \frac{d\phi}{d\lambda} = L \ . \tag{40}$$

where L, the total angular momentum, is the second conserved quantity. (For massless particles these can be thought of as the energy and angular momentum; for massive particles they are the energy and angular momentum per unit mass of the particle.)

Further note that the constancy of the angular momentum in Equation 40 is the GR equivalent of Kepler's second law (equal areas are swept out in equal times).

Armed with this information, we can now analyze the orbits of particles in Schwarzschild metric. We begin by writing explicitly Equation 35, using Equation 36,

$$-\left(1-\frac{2GM}{r}\right)\left(\frac{dt}{d\lambda}\right)^2+\left(1-\frac{2GM}{r}\right)^{-1}\left(\frac{dr}{d\lambda}\right)^2+r^2\left(\frac{d\phi}{d\lambda}\right)^2=-\epsilon\ . \tag{41}$$

Multiply this Equation by $(1-2GM/r)$ and use the expressions for E and L (Equations 38 and 40) to write

$$-E^2+\left(\frac{dr}{d\lambda}\right)^2+\left(1-\frac{2GM}{r}\right)\left(\frac{L^2}{r^2}+\epsilon\right)=0\ . \tag{42}$$

Clearly, we have made a great progress: instead of the 4 geodesic Equations, we obtain one differential equation for $r(\lambda)$.

We can re-write Equation 42 as

$$\frac{1}{2}\left(\frac{dr}{d\lambda}\right)^2+V(r)=\frac{1}{2}E^2\ , \tag{43}$$

where

$$V(r)=\frac{1}{2}\epsilon-\epsilon\frac{GM}{r}+\frac{L^2}{2r^2}-\frac{GML^2}{r^3}\ . \tag{44}$$

Equation 43 is identical to the classical equation describing the motion of a (unit mass) particle moving in a 1-dimensional potential $V(r)$, provided its "energy" is $\frac{1}{2}E^2$. (Of course, the true energy is E, but we use this form due to the potential).

Looking at the potential (Equation 44) we see that it only differs from the Newtonian potential by the last term (note that this potential is *exact*, not a power series in $1/r$!). The first term is just a constant ($\epsilon = 1, 0$), the 2^{nd} term corresponds exactly to the Newtonian gravitational potential, and the third term is a contribution from angular momentum which takes the same form in Newtonian gravity and general relativity. It is the last term, though, which contains the GR contribution, which turns out to make a great deal of difference, especially at small r.

It is important not to get confused, though: the physical situation is quite different from a classical particle moving in one dimension. The trajectories under consideration are orbits around a star or other object (see Figure 1). The quantities of interest to us are not only $r(\lambda)$, but also $t(\lambda)$ and $\phi(\lambda)$. Nevertheless, obviously it is great help that the radial behavior reduces this to a problem which we know how to solve.

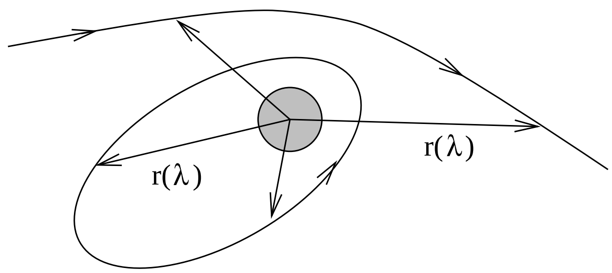

Fig. 1.— Trajectories of particles in a gravitational potential.

4.3. Massive Particles Trajectories

For massive particles, $\epsilon = 1$. The potential as a function or r (normalized to GM) for different values of the angular momentum L are shown in figures 2, 3 for the Newtonian potential (the first 3 terms in Equation 44) and for the full GR potential in Equation 44.

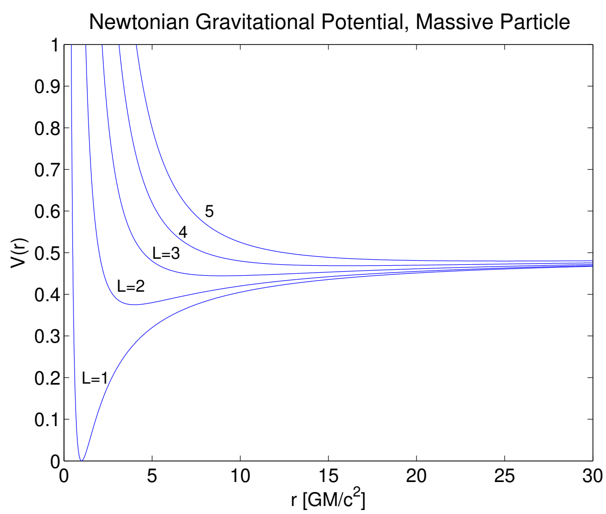

Fig. 2.— Gravitational potential for massive particles: the Newtonian case.

Clearly, for large r, the Newtonian potential and the GR potential match. It is the low r behavior that is very different.

Let us examine the kinds of possible orbits, as illustrated in the figures. There are different curves $V(r)$ for different values of L; for any one of these curves, the behavior of the orbit can be judged by comparing the $\frac{1}{2}E^2$ to $V(r)$. The general behavior of the particle will be to move in the potential until it reaches a "turning point" where $V(r) = \frac{1}{2}E^2$, where

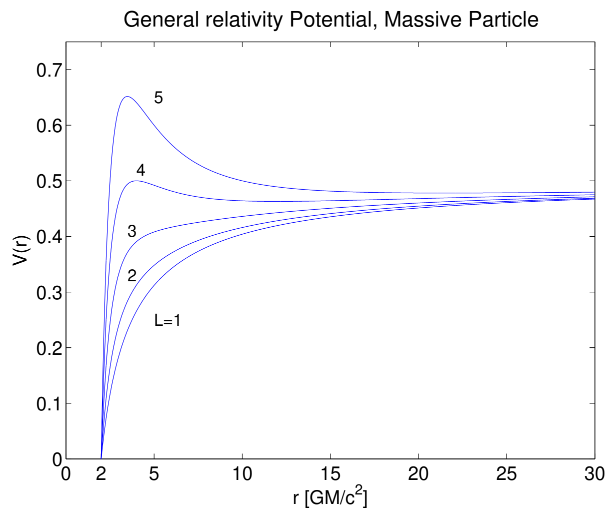

Fig. 3.— Gravitational potential for massive particles: full GR potential

it will begin moving in the other direction. Sometimes there may be no turning point to hit, in which case the particle just keeps going.

In the particular case of flat potential, $dV/dr = 0$, a particle will move in a circular orbit, at radius $r_c = const$. This radius is easily found by differentiating the potential (Equation 44),

$$\epsilon GM r_c^2 - L^2 r_c + 3GML^2\gamma = 0 \ , \tag{45}$$

where $\gamma = 0$ in Newtonian gravity and $\gamma = 1$ in general relativity.

Circular orbits will be **stable** if they correspond to a minimum of the potential, and unstable if they correspond to a maximum. Bound orbits which are not circular will oscillate around the radius of the stable circular orbit.

In Newtonian gravity (Figure 2), we find that circular orbits appear at

$$r_c = \frac{L^2}{\epsilon GM} \ , \tag{46}$$

and are always stable.

