Phase-plane analysis of perihelion precession and Schwarzschild orbital dynamics

Bruce Dean[a)]
Physics Department, West Virginia University, Morgantown, West Virginia 26506-6315

(Received 22 December 1997; accepted 11 June 1998)

A calculation of perihelion precession is presented that utilizes a phase-plane analysis of the general relativistic equations of motion. The equations of motion are reviewed in addition to the phase-plane analysis required for the calculation. ''Exact'' phase planes for orbital dynamics in the Schwarzschild geometry are discussed, and bifurcations are identified as a dimensionless parameter involving the angular momentum is varied. © *1999 American Association of Physics Teachers.*

I. INTRODUCTION

The perihelion precession of planetary orbits has provided one of the earliest experimental tests of Einstein's general theory of relativity. In the standard textbook presentation of this calculation[1] there are essentially two approaches taken to calculate its value from the nonlinear equations of motion:

(a) approximate an elliptic integral,
(b) find a perturbative solution to the general relativistic equations.

Although (a) and (b) are the most common methods appearing in the literature, other approximation methods do exist. Wald,[1] for instance, considers small oscillations about an elliptical orbit; Misner, Thorne, and Wheeler (MTW)[1] consider nearly circular orbits and then later use the PPN (''Parametrized Post-Newtonian'') formalism. The purpose of this paper is to illustrate how the perihelion calculation may be performed and to present an analysis of the Schwarzschild orbital dynamics based on a standard technique of nonlinear analysis: the phase-plane approach (see also Refs. 2 and 3). Not only is the calculation simpler to perform in the modern setting of phase-plane analysis, but there is more physics to be learned with less algebra compared with the standard procedures.

The contents of this paper are organized as follows. In Sec. II the general relativistic equations of motion are derived. The goal here is to not only make the presentation as self-contained as possible, but to ''tailor'' the derivation toward a discussion emphasizing the phase-plane analysis, and for easy comparison with the corresponding Newtonian calculation. In Sec. III, the phase-plane analysis is developed and in Sec. IV applied to obtain the well-known value of perihelion precession. In Sec. V, a discussion of the Schwarzschild orbital dynamics is given based upon an ''exact'' general relativistic phase plane. The standard results are discussed, but also an alternative viewpoint for analyzing the orbital dynamics is presented based upon the separatrix structure of the phase plane. In this approach, the critical relationship that holds between energy and momentum at the unstable orbital radius (i.e., the separatrix) summarizes the range of physically possible orbits, and demonstrates a saddle-center bifurcation as a dimensionless parameter involving the angular momentum is varied.

In Sec. VI, a phase-plane analysis of the dynamical invariance between the coordinate and proper time reference frames is given. Although the dynamical structure (i.e., the effective potential) is invariant between the two reference frames, the phase diagrams in each case are not identical. This is due to the existence of an additional phase-plane fixed point that appears in the coordinate reference frame at the event horizon. This fixed point is obviously coordinate dependent, but must exist to explain the apparent ''slowing down'' of objects (and redshift of signals) approaching the horizon as seen by an observer in the coordinate reference frame.

For comparison with the relativistic case, the corresponding Newtonian phase-plane results are discussed in an Appendix. Not only does this analysis complement the dynamics considered in Sec. VI, but it is shown that an analysis of Newtonian orbits using time as an independent variable is just as instructive and no more complicated in principle than using the equatorial angle as the independent variable (however, the opposite is true when using the standard methods of analysis, e.g., Ref. 4). Furthermore, by considering t as the independent variable rather than φ, an additional fixed point appears at infinity. But more importantly, the emphasis of the analysis is shifted from trying to find an explicit closed form solution (i.e., the standard approach) to a more intuitive and qualitative description based on the energy method.

Finally, the phase-plane analysis is applied to the kinematics of light rays in the Schwarzschild black hole spacetime. The standard results are discussed and then compared with the timelike phase-plane results. The added significance of the photon orbits (in the phase-plane context) is that the equilibrium points of the differential equations exhibit a transcritical bifurcation (i.e., a change in stability) at these parameter values.

II. GENERAL RELATIVISTIC ORBITS

The general relativistic equations of motion for a point mass with rest mass, m_0, orbiting a mass, M (assuming for simplicity that $m_0 \ll M$), originate from the Schwarzschild line element:

$$ds^2 = c^2 \Lambda dt^2 - \Lambda^{-1} dr^2 - r^2 d\Omega^2,$$
$$\Lambda = 1 - r_S/r, \quad d\Omega^2 = d\theta^2 + \sin^2\theta d\varphi^2. \tag{1}$$

Equation (1) is expressed using spherical coordinates and r_S is the Schwarzschild radius (G is Newton's gravitational constant and c is the speed of light):

$$r_S = 2MG/c^2. \tag{2}$$

The Lagrangian is a constant of the motion:

 © 1999 American Association of Physics Teachers

$$L=\tfrac{1}{2}m_0(ds/d\tau)^2=\tfrac{1}{2}m_0c^2;\quad \tau\equiv\text{proper time}, \tag{3}$$

and if the orbit is confined to the equatorial plane, i.e., $\theta=\pi/2$, L takes the explicit form ($\dot{t}=dt/d\tau$, etc.):

$$L=\tfrac{1}{2}m_0c^2=\tfrac{1}{2}m_0c^2\Lambda\dot{t}^2-\tfrac{1}{2}m_0\Lambda^{-1}\dot{r}^2-\tfrac{1}{2}m_0r^2\dot{\varphi}^2. \tag{4}$$

From the Euler–Lagrange equations there are two additional constants of motion [(4) is trivial; (5) and (6) are first integrals]:

$$\partial L/\partial t=0\Rightarrow\partial L/\partial\dot{t}=E=m_0c^2\Lambda\dot{t}, \tag{5}$$

$$\partial L/\partial\varphi=0\Rightarrow\partial L/\partial\dot{\varphi}=J=m_0r^2\dot{\varphi}. \tag{6}$$

Physically, E is the energy required for an observer at infinity to place m_0 in orbit about M [it is left as an exercise to check this physical interpretation by considering radial motion in (4) and then combining with (5) in the Newtonian limit]. J is the angular momentum of the system and since this is constant, there will be no precession of the equatorial plane.

Continuing with the equations of motion, using (5) and (6) to eliminate $\dot{t}$ and $\dot{\varphi}$ from (4) and then rearranging algebraically gives the following result:

$$\dot{r}^2/c^2=(dr/ds)^2=\hat{E}^2-(1+J^2/m_0^2c^2r^2)\Lambda, \tag{7}$$

where $\hat{E}\equiv E/m_0c^2$ defines the total energy per unit rest energy. Noting the functional dependence of r on the equatorial angle [i.e., $r=r(\varphi)\Rightarrow\dot{r}=(dr/d\varphi)\dot{\varphi}$] allows (7) to be further expressed in terms of the constant J. Furthermore, the degree of this equation (in r) is reduced by making the usual change of variable to $u=r_S/r$. Simplifying algebraically gives the following result:

$$(du/d\varphi)^2=2\sigma\hat{E}^2-(2\sigma+u^2)\Lambda, \tag{8}$$

where σ defines the dimensionless parameter:

$$\sigma=\tfrac{1}{2}(m_0cr_s/J)^2=2(GMm_0/cJ)^2=\tfrac{1}{2}(r_S/J)^2. \tag{9}$$

Equation (9) is expressed on the far right-hand side in the ''geometrized'' system of units (i.e., $G=c=1$, $r_S=2M$; see, e.g., Shutz,[1] p. 198) with m_0 taken as unity for later comparison with the standard results.

Differentiating (8) with respect to φ then gives the familiar second-order equation in dimensionless form:

$$d^2u/d\varphi^2+u=\sigma+\tfrac{3}{2}u^2. \tag{10}$$

As previously discussed, there are two common procedures for calculating the value of perihelion precession. The first procedure [procedure (A)] is to approximate an elliptic integral obtained by separating variables in (8) (see, for example, Ref. 5). In procedure (B), (10) is solved perturbatively or by using other approximation techniques. However, at this point we shall deviate from these approaches and consider (10) from the viewpoint of phase-plane analysis. Additional discussion and comparison with the Newtonian case is given in the Appendix.

III. PHASE-PLANE ANALYSIS

To begin the phase-plane analysis of (10) (see, e.g., Strogatz[6] or Tabor[7] for an introduction), let us convert this second-order, nonlinear, inhomogeneous, differential equation to two first-order equations by defining new variables: $x=u$ and $y=du/d\varphi$. In 2-d form, (10) is equivalent to (primes will denote derivatives with respect to φ)

$$\begin{aligned} x'&=f(x,y)=y,\\ y'&=g(x,y)=\tfrac{3}{2}x^2-x+\sigma. \end{aligned} \tag{11}$$

To find the fixed points of (11) (i.e., equilibrium points of the solution) we solve simultaneously: $x'=y'=0$, for x and y. Therefore, the fixed points of (11) are given by

$$\vec{x}_1^*=\left(\frac{1+\sqrt{1-6\sigma}}{3},0\right),\quad \vec{x}_2^*=\left(\frac{1-\sqrt{1-6\sigma}}{3},0\right). \tag{12}$$

Alternatively, by expressing y in terms of x using (8):

$$x'=y=\pm[2\sigma\hat{E}^2-(2\sigma+x^2)(1-x)]^{1/2}=0, \tag{13}$$

and then solving simultaneously: $x'=y'=0$, for $\hat{E}^2$ and x rather than x and y, the corresponding energies at each fixed point are expressed solely in terms of σ:

$$\begin{aligned} \hat{E}_1^2-1&=\frac{2\sigma[1-4\sigma-(1-6\sigma)^{1/2}]}{[(1-6\sigma)^{1/2}-1]^3},\\ \hat{E}_2^2-1&=\frac{2\sigma[-1+4\sigma-(1-6\sigma)^{1/2}]}{[(1-6\sigma)^{1/2}-1]^3}, \end{aligned} \tag{14}$$

respectively. Therefore, solving simultaneously for $\hat{E}^2$ and x gives additional information on the dynamics. Furthermore, the phase-plane equations analogous to (11) that result from the proper and coordinate time analysis considered in Sec. VI (and also in the Newtonian case) give nonphysical roots when solving only for x and y (i.e., they do not correspond to the effective potential extrema). However, these additional roots are eliminated by solving for $\hat{E}^2$ and x as illustrated above and as discussed in Sec. VI.

To give a general classification of the fixed points (12) a linear stability analysis must be performed. Essentially, this amounts to series expanding (11) about an arbitrary fixed point in the small parameters: $\delta x=x-x^*$ and $\delta y=y-y^*$. Dropping second-order terms, the resulting first-order linear equations are expressed in matrix form:

$$\begin{aligned} \begin{pmatrix}\delta x'\\ \delta y'\end{pmatrix}&\approx\left.\begin{pmatrix}\partial_x f & \partial_y f\\ \partial_x g & \partial_y g\end{pmatrix}\right|_{\vec{x}=\vec{x}^*}\begin{pmatrix}\delta x\\ \delta y\end{pmatrix}\\ &=\left.\begin{pmatrix}0 & 1\\ 3x-1 & 0\end{pmatrix}\right|_{\vec{x}=\vec{x}^*}\begin{pmatrix}\delta x\\ \delta y\end{pmatrix}\equiv A|_{\vec{x}=\vec{x}^*}\delta\vec{x}. \end{aligned} \tag{15}$$

The general solution of (15) is an exponential whose stability at each fixed point is analyzed by classifying the eigenvalues of the matrix **A**. Solving the eigenvalue problem, we find roots to

$$|A-\lambda I|=0, \tag{16}$$

but since **A** is 2×2, the characteristic polynomial may be expressed:

$$\lambda^2-\tau\lambda+\Delta=0, \tag{17}$$

where $\tau\equiv$trace A, and $\Delta\equiv$determinant A. The eigenvalues are roots of (17):

$$\lambda=\tfrac{1}{2}(\tau\pm\sqrt{\tau^2-4\Delta}), \tag{18}$$

and accordingly, the exponential solutions of (15) are classified by various regions of Fig. 1 [dots mark the location of the fixed points given in (12)].

Briefly, region I corresponds to a ''saddle-node'' fixed point, whose stable and unstable manifolds [corresponding to

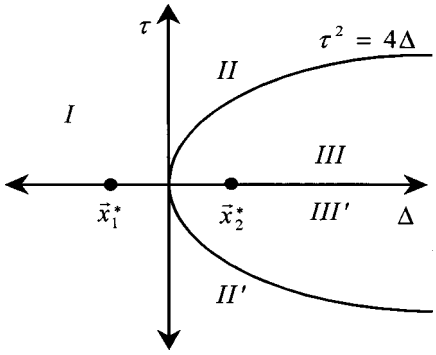

Fig. 1. Eigenvalue classification.

positive and negative (real) eigenvalues, respectively] are given by the eigenvectors of (15). Region II represents an ''unstable node,'' i.e., two positive real eigenvalues with $\tau^2-4\Delta>0$; region III gives solutions having one (positive) real and one complex eigenvalue (''unstable spirals''), while regions II′ and III′ are the complimentary stable solutions of regions II and III, respectively. The ''boundary'' cases are given by $\tau^2=4\Delta$ (degenerate nodes and lines of fixed points) and $\tau=0,\ \Delta>0$ are ''centers'' giving periodic orbits in the phase plane. A complete discussion of these cases will not be given here, we simply use these results. Refer, for instance, to Strogatz[6] for additional analysis and details.

Evaluating the matrix **A** in (15) at each fixed point of (12), the following classifications are obtained:

$$A|_{\vec{x}_1^*}=\begin{pmatrix}0 & 1\\ \sqrt{1-6\sigma} & 0\end{pmatrix}\Rightarrow\underbrace{\tau=0;\Delta=-\sqrt{1-6\sigma}}_{\text{''Saddle Node''}}$$
$$A|_{\vec{x}_2^*}=\begin{pmatrix}0 & 1\\ -\sqrt{1-6\sigma} & 0\end{pmatrix}\Rightarrow\underbrace{\tau=0;\Delta=+\sqrt{1-6\sigma}}_{\text{''Center''}}, \tag{19}$$

corresponding to a ''saddle'' and ''center-node'' fixed point, respectively (see Fig. 1 for the placement of these points). As previously discussed, the linear stability analysis gives an exponential solution about each fixed point with the phase-plane trajectories shown in Fig. 2 [the directions follow from (11) and are indicated by arrows].