If the energy is greater than the asymptotic value $E = 1$, the orbits will be unbound, describing a particle that approaches the star and then recedes to infinity. We know that the orbits in Newton's theory are conic sections: bound orbits are either circles or ellipses, while unbound ones are either parabolas or hyperbolas (we won't show that here).

In general relativity (Figure 3) the situation is different, but only for r sufficiently small.

Since the difference resides in the term $-GML^2/r^3$, as $r \to \infty$ the behaviors are identical in the two theories. But as $r \to 0$ the potential goes to $-\infty$ rather than $+\infty$ as in the Newtonian case. At $r = 2GM$ the potential is always zero; inside this radius is the black hole, which we will discuss more thoroughly later.

Solving for the radius for which $dV/dr = 0$ (Equation 45), we find the radii of circular orbits

$$r_c = \frac{L^2 \pm \sqrt{L^4 - 12G^2M^2L^2}}{2GM} \ . \tag{47}$$

For large L there will be two circular orbits, one stable and one unstable. In the $L \to \infty$ limit their radii are given by

$$r_c = \frac{L^2 \pm L^2(1 - 6G^2M^2/L^2)}{2GM} = \left(\frac{L^2}{GM}, \ 3GM\right) \ . \tag{48}$$

In this limit the stable circular orbit becomes farther and farther away as L increases, while the unstable one approaches $3GM$, behavior which parallels the massless case, as we shell see below.

As L decreases, the two circular orbits come closer together; they coincide when the discriminant in Equation 47 vanishes, at

$$L = \sqrt{12}GM \ , \tag{49}$$

for which

$$r_c = r_{ISCO} = 6GM \ , \tag{50}$$

and disappear entirely for smaller L. *Thus* $6GM$ *is the smallest possible radius of a stable circular orbit in the Schwarzschild metric.* Hence the name "inner most stable circular orbit", and the sign r_{ISCO}. The radius r_{ISCO} is important for practical reasons: it a *measurable* quantity of accretion disks that form around massive black holes. In fact (we will not show it here), measurements of r_{ISCO} can probe the *spin* of the black hole, since r_{ISCO} is different for Schwarzschild metric and for Kerr metric which describes rotating black holes. The technology for making such measurements have only matured in the past 2-3 years.

There are also unbound orbits, which come in from infinity and turn around, and bound but non-circular ones, which oscillate around the stable circular radius. Note that such orbits, which would describe exact conic sections in Newtonian gravity, will not do so in GR, although we would have to solve the equation for $d\phi/dt$ to demonstrate it. Finally, there are orbits which come in from infinity and continue all the way in to $r = 0$; this can happen either if the energy is higher than the barrier, or for $L < \sqrt{12}GM$, when the barrier goes away entirely.

We have therefore found that the Schwarzschild solution possesses stable circular orbits for $r > 6GM$ and unstable circular orbits for $3GM < r < 6GM$. It's important to remember that these are only the geodesics; there is nothing to stop an accelerating particle from dipping below $r = 3GM$ and emerging, as long as it stays beyond $r = 2GM$. Massive particle trajectories are plotted in Figures 4, 5.

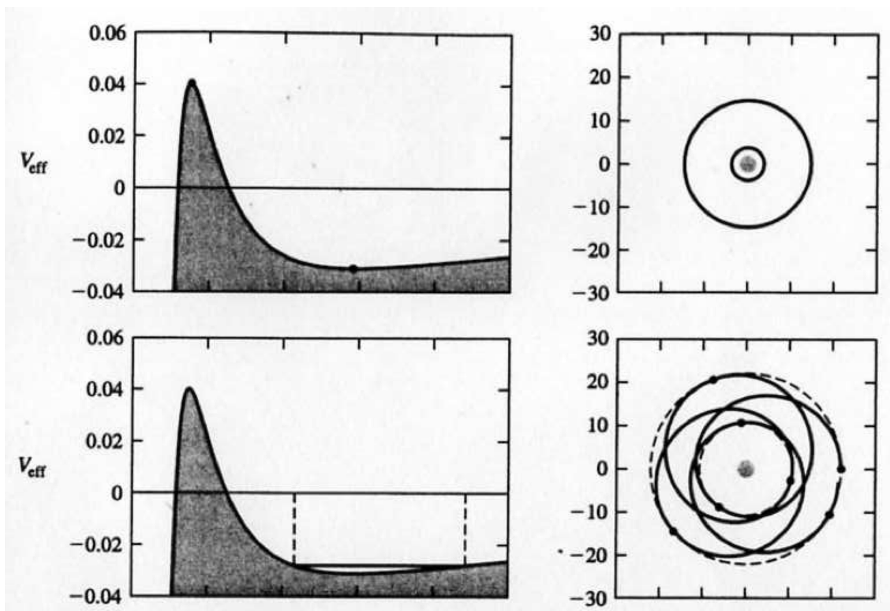

Fig. 4.— Massive particle orbits in Schwarzschild geometry (part I: bound orbits); $L/M = 4.3$, different values of E. Left: potential, with horizontal lines represent E. Right: orbits in polar coordinates (r-ϕ plane). Upper: two circular orbits ($dV/dr = 0$) outer orbit is stable, inner one is not. Lower: bound orbit between two turning radii.

4.4. massless particles

The treatment of light deflection is very similar to that of massive particles; We simply put $\epsilon = 0$ in the potential (Equation 44) and the condition for a flat potential (Equation 45). The potentials in the Newtonian case and in the GR case are presented in Figures 6, 7.

In the Newtonian case (Figure 6), Equation 45 shows us that there are no circular orbits, as we already know. Figure 6 indeed shows that there are no bound orbits of any sort.

Although it is somewhat obscured in this coordinate system, don't forget that since the Newtonian gravitational force on a massless particle is zero, massless particles actually move

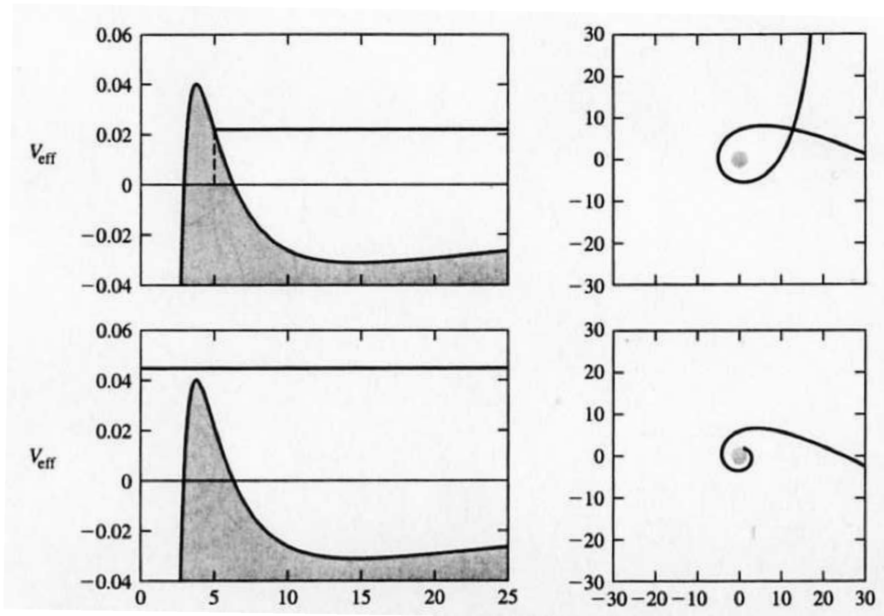

Fig. 5.— Massive particle orbits in Schwarzschild geometry (part II: unbound orbits); $L/M = 4.3$, different values of E. Upper: particle comes from infinity, passes around the center of attraction and moves out to infinity again. Note that the orbit **does not** look like a parabola. Lower: particles comes from infinity and plunges into the center.

in a straight line.