Physically, trajectories about the center node correspond to elliptical orbits and we will use this fact in a moment to obtain the value for perihelion precession. However, the

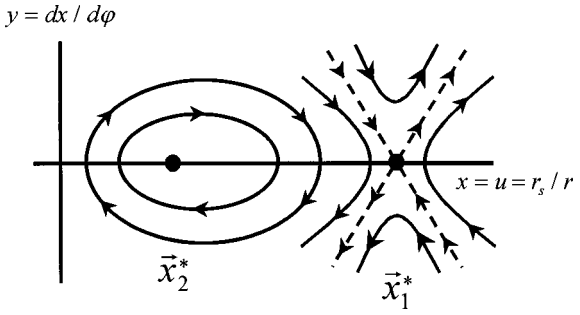

Fig. 2. Linear stability phase plane.

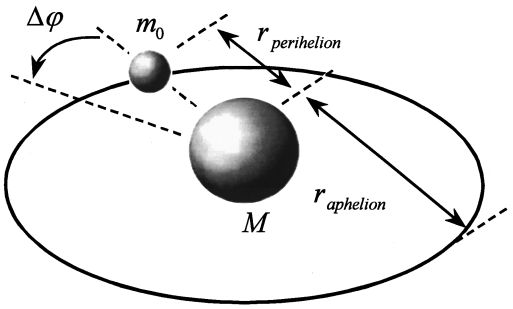

Fig. 3. Schematic of perihelion precession.

saddle node that appears is not predicted by Newtonian theory, but is due to an unstable orbital radius originating from the r^{-3} term of the effective potential derived from (8) [see also (33) and the Appendix]:

$$\hat{V}_{\text{eff}}^2=(1+x^2/2\sigma)(1-x). \tag{20}$$

As a result of this instability, there are orbital effects not present in the Newtonian theory which have been summarized in the literature (see, e.g., MTW,[1] p. 637). But what is nice in the phase-plane approach is that this result comes out very quickly in the analysis as a secondary fixed point (see Sec. V).

IV. PERIHELION PRECESSION

Physically, perihelion precession means that the distance of closest approach (i.e., the perihelion; *-helion* refers to the sun) between two orbiting bodies begins to revolve about the orbit in the same sense as the orbiting body (see Fig. 3). Therefore, planetary orbits do not close but are separated by a small correction, $\Delta\varphi$, after completing a single orbit. For the planet Mercury this is observed to be approximately 574 arcsec per century with 532″ accounted for by a Newtonian analysis of external planetary perturbations. However, there are 43″ not explained from a Newtonian analysis, but predicted to within experimental error by Einstein's theory.

To perform the perihelion calculation simply solve (15) about $\vec{x}_2^*$. This first-order system is rewritten here as

$$\delta x'=\delta y,\quad \delta y'=-\omega^2\delta x,\quad \omega=(1-6\sigma)^{1/4}. \tag{21}$$

The solutions are centers corresponding to precessing elliptical orbits (see Fig. 2):

$$\begin{aligned}\delta x(\varphi)&=A\cos\omega\varphi+B\sin\omega\varphi,\\ \delta y(\varphi)&=-\omega A\sin\omega\varphi+\omega B\cos\omega\varphi,\end{aligned} \tag{22}$$

with A and B arbitrary constants. Choosing initial conditions at the position of perihelion:

$$\delta x(0)=u_0,\quad \delta y(0)=du(0)/d\varphi=0, \tag{23}$$

(22) becomes

$$\begin{aligned}\delta x(\varphi)&=u(\varphi)=u_0\cos\omega\varphi,\\ \delta y(\varphi)&=u'(\varphi)=-\omega u_0\sin\omega\varphi,\end{aligned} \tag{24}$$

giving a typical ''center'' solution about the fixed point $\vec{x}_2^*$.

As previously discussed, in ''physical'' space the orbit of m_0 about M does not close. However, the phase-plane trajectory given by (24) must close after a single orbit since the system is conservative (ignoring radiative effects). There-

fore, the period of a single orbit, Φ, is defined from the period of the phase space trajectory given by (24):

$$\omega\Phi = 2\pi. \tag{25}$$

Solving for Φ, and then substituting for ω in the limit of small σ gives the following result:

$$\Phi = 2\pi\omega^{-1} \approx 2\pi + 3\pi\sigma. \tag{26}$$

The Newtonian calculation gives only the first term, $\Phi = 2\pi$, as expected. However, as seen in (26), the Schwarzschild solution gives the correction:

$$\Delta\varphi = 3\pi\sigma = 6\pi(GMm_0/cJ)^2, \tag{27}$$

which is the usual value (expressed in MKS units) obtained from procedures (A) or (B) [compare also with (9) for the standard expression in the ''geometrized'' system of units].

V. AN ''EXACT'' PHASE PLANE

In Sec. IV a linear stability analysis has been considered about each fixed point. However, this procedure gives only ''local'' information on the general relativistic orbits, and is in fact one shortcoming of the linear stability analysis. Therefore, no correspondence can be made with parabolic, hyperbolic, or orbits near the black hole event horizon using Fig. 2 alone. However, since the equations of motion are integrable due to so many constants of motion [i.e., (4)–(6)], an exact phase plane can be constructed and then several ''global'' features of these orbits may be deduced as a result. In addition, other qualitative features of the Schwarzschild orbital dynamics may be derived from this diagram (Fig. 4) as discussed below. (Note: Since the equations are integrable no chaos exists here. However, if additional degrees of freedom are allowed, the possibility for chaos exists; see, e.g., Ref. 8 for a discussion of chaos in relativistic orbital dynamics).

To obtain an exact phase-plane diagram, consider the ''level curves'' found by taking the ratio of x' and y' from (11), and then integrating to get a conserved quantity:

$$dy/dx = (\tfrac{3}{2}x^2 - x + \sigma)/y \Rightarrow y^2 = \beta + x^3 - x^2 + 2\sigma x. \tag{28}$$

The value of the constant β is easily found by comparison with (8):

$$\beta = 2\sigma(\hat{E}^2 - 1), \tag{29}$$

so that (28) may be alternatively expressed:

$$\hat{E}^2 - 1 = (y^2 + x^2 - x^3)/2\sigma - x. \tag{30}$$

In Fig. 4, the level curves corresponding to different values of $\hat{E}$ in (30) are shown with the effective potential (20) (with $\sigma = \frac{1}{9}$). These curves are exact solutions of (11) for various energies and initial conditions, and should be compared with the approximate solutions given by the linear stability analysis of Fig. 2. The vertical dotted line at $r_S/r = 1$, labels the black hole event horizon.

The value of σ used in (30) has been greatly exaggerated to better illustrate the qualitative features of the exact phase plane. For a more realistic value of σ consider Mercury's orbit—taking the value of $\Delta\varphi$ over a single orbit and then using (27):

$$3\pi\sigma \approx 0.104'' \Rightarrow \sigma \approx 5.4\times 10^{-8}, \tag{31}$$

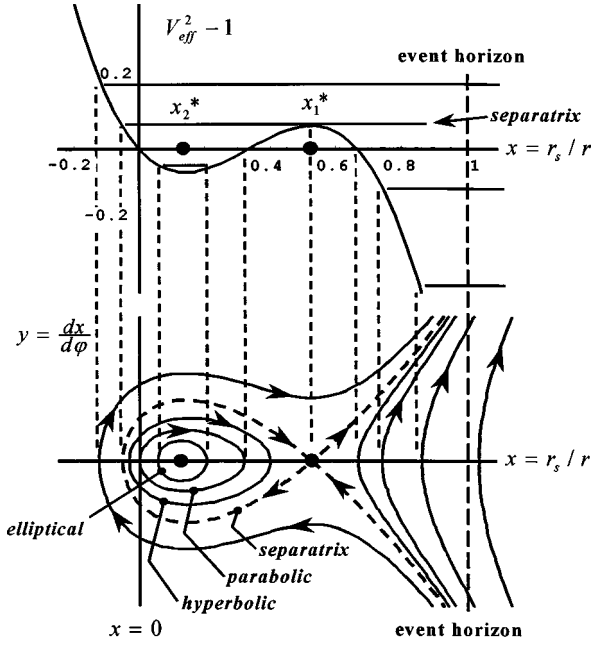

Fig. 4. An ''exact'' phase plane for $\sigma = 1/9$.

or for the binary pulsar system discovered by Hulse and Taylor:[9]

$$3\pi\sigma \approx 4° \Rightarrow \sigma \approx 7.4\times 10^{-3}. \tag{32}$$

To check that $\sigma = \frac{1}{9}$ is a reasonable value in Fig. 4, an upper bound may be placed on σ for the existence of stable or unstable orbits from either phase-plane fixed point. By inspection of (12), if $\sigma > \frac{1}{6}$ then no (real) fixed points exist for a given value of energy and angular momentum. To trace the physical origin of this value and to understand the topological structure of Fig. 4 from a more general viewpoint, note that when $y = 0$ in (30) the effective potential is obtained:

$$\hat{V}_{\text{eff}}^2 - 1 = (x^2 - x^3)/2\sigma - x. \tag{33}$$

The locations of the stable and unstable orbits are found as usual by solving $\partial_x \hat{V}_{\text{eff}} = 0$ for x, which gives identically (12). From (12), no extrema exist for $\sigma > \frac{1}{6}$, establishing an upper bound on σ for stable or unstable orbits. For $\sigma = \frac{1}{6}$, stable orbits (smallest value of) and unstable orbits (largest value of) coincide at

$$r_{1,2} = 3r_S, \tag{34}$$

providing an inflection point in the plot of $\hat{V}_{\text{eff}}^2 - 1$ vs x as shown in Fig. 5 for several values of σ (Ref. 10) (Note: The standard presentations on this diagram are commonly displayed as $\hat{V}_{\text{eff}}^2 - 1$ vs $1/x$; see, e.g., Wald, Ohanian, and Ruffini, or MTW[1] for an alternative parametrization using r_S and J/m_0c^2). But another critical value of σ occurs when $\hat{E}^2 = 1$. To see this, solve $\hat{E}_1^2 - 1 = 0$ using (14) to get $\sigma = \frac{1}{8}$, as displayed in Fig. 5. Therefore, qualitatively distinct orbits exist based upon the following values of σ:

$$0 < \sigma < \tfrac{1}{8}, \quad \sigma = \tfrac{1}{8}, \quad \tfrac{1}{8} < \sigma < \tfrac{1}{6}, \quad \sigma = \tfrac{1}{6}, \quad \sigma > \tfrac{1}{6}. \tag{35}$$

For $\sigma > \frac{1}{6}$ there is insufficient angular momentum for m_0 to sustain an orbit, therefore the mass simply falls into M and correspondingly, $\hat{V}_{\text{eff}}$ has no extrema. The physical signifi-

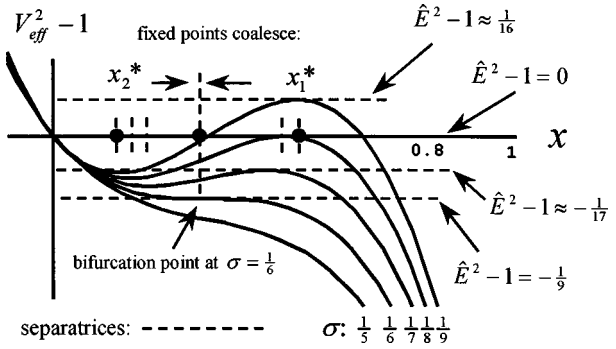

Fig. 5. Schwarzschild effective potential.

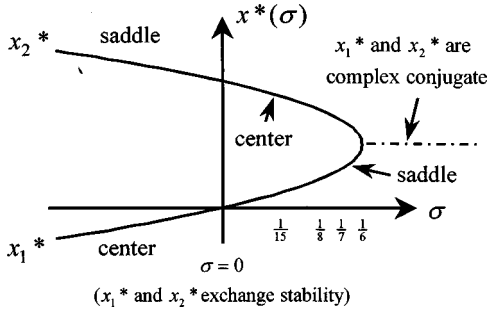

Fig. 7. Schwarzschild bifurcation diagram.

cance of $\sigma = \frac{1}{6}$ is discussed above (34). The physical meaning of the other values in (35) are understood by analyzing the separatrix[11] structure of (30). Essentially, this corresponds to a limitation placed upon the types of orbits that may exist before an unstable orbit is reached and the kinematic classification of the separatrices as distinct unstable orbits.

In essence, the separatrix gives a graphic representation of the critical relationship between energy and angular momentum at the unstable orbital radius (see Fig. 6). For a given angular momentum (σ), the critical energy of the unstable orbit is calculated from $\hat{E}_1$ of (14). For the values of σ plotted in Fig. 5, these energies are computed and marked with horizontal lines. Substituting these values of $\hat{E}^2 - 1$ into (30), the separatrices corresponding to (35) are plotted in Fig. 6.

These distinct separatrices divide the phase plane into four regions of motion for $0 < \sigma < \frac{1}{6}$ ($\sigma = \frac{1}{9}$ is just one special case in Fig. 4). To begin, consider Fig. 4 in the region surrounding the stable fixed point $\vec{x}_2^*$. The oval trajectory in this region corresponds to an elliptical orbit and was used earlier to find the value for perihelion precession. A unique parabolic orbit occurs as the phase-plane trajectory just touches the y axis and separates the hyperbolic and elliptic orbits. The hyperbolic orbit[12] is characterized by a trajectory approaching M from infinity, but then returning to infinity with constant $dr/d\varphi$. Therefore, the separatrix of Fig. 4 (typical

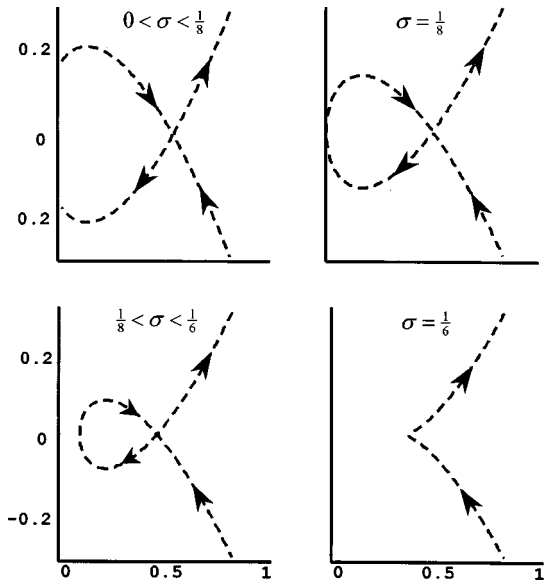

Fig. 6. Separatrices for selected values of σ.