In terms of the effective potential, a photon with a given energy E will come in from $r = \infty$ and gradually "slow down" (actually $dr/d\lambda$ will decrease, but the speed of light isn't changing) until it reaches the turning point, when it will start moving away back to $r = \infty$. The lower values of L, for which the photon will come closer before it starts moving away, are simply those trajectories which are initially aimed closer to the gravitating body.

In general relativity (Figure 7) the situation is different. Except for the case $L = 0$, there is always a barrier; however, a sufficiently energetic photon will nevertheless go over the barrier and be dragged inexorably down to the center. (Note that "sufficiently energetic" means "in comparison to its angular momentum" — in fact the frequency of the photon is immaterial, only the direction in which it is pointing.) Thus, in reality, if a photon reaches from infinity and does not cross the barrier, its trajectory is tilted, and it continues to infinity (analogue to a Coulomb scattering between electron and proton). If it does cross the barrier, it falls into the potential well and never returns.

At the top of the barrier there are unstable circular orbits. For $\epsilon = 0$, $\gamma = 1$, Equation

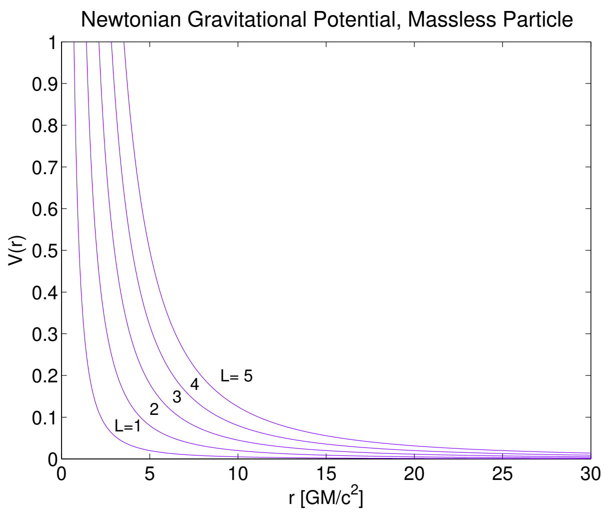

Fig. 6.— Gravitational potential for massless particles: the Newtonian case.

45 (which was written in a general form to include photons as well), results in

$$r_c = 3GM \ , \tag{51}$$

just as in the limit case for the massive particles, see Equation 48. This is borne out by Figure 7, which shows a maximum of $V(r)$ at $r = 3GM$ for every L. This means that a photon can orbit forever in a circle at this radius, but any perturbation will cause it to fly away either to $r = 0$ or $r = \infty$. Possible photon trajectories are presented in Figure 8.

5. Classical tests of GR

Most experimental tests of general relativity involve the motion of test particles in the solar system, and hence geodesics of the Schwarzschild metric. The exception is the gravitational red-shift of emission lines, which is measured in distant objects (pulsars).

Einstein himself suggested three tests of general relativity:

1. The gravitational red shift of spectral lines;

2. Deflection of light by the sun.

3. Precession of the perihelia of the inner planets (perihelia = point in the orbit of a planet in which it is closest to the sun).

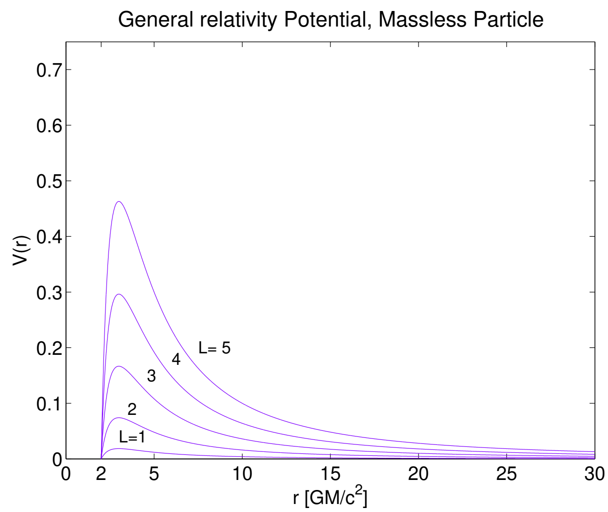

Fig. 7.— Gravitational potential for massless particles: full GR potential

Since then, another suggested test is

4. Gravitational time delay of radar signal sent (and reflected) in a gravitational field (Shapiro time delay).

Unfortunately, we don't have the time to fully analyze all these tests; they were all measured experimentally, and the results are found in excellent agreement with the theoretical predictions.

We will analyze here the gravitational redshift, and very briefly the other effects. Full analysis of the other tests can be found in Weinberg and Hartle.

5.1. The Gravitational Redshift

We have already discussed gravitational redshift, in the context of weak field approximation. Let us generalize the discussion here for the Schwarzschild metric.

We consider two observers who are not moving on geodesics, but are stuck at fixed spatial coordinate values (r_1, θ_1, ϕ_1) and (r_2, θ_2, ϕ_2). According to Equation 41, the proper time of observer i will be related to the coordinate time t by

$$\frac{d\tau_i}{dt} = \left(1 - \frac{2GM}{r_i}\right)^{1/2} . \tag{52}$$

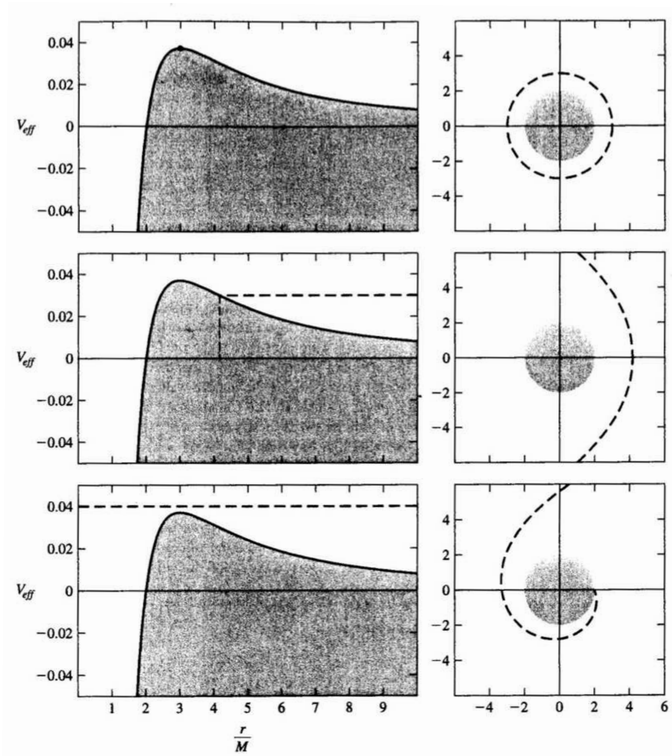

Fig. 8.— Massless particles trajectories in Schwarzschild geometry. Upper: circular orbit; middle: scattered orbit; bottom: plunge orbit.