for $\sigma < \frac{1}{8}$) corresponds to a critically unstable hyperbolic orbit that separates trajectories spiraling into M (above the separatrix) or escaping to infinity (below the separatrix).[13]

Similarly, as illustrated in Fig. 6 these separatrices are summarized according to the following values of σ as distinct unstable orbits:

$$\begin{aligned} &0 < \sigma < \tfrac{1}{8} \Rightarrow \text{unstable hyperbolic,} \\ &\sigma = \tfrac{1}{8} \Rightarrow \text{unstable parabolic,} \\ &\tfrac{1}{8} < \sigma < \tfrac{1}{6} \Rightarrow \text{unstable elliptic.} \end{aligned} \tag{36}$$

It is obvious from Fig. 6 that for σ in the range: $\frac{1}{8} \leqslant \sigma < \frac{1}{6}$, only elliptical orbits are possible (about $\vec{x}_2^*$) before the unstable orbit is reached, while the case $0 < \sigma < \frac{1}{8}$ allows all three: hyperbolic, parabolic, and elliptic as discussed above. But these results are consistent with the orbital motion obtained from inspection of the effective potential for different values of σ in Fig. 5. These qualitative differences over the range of unstable orbits have not been pointed out in the literature, but this is not to imply that the Schwarzschild orbital dynamics are poorly understood; see, e.g., Chandrasekhar[14] for an alternative but lengthy analysis.

It should be noted that a physical orbit corresponding to the separatrix can never be achieved in finite proper time. To do so would imply that the phase-plane trajectories change direction at $\vec{x}_1^*$, which is not possible in a deterministic system. To see this, consider the proper time equivalent of (30) [this is (7) after rewriting the equation using the definition of σ in (9) and again using $x = r_S/r$]:

$$(dr/ds)^2 = \hat{E}^2 - 1 + (x^3 - x^2)/2\sigma + x. \tag{37}$$

Separating variables gives an elliptic integral:

$$c\tau = \pm \int dr/\sqrt{(\hat{E}^2 - 1) + x - x^2/2\sigma + x^3/2\sigma}, \tag{38}$$

which diverges to $\pm\infty$ as r approaches the unstable orbital radius r_1 of (12) [and (14) is substituted for $\hat{E}^2 - 1$], i.e., for a particle approaching the saddle point along the separatrix.

From the separatrix analysis it is apparent that a bifurcation occurs at the critical value $\sigma = \frac{1}{6}$, i.e., the topological structure of the phase plane changes as the two fixed points move together, coalesce into a single fixed point, and then disappear from the phase plane as σ is further increased above the critical value 1/6. Therefore, the Schwarzschild orbital dynamics may be interpreted and analyzed as a conservative 2-*d* bifurcation phenomena. Specifically, this bifurcation is a saddle-center bifurcation[15] (see Fig. 7), and summarizes the range of physically possible orbits that may

occur as the energy and angular momentum are varied for $\sigma>0$. But from a more general viewpoint one should also consider negative values of σ [although it is clear that $\sigma<0$ has no physical interpretation since σ must be positive definite according to (9); note also that $\sigma=0$ in (11) gives the phase-plane equations for light rays—see (44) below]. For $\sigma<0$ the two fixed points [Eq. (12)] change stability at $\sigma=0$ as shown in Fig. 7. Therefore, another (transcritical) bifurcation occurs at $\sigma=0$ (see, e.g., Strogatz,[6] pp. 50–52), followed by the saddle-center bifurcation at $\sigma=\frac{1}{6}$.

Finally, an interpretation of the phase-plane trajectories to the right of the separatrix should be given, namely those trajectories leaving and then returning through the event horizon. These trajectories are clearly nonphysical since it is impossible for any classical particle or light ray to escape from within the black hole horizon. The origin of these trajectories may be understood as a consequence of the symmetry of (11) under the interchange: $\varphi\rightarrow-\varphi$; $y\rightarrow-y$, where $\varphi\rightarrow-\varphi$ is due to the time-reversal symmetry of the Schwarzschild dynamics. As a consequence, this system is classified as *reversible* and gives the symmetry of Fig. 4 (and Fig. 6) about the x axis, but with the vector field below the x axis reversing direction.[16]

VI. PROPER AND COORDINATE TIME PHASE DIAGRAMS

In the standard analysis on relativistic orbital dynamics, the proper time parameter is replaced by the equatorial angle as the independent variable. One advantage of this replacement is to simplify the algebra of a perturbative analysis, and is a carryover from the standard techniques applied in the Newtonian case (see the Appendix). However, as far as the phase-plane analysis is concerned, there are no essential difficulties analyzing the dynamics using the proper time (or coordinate time) as independent variables. In fact, there is additional information available which also gives a nontrivial introduction to dynamical invariance.

To demonstrate the invariance of the effective potential between the proper and coordinate time reference frames, start with (7) to obtain the proper time result (Note: Using φ rather than τ eliminates the x^4 leading term appearing below):

$$(r_S/c)^2\dot{x}^2=x^4[\hat{E}^2-(1+x^2/2\sigma)\Lambda]. \tag{39}$$

The corresponding coordinate time expression is obtained using: $\dot{x}=(dx/dt)\dot{t}$, in combination with (5) which gives

$$(r_S/c)^2(dx/dt)^2=x^4(\Lambda/\hat{E})^2[\hat{E}^2-(1+x^2/2\sigma)\Lambda]. \tag{40}$$

Solving (40) for $\hat{E}^2$ gives the coordinate frame expression for the total energy:

$$\hat{E}^2=\frac{x^4(1+x^2/2\sigma)\Lambda^3}{x^4\Lambda^2-(r_S/c)^2(dx/dt)^2}. \tag{41}$$

By inspection of (41), as $dx/dt\rightarrow0$, the effective potential (20) is recovered, i.e., $\hat{V}^2_{\rm eff}\rightarrow\hat{V}'^2_{\rm eff}=\hat{V}^2_{\rm eff}$ is invariant between the proper and coordinate time reference frames. Therefore, the dynamical structure is invariant, or alternatively stated, the extrema of $\hat{V}^2_{\rm eff}$ are identical in either reference frame. However, the phase diagrams in each case are not identical due to the existence of an additional ''frame-dependent'' fixed point that appears in the coordinate reference frame at the event horizon (see Fig. 8).

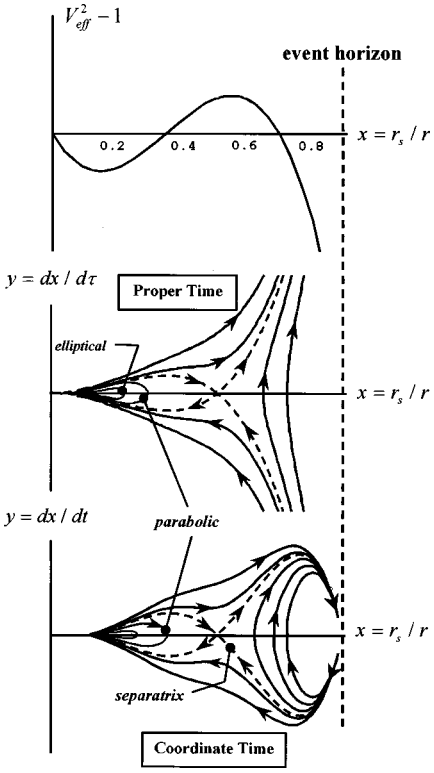

Fig. 8. Proper and coordinate time phase diagrams.

To summarize these results, the corresponding phase-plane equations analogous to (11) in both the proper and coordinate time reference frames are derived by differentiating (39) and (40), respectively. In each case the results are given by

$$\begin{aligned} dx/d\tau&=y=\pm x^2[\hat{E}^2-(1+x^2/2\sigma)\Lambda]^{1/2},\\ dy/d\tau&=x^3[7x^3-6x^2+10x\sigma+8\sigma(\hat{E}^2-1)]/4\sigma, \end{aligned} \tag{42}$$

and

$$\begin{aligned} dx/dt&=y=\pm x^2\Lambda^{1/2}[\hat{E}^2-(1+x^2/2\sigma)\Lambda]^{1/2}/\hat{E},\\ dy/dt&=x^3\Lambda[9x^4-15x^3+2x^2(3+7\sigma)\\ &\quad+2x\sigma(6\hat{E}^2-11)-8\sigma(\hat{E}^2-1)]/4\sigma. \end{aligned} \tag{43}$$

Although (42) and (43) are more complicated algebraically than (11), the simultaneous solution of $\dot{x}=\dot{y}=0$ for $\hat{E}^2$ and x in each case reduces to (12) and (14) identically, but with another fixed point, $x=0$, at infinity and at $x=1$ in the case of (43). However, the fixed point at infinity exists for the Newtonian case as well, and is discussed in the Appendix. The fixed point at the event horizon is obviously coordinate dependent and does not correspond to any extrema of the effective potential. Nevertheless, this fixed point has physical consequences for observers in the coordinate reference frame—explaining the slowing down of objects and redshift of signals approaching the event horizon.

As discussed below (14), there are additional nonphysical roots obtained when solving $\dot{x}=\dot{y}=0$, only for x and y. The nonphysical nature of these fixed points is due to the fact that there must be a constraint placed upon $\hat{E}$ when τ or t is used

as the independent variable. Solving simultaneously the expressions for y given in (42) or (43) gives the proper constraint on $\hat{E}$ and as a result forces these fixed points to coincide with the extrema of the effective potential. This is also a feature of the Newtonian dynamics when using t as the independent variable.

VII. LIGHT RAYS

The analysis of photon orbits in the Schwarzschild spacetime is a straightforward application of the techniques discussed for timelike orbits. For light rays, $d\tau=0$, which in turn implies that both E and J are divergent from (5) and (6), although their ratio remains finite. As a result, $\sigma\to 0$, and the phase-plane equations for light rays follow as a special case of (11):

$$x' = y = \pm[1/b^2 - x^2(1-x)]^{1/2},$$
$$y' = \tfrac{3}{2}x^2 - x, \tag{44}$$

where $1/b^2 \equiv 2\sigma\hat{E}^2$ is a constant expressing the dimensionless impact parameter, b, as the finite ratio of $\hat{E}$, J, and r_S.

The simultaneous solution of $x'=y'=0$ for $1/b^2$ and x results in two fixed points and the corresponding values of the impact parameter:

$$\{x_1=\tfrac{2}{3}; 1/b^2=\tfrac{4}{27}\}, \quad \{x_2=0; 1/b^2=0\}, \tag{45}$$

giving the standard results for the unstable orbital radius, x_1, and the impact parameter at which this instability occurs. The fixed point, x_2, is a center node (at infinity) about which the hyperbolic orbits ''precess'' (see Ref. 12 for a comment on the timelike case) and gives the standard result for light bending. Therefore, the perihelion precession of timelike orbits and light bending are actually special cases of one another: In the timelike case this center node fixed point is at finite r and allows ''real'' circular orbits; but for light rays this fixed point moves to infinity and gives the precessing hyperbolic orbits noted above. However, a phase-plane calculation of light bending analogous to that discussed in Sec. III does not work here. This is due to the fact that a linear stability analysis (15) ''kills'' the necessary terms; namely, the impact parameter disappears from the matrix **A** (a similar result occurs when calculating the period of a simple pendulum for large angles using this technique).

The phase-plane level curves for light rays in Fig. 9 correspond to different values of $1/b^2$. These are shown together with the locations of the fixed points and photon effective potential: $x^2(1-x)$. The most striking difference between the photon and timelike phase-plane dynamics (comparing Figs. 9 and 4) is that the center node fixed point moves to the origin as $\sigma\to 0$ (as discussed above). As a result, circular photon orbits do not exist in any ''dynamical'' sense, but become circular in geometry as the orbits approach the separatrix. To see this use the definition of y in the first equation of (44), and then separating variables shows that $\varphi\to\infty$ as $x\to x_1$ and $1/b^2\to\frac{4}{27}$ [this result is analogous to the proper time divergence pointed out in (38)]. Therefore, the separatrix corresponds to the unstable ''photon sphere'' that is commonly discussed in the literature (see, e.g., Ohanian and Ruffini,[1] p. 410).

The physical interpretation of the various phase-plane regions of Fig. 9 is similar to that of Fig. 4, but there are important differences. For light rays with impact parameter

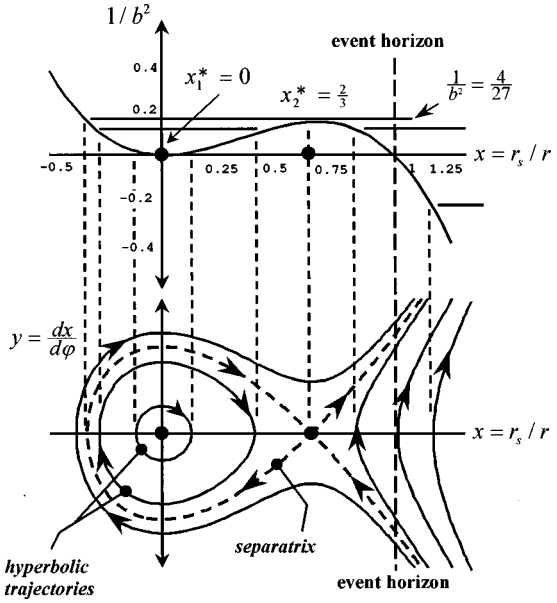

Fig. 9. Null geodesics phase plane.

$1/b^2=0$, these orbits just graze the event horizon from the inside and simultaneously (in an unrelated trajectory) reach the center node fixed point of the effective potential (see Fig. 9). For $1/b^2<0$, b loses its interpretation as an impact parameter since the trajectories in this case originate from the singularity at $r=0$ and lie within the horizon. For $1/b^2 < \frac{4}{27}$, the trajectories are confined to within the separatrix and correspond to the light rays arriving from infinity, reaching a turning point [given by the appropriate root of the first equation in (44)], and then return to infinity as discussed below (45). For $1/b^2 > \frac{4}{27}$, a photon arrives from infinity (above the separatrix) and then falls through the event horizon. The corresponding time-reversed trajectories are given below the separatrix.