Equation 52 can of course also be obtained directly from the Schwarzschild metric by recalling that the proper time = the time measured by an observer carrying the clock with her, is related to the interval via $(\Delta\tau)^2 = -(\Delta s)^2/c^2$ (see, "Special relativity", Equation 21) and using the Schwarzschild metric (equation 27) for an observer fixed in space ($dr = d\Omega = 0$). Also, recall that $U^0 = \frac{dt}{d\tau}$, where τ is the proper time and t is the *coordinate time*, which is, in general and in the specific case of Schwarzschild metric, different than the proper time. (Alongside with the spatial coordinates, the coordinate time is what is used to describe an event).

Suppose that the observer $\mathcal{O}_1$ emits a light pulse which travels to the observer $\mathcal{O}_2$, such that $\mathcal{O}_1$ measures the time between two successive crests of the light wave to be $\Delta\tau_1$. Each crest follows the same path to $\mathcal{O}_2$, except that they are separated by a coordinate time

$$\Delta t = \left(1 - \frac{2GM}{r_1}\right)^{-1/2} \Delta\tau_1 \ . \tag{53}$$

This separation in coordinate time does not change along the photon trajectories (the Killing vector associated with time change is conserved !), but the second observer measures a time between successive crests (in his clock, hence proper time) given by

$$\begin{aligned} \Delta\tau_2 &= \left(1 - \frac{2GM}{r_2}\right)^{1/2} \Delta t \\ &= \left(\frac{1-2GM/r_2}{1-2GM/r_1}\right)^{1/2} \Delta\tau_1 \ . \end{aligned} \tag{54}$$

Since these intervals $\Delta\tau_i$ measure the proper time between two crests of an electromagnetic wave, the observed frequencies will be related by

$$\begin{aligned} \frac{\omega_2}{\omega_1} &= \frac{\Delta\tau_1}{\Delta\tau_2} \\ &= \left(\frac{1-2GM/r_1}{1-2GM/r_2}\right)^{1/2} \ . \end{aligned} \tag{55}$$

This is an exact result for the frequency shift; in the limit $r >> 2GM$ we have

$$\begin{aligned} \frac{\omega_2}{\omega_1} &= 1 - \frac{GM}{r_1} + \frac{GM}{r_2} \\ &= 1 + \Phi_1 - \Phi_2 \ . \end{aligned} \tag{56}$$

This tells us that the frequency goes down as Φ increases, which happens as we climb out of a gravitational field; thus, a redshift. You can check that it agrees with our previous calculation based on the equivalence principle.

The origin of this is of course that for a static observer in the Schwarzschild metric the proper time interval $(\Delta\tau)$ between two events is a function of the coordinate r - the clocks of two static observers located at different radii in Schwarzschild metric tick at different paces, and the time delay is given by Equation 52.

5.2. Deflection of Light

Historically, this was the first independent test of GR. While in Newtonian gravity photons move in straight lines, in GR their paths are deflected. This can be observed when we look at the light coming from a distant star which is "nearly behind" the sun, and 1/2 a year later when the earth is in the other side of the sun. From practical reasons, the first measurement can be done only during solar eclipse. The location of the star in the sky (=relatively to other stars) will change.

Consider a light ray that approaches from infinity. Using Equations 43 and 44, we find that (with $\epsilon = 0$)

$$\frac{1}{L^2}\left(\frac{dr}{d\lambda}\right)^2 + \frac{1}{r^2}\left(1 - \frac{2GM}{r}\right) = \frac{E^2}{L^2}. \tag{57}$$

Defining the impact parameter as

$$b \equiv \frac{L}{E}, \tag{58}$$

and using Equation 40 ($L = r^2 d\phi/d\lambda$), we find

$$\frac{d\phi}{dr} = \pm \frac{1}{r^2}\left[\frac{1}{b^2} - \frac{1}{r^2}\left(1 - \frac{2GM}{r}\right)\right]^{-1/2} \tag{59}$$

(see Figure 9).

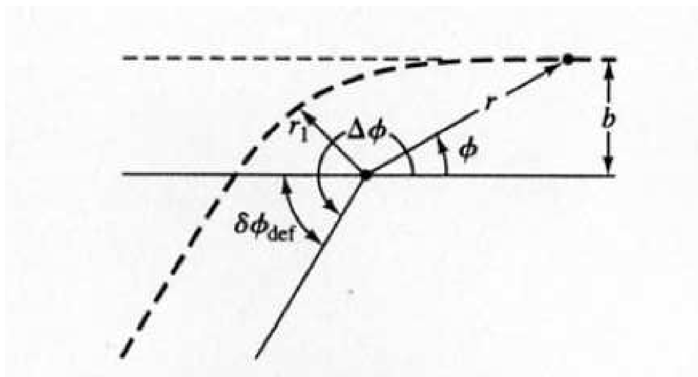

Fig. 9.— Deflection of light by angle $\delta\phi_{def}$.

Getting the maximum deflection angle is now a matter of simple integration,

$$\Delta\phi = 2\int_{r_1}^{\infty} \frac{dr}{r^2}\left[\frac{1}{b^2} - \frac{1}{r^2}\left(1 - \frac{2GM}{r}\right)\right]^{-1/2} \tag{60}$$

where $r = r_1$ is the turning point, which is the radius where $1/b^2 = (1 - 2GM/r)/r^2$.

For deflection of light by the sun, the impact parameter b cannot be smaller than the stellar radius, $b \geq R_\odot \approx 7 \times 10^{10}$ cm, and thus $2GM_\odot/c^2 b \leq 10^{-6}$.

Hartle shows that in this case, equation 60 can be solved by (i) change of parameters, $w = b/r$, and (ii) expansion in GM/b in a Taylor series and keeping the first order. The result is

$$\Delta\phi \approx \pi + \frac{4GM}{b}. \tag{61}$$

The actual deflection is

$$\delta\phi_{def} = \Delta\phi - \pi \approx \frac{4GM}{c^2 b}, \tag{62}$$

which is $\lesssim 1.75''$ (arc-sec).

This effect is also seen outside our solar system, as part of what is known as "gravitational lensing".

5.3. Precession of the Perihelia

The basic idea should be clear: since the shape of the gravitational potential in GR is different than in Newtonian mechanics, the perihelia of bound orbits (= stellar orbits, which are nearly ellipses) precess. Historically, the first test of GR was in fact its ability to explain the precession of mercury (which was known already in Einstein's time to deviate from Newtonian expectations).

The calculation is very similar to the one of the light deflection, and I will not put it here; you can find it in Hartle (or Carroll). We basically solve for $d\phi/dr$ (similar to equation 59), and integrate between successive turning points. For small deflection angle, the result is

$$\delta\phi_{prec} = \frac{6\pi GM}{c^2 a(1-\varepsilon^2)}, \tag{63}$$

where a is the semi-major axis and ε is the eccentricity. Obviously, this effect is largest for small a. For mercury, it predicts 43 arc-sec per century, which is consistent with observations.

5.4. Shapiro Time Delay

(Just a quick sketch). Instead of solving for $d\phi/dr$, one can solve for dt/dr (using equation 38). The result is

$$\frac{dt}{dr} = \pm\frac{1}{b}\left(1-\frac{2GM}{r}\right)^{-1}\left[\frac{1}{b^2}-\frac{1}{r^2}\left(1-\frac{2GM}{r}\right)\right]^{-1/2} \tag{64}$$

Thus, if one sends an electromagnetic signal in a gravitational field, which can be reflected back, and one knows the location of the emitter and reflector, it is possible to calculate the GR correction to the time of the returned signal. As emitter we take a radar located at earth, and as reflector we take one of the inner planets (mercury or Venus). The GR excess time delay is

$$(\Delta t)_{excess} \approx \frac{4GM}{c^3}\left[\log\left(\frac{4r_R r_{earth}}{r_1^2}\right)+1\right] \tag{65}$$

where $r_1 = b$ is the radius of closest approach to the center.

This experiment was proposed and carried by Irwin Shapiro, and is known as **Shapiro time delay**. In the solar system, the effect is a few hundreds microseconds, over a period of ~ 1 hour it takes the signal to reach the planets and return; thus a minimum accuracy of $\sim 10^{-7}$ is needed to see this effect!.

The experiment is conducted by sending a radar signal to venus which reflects the signal, and measure the signal's arrival time. The results are then compared to the expected from Keplerian orbit (see Figure 10). When earth and venus are in the opposite sides of the sun (as occured on Jan. 25th, 1970) the signal passes closest to the sun and the delay is maximal. When earth and venus are at the same side of the sun, the delay is minimal.

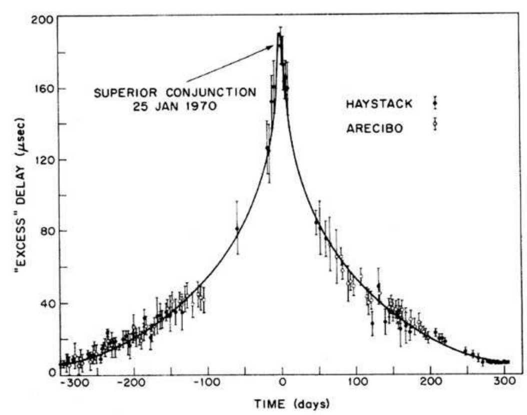

Fig. 10.— Excess of time delay of radar reflection between earth and venus with respect to the expected from Keplerian orbits, over 2 years of measurements.

6. Black Holes

So far we have discussed trajectories of particles - both massive and massless (geodesics) outside the troublesome radius $r = 2GM$. I should stress that this is the regime of interest for the solar system and most other astrophysical situations.

Now let us turn to study objects which are described by the Schwarzschild solution even at radii smaller than $2GM/c^2$ — black holes. (We will use the term "black hole" for the moment, even though we haven't introduced a precise meaning for such an object.)

Consider a light ray propagating in the radial direction: θ and ϕ are constant and $ds^2 = 0$. The Schwarzschild metric (Equation 27) thus implies

$$ds^2 = 0 = -\left(1 - \frac{2GM}{r}\right) \mathrm{d}t^2 + \left(1 - \frac{2GM}{r}\right)^{-1} \mathrm{d}r^2 \ , \tag{66}$$

from which we see that

$$\frac{dt}{dr} = \pm \left(1 - \frac{2GM}{r}\right)^{-1} \ . \tag{67}$$

For large r the slope $dt/dr = \pm 1$, as it would be in flat space, while as we approach $r = 2GM/c^2$ we get $dt/dr \to \pm\infty$. Thus a light ray which approaches $r = 2GM/c^2$ never seems to get there, at least in this coordinate system; instead it seems to asymptote to this radius (see Figure **??**).

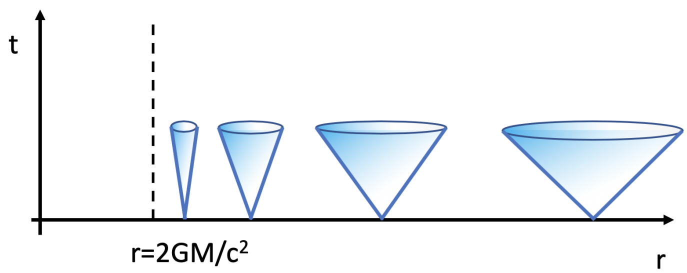

Fig. 11.— In Schwarzschild coordinates, light cones appear to close up as we approach $r = 2GM/c^2$.

This is really an illusion; **the light ray (or a massive particle) actually has no trouble reaching** $r = 2GM/c^2$. But an observer far away would never be able to tell.

If we stayed outside while an intrepid observational general relativist dove into the black hole, sending back signals all the time, we would simply see the signals reach us more and more slowly. This is of course consistent with the gravitational time delay, Equation 55. As infalling astronauts approach $r = 2GM/c^2$, any fixed interval $\Delta\tau_1$ of their proper time corresponds to a longer and longer interval $\Delta\tau_2$ from our point of view. This continues forever; we would never see the astronauts cross $r = 2GM/c^2$, we would just see them move more and more slowly (and become redder and redder, almost as if they were embarrassed to have done something as stupid as diving into a black hole).

Indeed, we would never be able to see the infalling astronauts reach $r = 2GM/c^2$. But the astronauts will still be there, in a finite amount of their proper time. To show this, the best way is to switch to a different coordinate system, which is better behaved at $r = 2GM/c^2$. While of course one cannot "derive" a coordinate system, we will motivate the choice using several steps as follows.

Equation 67 shows that the problem with the current coordinates is that $dt/dr \to \infty$ along radial null geodesics which approach $r = 2GM/c^2$. Thus, progress in the r direction becomes slower and slower with respect to the coordinate time t. This suggests a way to fix it: replace the t with a different coordinate, which "moves more slowly" along null geodesics.

This is done in three steps.

The first thing is to simply solve Equation 67. This gives

$$t = \pm r^* + \text{ constant } , \tag{68}$$

where the **tortoise coordinate** r^* is defined by[2]

$$r^* = r + 2GM \ln\left(\left|\frac{r}{2GM} - 1\right|\right) \ . \tag{69}$$

(simply plug that in Equation 68 to obtain the result in Equation 67; note that $dr^* = dr/(1 - 2GM/r)$). In terms of the tortoise coordinate (replace $r \to r^*$) the Schwarzschild metric becomes

$$ds^2 = \left(1 - \frac{2GM}{r}\right)\left(-\mathrm{d}t^2 + \mathrm{d}r^{*2}\right) + r^2 d\Omega^2 \ , \tag{70}$$

where r is thought of as a function of r^*. OK. We have some progress - none of the metric coefficients becomes infinite at $r = 2GM$ (although both g_{tt} and $g_{r^*r^*}$ become zero). Furthermore, the lightcones never don't close up (see Figure 12). The price we pay, however, is that the surface of interest at $r = 2GM/c^2$ has just been pushed to infinity: at $r = 2GM, r^* \to -\infty$. Note that the tortoise coordinates do not make much physical sense in the $r < 2GM$ regime, but we still define them there for ease of the following math.