The trajectories to the right of the unstable orbital radius of Fig. 9 are also interpreted as time-reversed paths that reach a maximum distance from the event horizon and then return to the singularity. But another interpretation is possible using simple energy considerations: A photon and its time-reversed counterpart originate from a point of maximum distance from the horizon, and then both proceed simultaneously from this point into the horizon. The analogous interpretation is also possible in the timelike case for massive particles (see Fig. 4 and the discussion in the final paragraph of Sec. V). However, these are classical interpretations and should not be identified with quantum phenomena.

VIII. DISCUSSION

We have considered an alternative procedure for calculating the value of perihelion precession and have summarized the Schwarzschild orbital dynamics in the modern setting of phase-plane analysis. Contrasting these calculations with the standard textbook procedures, the main results are obtained very quickly while minimizing the algebra, but placing more emphasis on the physics. For example, by calculating the value for perihelion precession using a perturbative solution, a departure is made from an analysis based on physical concepts to an exercise in algebra. However, in the phase-plane approach, the physical concepts are given greater emphasis

and made more accessible to beginning students. This is due to the fact that the phase-plane technique itself is based essentially on the ''energy-method'' diagrams taught in introductory mechanics courses. The analysis presented in Sec. VI demonstrates that important topics such as dynamical invariance are easily handled using the phase-plane techniques. These provide nontrivial and physically interesting examples which normally are difficult conceptually for beginning students.

In addition, the traditional analysis of the effective potential could be augmented with discussion on the ''exact'' Schwarzschild phase plane, or more specifically its separatrix structure. Essentially, the separatrix gives a geometric representation of the critical relationship occurring between energy and angular momentum, and as such, divides the phase plane into physically distinct regions of motion. By varying a dimensionless parameter involving the angular momentum, a saddle-center bifurcation occurs as the two fixed points coalesce and disappear—altering the phase-plane topology. For the case of light rays the separatrix corresponds physically to an unstable ''photon sphere'' as discussed earlier. As a special case of (11), the photon orbits also provide a transcritical bifurcation point of the dynamics—exchanging stability at $\sigma=0$.

For additional applications it would be interesting to analyze solutions other than the Schwarzschild case, e.g., the Reissner–Nordstrom (a charged, spherically symmetric black hole), the Kerr solution (a rotating black hole), or the Kerr–Newman solution (a charged, rotating black hole). Further applications would include an analysis of cosmological solutions and nonconservative orbital dynamics (i.e., systems emitting gravitational radiation), and also solutions stemming from alternative theories of general relativity. Analysis of these topics will appear elsewhere.

In summary, constructing an exact phase plane for an arbitrary solution will only be possible if the fixed point algebraic equation, $x'=y'=0$, is of fourth order or less (and in addition if a sufficient number of first integrals exists). Otherwise, finding roots will be difficult if not impossible. However, a numerical approach could always be taken, and would be motivated by the interesting pictures that result from combining the fixed point structure of general relativity state space into a diagram that includes the event horizon.

ACKNOWLEDGMENTS

I thank Jim Crawford and Richard Treat for many helpful comments and discussion on the analysis. I also thank Charles Jaffe and Steve Strogatz for pointing out references, and Boyd Edwards for discussion in Physics 401A, Spring 1997. In addition, I thank Abhay Ashtekar, Beverly Berger, Bryce DeWitt, Bill Hiscock, Ted Newman, Jorge Pullin, Carlo Rovelli, Lee Smolin, and Clifford Will for helpful comments on the existing literature.

APPENDIX: NEWTONIAN PHASE-PLANE ANALYSIS

As discussed earlier in Sec. I, the standard analysis of the Newtonian orbital dynamics is based on the change of independent variable, $t\rightarrow\varphi$, for the purpose of finding a closed form solution describing the orbital geometry. But a phase-plane analysis of the differential equations using time as the independent variable is no more complicated in principle than using φ. Furthermore, there are results shared by the relativistic case (discussed in Sec. VI) that are clarified in this analysis.

To begin, consider the Newtonian limit of the equations derived in Sec. II. The effective potential/(unit rest energy) is given in (20), and is defined as $\hat{V}_{\rm eff}^2$ which gives the proper Newtonian limit for $\hat{V}_{\rm eff}$ (to within an additive constant) in the limit of large r. As a result, the Newtonian limit of (20) is given by

$$\hat{V}_{\rm eff}=[1-x+(x^2-x^3)/2\sigma]^{1/2}\approx 1-x/2+x^2/4\sigma, \tag{46}$$

which differs from the standard Newtonian form by an additive constant (corresponding to the rest mass energy of m_0). The standard Newtonian effective potential energy is chosen to be zero at infinity, giving the usual expression:

$$\hat{V}_{\rm eff}=x^2/4\sigma-x/2, \tag{47}$$

compared to the relativistic limit where the energy at infinity corresponds to the rest mass energy. But this additive constant is of little consequence insofar as the dynamics are concerned, and so we adopt (47) for the remaining discussion [Note: From a certain viewpoint (see Kompaneyets,[4] p. 44) one may regard this difference as a choice of ''gauge:'' i.e., the ''Newtonian gauge'' takes $\hat{V}_{\rm eff}=0$ at infinity, while the special ''relativistic gauge'' is $\hat{V}_{\rm eff}=1$].

The corresponding Newtonian expression for (39) is derived using the standard Lagrangian and Hamiltonian results:

$$(r_S/c)^2\dot{x}^2=2x^4[\hat{E}-\hat{V}_{\rm eff}], \tag{48}$$

where $\hat{V}_{\rm eff}$ is given by (47) and $x=r_S/r$. Although the choice of units seems odd at first, this form gives the most straightforward comparison with the relativistic case. As a check, (48) reduces to [after substituting (2), (47), and then (9)]

$$(du/dt)^2=-u^4(J/m)^2[u^2-2u_0u-b_u^2], \tag{49}$$

where $u=1/r$ and $u_0=GMm^2/J^2$ gives the standard radius of a circular orbit. The constant: $b_u^2\equiv 1/b^2=2mE/J^2$ expresses the impact parameter (for a particle approaching from infinity) in terms of E and J. The zeroes of (49) give the standard turning points of the effective potential (aside from $u=0$). Furthermore, substituting $u=u(\varphi)$ and then (6) into (49) (the Newtonian expression for J is identical in form to the relativistic case) leads to the standard second-order differential equation that is commonly evaluated for the analysis of these orbits.