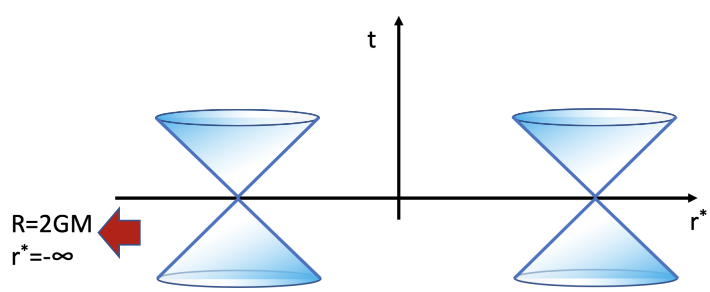

Fig. 12.— Light cones in tortoise coordinates in Schwarzschild metric are nondegenerate, but the surface $r = 2GM/c^2$ has been pushed to $r^* \to -\infty$.

---

[2]Defined this way, i.e., with the absolute value, the tortoise coordinates are well defined for all values of r, except $r = 2GM$. However, one can smoothly connect the two regions, $r > 2GM$ and $r < 2GM$. On the other hand, as $r \to 2GM$, $r^* \to -\infty$ and the region $r < 2GM$ does not carry physical meaning in these coordinates.

The 2^{nd} step is to define coordinates which are naturally adapted to the null geodesics. If we let

$$\begin{aligned} \tilde{u} &= t + r^* \\ \tilde{v} &= t - r^* \ , \end{aligned} \tag{71}$$

then *infalling* radial ($d\Omega = 0$) null ($ds^2 = 0$) geodesics are characterized by $\tilde{u} =$ constant, while the *outgoing* ones satisfy $\tilde{v} =$ constant. (the fact that $\tilde{u} = const$ refers to an infalling particle is seen from the fact that in the regime $r > 2GM$, when r decreases, r^* decreases, while the time t increases. We use $d\tilde{u} = dt + dr^*$ and Equation 68, $dt = -dr^*$ for infalling particle to conclude that $d\tilde{u} = 0$ in this case).

The third step is to replace the timelike coordinate t with the new coordinate $\tilde{u}$, but keep the r coordinate. These are known as **Eddington-Finkelstein coordinates**. In terms of them the metric is

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) \mathrm{d}\tilde{u}^2 + (\mathrm{d}\tilde{u}\mathrm{d}r + \mathrm{d}r\mathrm{d}\tilde{u}) + r^2 d\Omega^2 \ . \tag{72}$$

Even though the metric coefficient $g_{\tilde{u}\tilde{u}}$ vanishes at $r = 2GM$, there is no real degeneracy; the determinant of the metric is

$$g = -r^4 \sin^2\theta \ , \tag{73}$$

which is perfectly regular at $r = 2GM/c^2$. Therefore *the metric is invertible*, and we see directly that $r = 2GM/c^2$ is simply a coordinate singularity in our original (t, r, θ, ϕ) system.

This regularity enables us to extend the range of r across $2GM/c^2$ to all $r > 0$. To convince you, note that Equation 72 is a perfect solution to Einstein's field equation in vacuum, with a metric that is simply described by the independent coordinates $(\tilde{u}, r, \theta, \phi)$.

In the Eddington-Finkelstein coordinates the condition for radial null curves is solved by

$$\frac{d\tilde{u}}{dr} = \begin{cases} 0 \ , & \text{(infalling)} \\ 2\left(1 - \frac{2GM}{c^2 r}\right)^{-1} \ . & \text{(outgoing)} \end{cases} \tag{74}$$

(To obtain the outgoing relation, we use the radial ($d\Omega = 0$), null ($ds^2 = 0$) condition in Equation 72.) We can therefore see what has happened: in this coordinate system the light cones remain well-behaved at $r = 2GM/c^2$, and this surface is at a finite coordinate value. There is no problem in tracing the paths of null or timelike particles past the surface.

On the other hand, something interesting is certainly going on. If we look at trajectories of outgoing photons, for $r < 2GM/c^2$, $d\tilde{u}/dr < 0$: thus, all future-directed paths are in the direction of **decreasing** r (see Figure 13).

The surface $r = 2GM/c^2$, while being locally perfectly regular, *globally functions as a point of no return* — once a test particle dips below it, it can never come back. For this

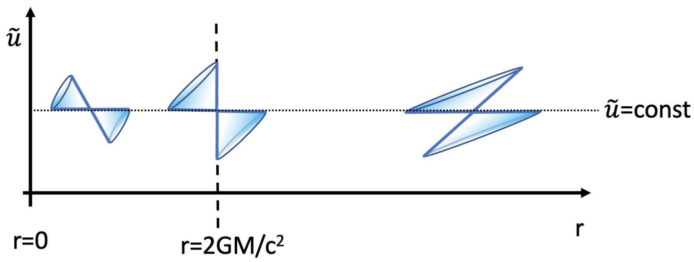

Fig. 13.— Lightcone in Eddington-Finkelstein coordinates, $\tilde{u}$ and r. If light is emitted at $r < 2GM/c^2$, all its future-directed paths are in the direction of decreasing r.

reason $r = 2GM/c^2$ is known as the **event horizon**; no event at $r \leq 2GM/c^2$ can influence any other event at $r > 2GM/c^2$. Notice that the event horizon is a null surface, not a timelike one. Notice also that since nothing can escape the event horizon, it is impossible for us to "see inside" — thus the name **black hole**.

Let us summarize our findings. We suspected that the t, r coordinates used do not give good description of the entire manifold. We have thus changed from our original coordinate t to the new one $\tilde{u}$. This new coordinate have a very nice property: if we decrease r along a radial null curve $\tilde{u}$ = constant (Equation 74), and we go right through the event horizon without any problems. *Indeed, a local observer actually making the trip would not necessarily know when the event horizon had been crossed !* — the local geometry is no different than anywhere else.

We therefore conclude that our suspicion was correct and our initial coordinate system didn't do a good job of covering the entire manifold. The region $r \leq 2GM$ should certainly be included in our spacetime, since physical particles can easily reach there and pass through.

6.1. Kruskal Coordinates

Apparently, this is not the end of it. Looking again at Equation 74, we see that in the $(\tilde{u}, r)$ coordinate system a particle can cross the event horizon on future-directed paths (getting in), but not on past-directed ones (getting out). However, we started with a time-independent solution: thus, we must be missing something.

Indeed, instead of choosing $\tilde{u}$ we could have chosen $\tilde{v}$ (Equation 71), in which case the metric would have been

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) \mathrm{d}\tilde{v}^2 - (\mathrm{d}\tilde{v}\mathrm{d}r + \mathrm{d}r\mathrm{d}\tilde{v}) + r^2 d\Omega^2 \ . \tag{75}$$

Now we can once again pass through the event horizon, but this time only along past-directed curves (!) (see Figure 14).

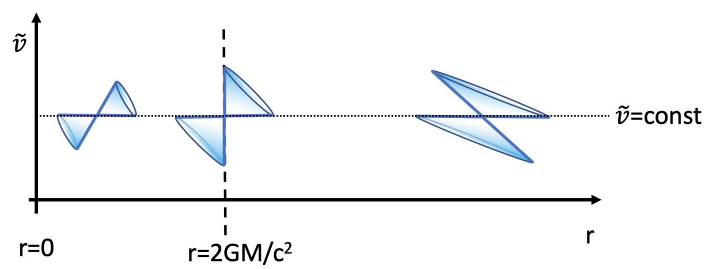

Fig. 14.— Lightcone in $\tilde{v}$ and r coordinates. Light emitted at $r < 2GM/c^2$, can always cross the event horizon outward, but light emitted outside the event horizon cannot enter it.

Really this is an extension of space-time in two directions: For if we keep $\tilde{u}$ constant and decrease r we must have $t \rightarrow +\infty$, while if we keep $\tilde{v}$ constant and decrease r we must have $t \rightarrow -\infty$. (The tortoise coordinate r^* goes to $-\infty$ as $r \rightarrow 2GM/c^2$.)

In order to get the coordinate system which uncovers the entire space-time, the way to proceed is to introduce new coordinates,

$$\begin{aligned} u' &= e^{\tilde{u}/4GM} \\ v' &= e^{-\tilde{v}/4GM} \ , \end{aligned} \tag{76}$$

which in terms of our original (t, r) system is

$$\begin{aligned} u' &= \left(\tfrac{r}{2GM} - 1\right)^{1/2} e^{(r+t)/4GM} \\ v' &= \left(\tfrac{r}{2GM} - 1\right)^{1/2} e^{(r-t)/4GM} \ . \end{aligned} \tag{77}$$

In the (u', v', θ, ϕ) system the Schwarzschild metric is

$$ds^2 = -\frac{16G^3M^3}{r} e^{-r/2GM} (\mathrm{d}u'\mathrm{d}v' + \mathrm{d}v'\mathrm{d}u') + r^2 d\Omega^2 \ . \tag{78}$$

Finally the nonsingular nature of $r = 2GM/c^2$ becomes completely manifest; in this form none of the metric coefficients behave in any special way at the event horizon.

Both u' and v' are null coordinates, in the sense that their partial derivatives $\partial/\partial u'$ and $\partial/\partial v'$ are null vectors. There is nothing wrong with this, since the collection of four partial derivative vectors (two null and two spacelike) in this system serve as a perfectly good basis for the tangent space. Nevertheless, we are somewhat more comfortable working in a system where one coordinate is timelike and the rest are spacelike. We therefore define

$$\begin{aligned} u &= \tfrac{1}{2}(u' - v') \\ &= \left(\tfrac{r}{2GM} - 1\right)^{1/2} e^{r/4GM} \cosh(t/4GM) \end{aligned} \tag{79}$$

and

$$\begin{aligned} v &= \tfrac{1}{2}(u' + v') \\ &= \left(\tfrac{r}{2GM} - 1\right)^{1/2} e^{r/4GM} \sinh(t/4GM) \ , \end{aligned} \tag{80}$$

in terms of which the metric becomes

$$ds^2 = \frac{32G^3M^3}{r} e^{-r/2GM}(-\mathrm{d}v^2 + \mathrm{d}u^2) + r^2 d\Omega^2 \ , \tag{81}$$

where r is defined implicitly from

$$(u^2 - v^2) = \left(\frac{r}{2GM} - 1\right) e^{r/2GM} \ . \tag{82}$$

The coordinates (v, u, θ, ϕ) are known as **Kruskal coordinates**, or sometimes Kruskal-Szekres coordinates (in some textbooks, the notation is different - $(v, u) \rightarrow (T, R)$). Note that v is the timelike coordinate. Furthermore, from Equation 81, clearly, $dv/du = \pm 1$.

The Kruskal coordinates have a number of miraculous properties. Like the (t, r^*) coordinates, the radial ($d\Omega = 0$) null ($ds = 0$) curves look like they do in flat space:

$$v = \pm u + \text{constant} \ . \tag{83}$$

Unlike the (t, r^*) coordinates, however, the event horizon $r = 2GM/c^2$ is not infinitely far away; in fact it is defined by

$$v = \pm u \ , \tag{84}$$

consistent with it being a null surface. More generally, we can consider the surfaces $r =$ constant. From Equation 82 these satisfy

$$u^2 - v^2 = \text{ constant} \ . \tag{85}$$

Thus, they appear as hyperbolae in the u-v plane. Furthermore, the surfaces of constant t are given by

$$\frac{v}{u} = \tanh(t/4GM) \ , \tag{86}$$

which defines straight lines through the origin with slope $\tanh(t/4GM)$. Note that as $t \to \pm\infty$ this becomes the same as Equation 84; therefore these surfaces are the same as $r = 2GM/c^2$.

Now, our coordinates (v, u) should be allowed to range over every value they can take without hitting the real singularity at $r = 0$; the allowed region is therefore $-\infty \leq u \leq \infty$ and $v^2 < u^2 + 1$. We can now draw a spacetime diagram in the v-u plane (with θ and ϕ suppressed), known as a "Kruskal diagram", which represents the entire spacetime corresponding to the Schwarzschild metric. Such a diagram is plotted in Figure 15.

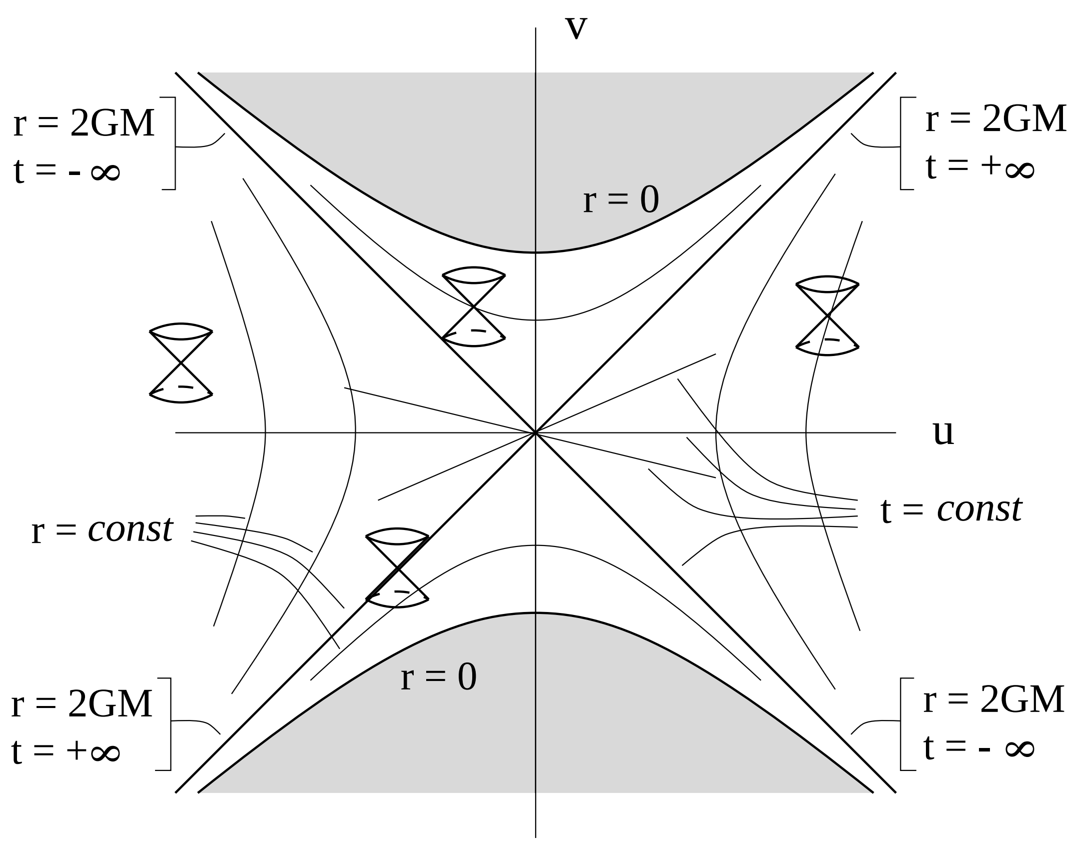

Fig. 15.— Kruskal diagram represents the entire space-time corresponding to Schwarzschild metric. Note that each point on the diagram is a two-sphere.