Continuing with the analysis, differentiating (48) gives the dimensionless phase-plane equations expressed using t as the independent variable:

$$\begin{aligned}\dot{x}&=y=\pm x^2c[4\sigma\hat{E}+2\sigma x-x^2)]^{1/2}/r_S\sqrt{2\sigma},\\ \dot{y}&=-x^3c^2(3x^2-5\sigma x-8\sigma\hat{E})/2\sigma r_S^2.\end{aligned} \tag{50}$$

Solving simultaneously, $\dot{x}=\dot{y}=0$ for $\hat{E}^2$ and x then gives the two fixed points:

$$\{x_1=\sigma;\hat{E}=-\sigma/4\},\quad x_2=0. \tag{51}$$

The first gives the standard results: a center node corresponding to a Newtonian circular orbit with radius r_1 and energy given by

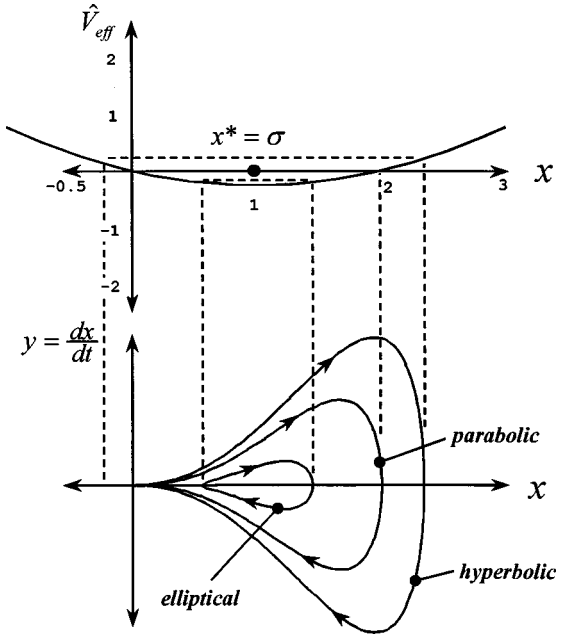

Fig. 10. Newtonian phase-plane diagram with t as the independent variable ($\sigma = 1$).

$$x_1 = \sigma \Rightarrow r_1 = r_S/\sigma = J^2/GMm^2,$$
$$\hat{E} = -\sigma/4 \Rightarrow E = -m(GMm/J)^2/2. \tag{52}$$

The second fixed point at infinity simply expresses the fact that it takes an infinite amount of time for the orbiting particle, m, to reach the turning point at infinity (in the case of parabolic and hyperbolic orbits)—a fixed point that is shared in the relativistic orbital dynamics. For comparison with the relativistic phase-plane results the Newtonian phase diagram for (50) is shown in Fig. 10.

[a]Present address: NASA Goddard Space Flight Center, Mailstop: 551.0, Greenbelt, MD 20771.

[1]M. Carmeli, *Classical Fields: General Relativity and Gauge Theory* (Wiley, New York, 1982), pp. 218–220; A. Eddington, *The Mathematical Theory of Relativity* (Cambridge U.P., Cambridge, 1924), 2nd ed., pp. 88–90; C. Misner, K. Thorne, and J. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), pp. 659–670, 1110–1116; H. Ohanian and R. Ruffini, *Gravitation and Spacetime* (Norton, New York, 1994), pp. 401–408; B. Schutz, *A First Course in General Relativity* (Cambridge U.P., Cambridge, 1990), pp. 275–284; H. Stephani, *General Relativity* (Cambridge U.P., Cambridge, 1990), pp. 105–108; R. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984), pp. 139–143; S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972), pp. 194–197, 230–233.

[2]N. Anderson and G. R. Walsh, ''Phase-Plane Methods for Central Orbits,'' Am J. Phys. **58** (6), 548–551 (1990).

[3]B. Davies, ''Derivation of Perihelion Precession,'' Am. J. Phys. **51** (10), 909–911 (1983); D. Ebner, ''Comment on B. Davies's 'Derivation of Perihelion Precession, ' '' *ibid.* **53** (4), 374 (1985); Daniel R. Stump, ''Precession of the Perihelion of Mercury,'' *ibid.* **56** (12), 1097–1098 (1988).

[4]A. P. Arya, *Introduction to Classical Mechanics* (Prentice–Hall, Englewood Cliffs, NJ, 1990), pp. 222–252; H. C. Corben and P. Stehle, *Classical Mechanics* (Dover, Mineola, NY, 1994), 2nd ed., pp. 90–100; G. Fowles and G. Cassiday, *Classical Dynamics of Particles and Systems* (Harcourt-Brace, Orlando, FL, 1993), 5th ed., pp. 191–216; H. Goldstein, *Classical Mechanics* (Addison–Wesley, Reading, MA, 1980), 2nd ed., pp. 94–102; A. S. Kompaneyets, *Theoretical Mechanics* (Dover, New York, 1962), 2nd ed., pp. 41–48; J. Marion and S. Thornton, *Analytical Mechanics* (Harcourt–Brace, Orlando, FL, 1995), 4th ed., pp. 291–321; A. Roy, *Orbital Dynamics* (Hilger, Redcliffe Way, Bristol, 1982), 2nd ed., pp. 69–100; K. Symon, *Mechanics* (Addison–Wesley, Reading, MA, 1971), pp. 128–134.

[5]J. L. Martin, *General Relativity—A First Course for Physicists* (Prentice–Hall International, Hertfordshire, UK, 1996), Revised ed., pp. 58–60.

[6]S. Strogatz, *Nonlinear Dynamics and Chaos—with Applications to Physics, Biology, Chemistry, and Engineering* (Addison–Wesley, New York, 1994). Also see problem #6.5.7, p. 186.

[7]M. Tabor, *Chaos and Integrability in Nonlinear Dynamics: An Introduction* (Wiley, New York, NY, 1989), pp. 13–30.

[8]L. Bombelli and E. Calzetta, ''Chaos around a Black Hole,'' Class. Quantum Grav. **9** (12), 2573–2599 (1992); S. Suzuki and K. Maeda, ''Chaos in Schwarzschild Space-Time: The Motion of a Spinning Particle,'' Phys. Rev. D **55**, 4848–4859 (1997).

[9]R. A. Hulse and J. H. Taylor, ''Discovery of a Pulsar in a Binary Star System,'' Astrophys. J. **195**, L51–L53 (1975).

[10]Using x rather than $1/x$ for the horizontal axis pushes the singularity at $r = 0$ to infinity. As a result, the relative locations of fixed points are more easily scaled and plotted in the phase plane.

[11]Also termed *homoclinic* orbit in the literature on nonlinear analysis.

[12]These are actually *precessing* hyperbolic orbits if one allows negative r values (although nonphysical—shown to the left of the y axis in Fig. 4).

[13]But for an ordinary massive object M, the trajectories never reach the event horizon as it lies within the object's surface.

[14]S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Oxford U.P., Clarendon, New York, 1983), pp. 96–122.

[15]J. Hale and H. Kocak, *Dynamics and Bifurcations* (Springer-Verlag, New York, 1991), p. 425; J. M. Mao and J. B. Delos, ''Hamiltonian Bifurcation Theory of Closed Orbits in the Diamagnetic Kepler Problem,'' Phys. Rev. A **45**, 1746–1761 (1992).

[16]Incidentally, $\vec{x}_2^*$ of Fig. 4 (and Fig. 3) is classified as a *nonlinear* center; see Strogatz (Ref. 6, p. 114).

THE AWE FACTOR

The other aspect of science, the one that I am more concerned with, is the wonder, the ''awe factor.'' I ask myself, what is the appeal of religion, what is the appeal of UFOs, what is the appeal of von Däniken or Velikovsky, all that nonsense? I suspect that a part of it is the kind of awesome romance that science ought to be the master of. Don't let us allow religion to walk away with the awe factor. Science has orders of magnitude more to offer in this field. Black holes are incomparably more wondrous, more romantic, than anything you read in the pseudoscientific literature, in New Age drivel, in the ''occult,'' in the Bible. Let's not sell science short.

Richard Dawkins, ''The 'Awe' Factor,'' Skeptical Inquirer **17**(3), 242–243 (1993).