Our original coordinates (t, r) were only good for $r > 2GM$, which is only a part of the manifold portrayed on the Kruskal diagram. It is convenient to divide the diagram into four regions (see Figure 16):

We are always in region I; by following future-directed null rays we reached region II, and by following past-directed null rays we reached region III. If we had explored spacelike

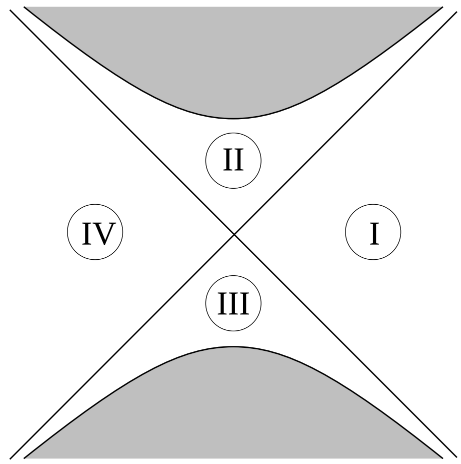

Fig. 16.— Kruskal diagram divides space-time into 4 regions.

geodesics, we would have been led to region IV.

Region II, of course, is what we think of as the black hole. Once anything travels from region I into II, it can never return. In fact, every future-directed path in region II ends up hitting the singularity at $r = 0$; once you enter the event horizon, you are utterly doomed. This is worth stressing; not only can you not escape back to region I, you cannot even stop yourself from moving in the direction of decreasing r, since this is simply the timelike direction. (This could have been seen in our original coordinate system; for $r < 2GM$, t becomes spacelike and r becomes timelike.) Thus you can no more stop moving toward the singularity than you can stop getting older. Since proper time is maximized along a geodesic, you will live the longest if you don't struggle, but just relax as you approach the singularity. Not that you will have long to relax. (Nor that the voyage will be very relaxing; as you approach the singularity the tidal forces become infinite. As you fall toward the singularity your feet and head will be pulled apart from each other, while your torso is squeezed to infinitesimal thinness.)

Regions III and IV might be somewhat unexpected. Region III is simply the time-reverse of region II, a part of spacetime from which things can escape to us, while we can never get there. It can be thought of as a "white hole." There is a singularity in the past, out of which the universe appears to spring. The boundary of region III is sometimes called the past event horizon, while the boundary of region II is called the future event horizon. Region IV, meanwhile, cannot be reached from our region I either forward or backward in time (nor can anybody from over there reach us). It is another asymptotically flat region of spacetime, a mirror image of ours. It can be thought of as being connected to region I by a "wormhole," a neck-like configuration joining two distinct regions.

6.2. Astrophysical black holes

In reality, we need not worry about regions III and IV. Remember that in real life, stellar radius is expected to be much larger than $2GM/c^2$, and thus the Schwarzschild solution is invalid; similarly, as we go backward in time, eventually we cannot ignore the fact that we are not in vacuum, and thus, again the Schwarzschild solution is invalid.

I should say a few words about the formation of astrophysical black holes from massive stars (you might have heard it already in astronomy courses).

The life of a star is a constant struggle between the inward pull of gravity and the outward push of pressure. When the star is burning nuclear fuel at its core, the pressure comes from the heat produced by this burning. (We should put "burning" in quotes, since nuclear fusion is unrelated to oxidation.) When the fuel is used up, the temperature declines and the star begins to shrink as gravity starts winning the struggle. Eventually this process is stopped when the electrons are pushed so close together that they resist further compression simply on the basis of the Pauli exclusion principle (no two fermions can be in the same state). The resulting object is called a **white dwarf**. If the mass is sufficiently high, however, even the electron degeneracy pressure is not enough, and the electrons will combine with the protons in a dramatic phase transition. The result is a **neutron star**, which consists of almost entirely neutrons (although the insides of neutron stars are not understood very well). Since the conditions at the center of a neutron star are very different from those on earth, we do not have a perfect understanding of the equation of state. Nevertheless, we believe that a sufficiently massive neutron star will itself be unable to resist the pull of gravity, and will continue to collapse. Since a fluid of neutrons is the densest material of which we can presently conceive, it is believed that the inevitable outcome of such a collapse is a black hole.

The process is summarized in figure 17, which is a diagram of radius vs. mass.

The point of the diagram is that, for any initial mass M, at the end of its life a star will decrease in radius until it hits the line. White dwarfs are found between points A and B, and neutron stars between points C and D. Point B is at a height of somewhat less than 1.4 solar masses; the height of D is less certain, but probably less than 2 solar masses. The process of collapse is complicated, and during the evolution the star can lose or gain mass, so the endpoint of any given star is hard to predict. Nevertheless white dwarfs are all over the place, neutron stars are not uncommon, and there are a number of systems which are

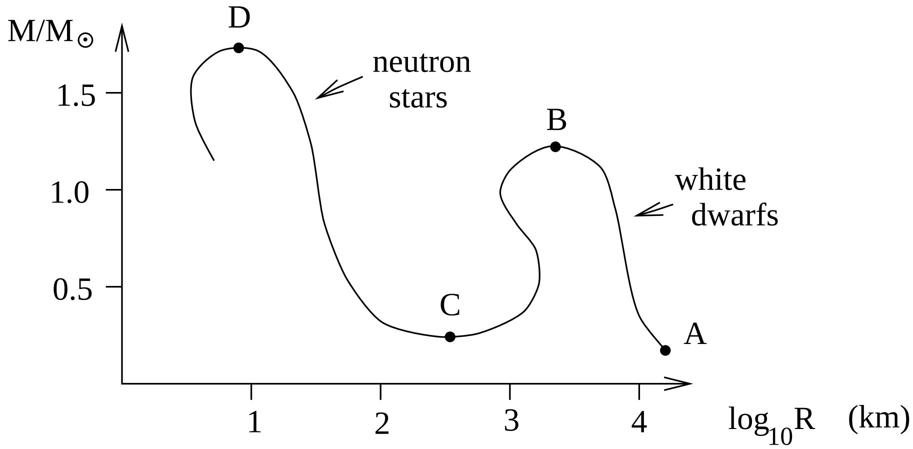

Fig. 17.— Rough sketch of R-M relation during the life cycle of a star.

strongly believed to contain black holes. (Of course, you can't directly see the black hole. What you can see is radiation from matter accreting onto the hole, which heats up as it gets closer and emits radiation.)

REFERENCES

[1] S. Carroll, *Lecture Notes on General Relativity*, part 7. The Schwarzschild Solution and Black Holes
(http://preposterousuniverse.com/grnotes/).

[2] J. Hartle, *Gravity: An Introduction to Einstein's General Relativity* (Addison-Wesley), chapters 9, 10 and 12.

[3] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley & Sons), chapters 8 and 11.

[4] C. Misner, K. Thorne and J. Wheeler, *Gravitation* (W.H. Freeman and Co.), chapters 31, 32 and 33.