A Nonintegrable Model in General Relativity

Richard Moeckel

School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA

Abstract. The geodesic flow of a perturbation of the Schwarzschild metric is shown to possess a chaotic invariant set. The perturbed metric is a relativistic analogue of Hill's problem in classical celestial mechanics in that it models the effects of a distant third body.

1. Introduction

This paper is concerned with a dynamical system which models the motion of a small mass near a black hole but also subject to the gravitational effects of a third, more distant mass. It is shown that under certain conditions, the system posesses hyperbolic periodic orbits with transverse homoclinic orbits. The resulting chaotic dynamics contrasts sharply with the behavior of the better-known general relativistic systems, which have turned out to be completely integrable.

In classical celestial mechanics, the motion of a small point mass around a large mass is modelled by the Kepler problem. The effects of a third mass could be treated as a perturbation. For example, a perturbing potential due to Hill is sometimes added to describe the effects of a distant third mass [Hill]. The Kepler problem is completely integrable. In fact, it possesses additional constants of motion beyond those required for integrability with the result that all of the bounded orbits are periodic rather than merely quasi-periodic as is typical for integrable systems. However, integrable systems are rare and delicate. One expects that most perturbations will destroy the integrability. This is difficult to show for perturbations of the Kepler problem because the unperturbed problem is so degenerate.

    The relativistic analogue of the Kepler problem is the Schwarzschild problem
which describes the motion of a small point mass around a spherically symmetric
black hole. Because of the symmetry, this problem is also completely integrable,
but it does not possess additional constants of motion like the Kepler problem
does. Several perturbations of the Schwarzschild problem have been studied.
Among them are models of axially symmetric rotating black holes (the Kerr
problem) and of electrically charged black holes (the Reissner-Nordstrom
problem). Oddly enough, these also lead to completely integrable dynamical
systems. One still expects that most perturbations will be nonintegrable. The
problem is that not many physically valid perturbations have been studied. One
reason for this is the difficulty of solving the Einstein equation (the relativistic
analogue of Laplace's equation for the gravitational potential). In particular, a
rigorous description of the effects of a third mass has not been found. In this
paper, an approximate model for these effects is obtained under the same
conditions as the classical Hill's problem. This model is a solution of the
linearized Einstein equation which behaves like the Hill problem in the classical
limit where gravity is relatively weak.
    Once the relativistic model has been constructed, it remains to study the
resulting dynamics. The nonintegrability is demonstrated by an application of
the Melnikov integral method for finding transverse homoclinic orbits. In the
unperturbed Schwarzschild problem there is a hyperbolic periodic orbit with
nontransverse homoclinic orbits (no such feature is present in the Kepler
problem). Certain branches of the stable and unstable manifolds of this orbit
coincide. The Melnikov integral allows one to prove by explicit computation
that the perturbation will split the stable and unstable manifolds apart so that
they intersect transversely. Then the existence of chaotic orbits follows in the
usual way. From the mathematical point of view, an interesting feature of the
computation is the fact that the size of the splitting of the stable and unstable
manifolds turns out to be exponentially small in a certain parameter.
    The paper is divided into sections. Section 2 compares the Kepler problem
with the Schwarzschild problem and establishes the existence of the hyperbolic
periodic orbit and homoclinic orbits for the unperturbed Schwarzschild problem.
This is all well-known. Section 3 contains the construction of the perturbation
analogous to Hill's problem. Section 4 is devoted to the computation of the
Melnikov integral and the observations about exponentially small splitting.
Finally, Sect. 5 contains the rather colorful interpretation of the results in
physical terms.

2. The Kepler Problem vs. the Schwarzschild Problem

The Kepler problem is the classical (non-relativistic) model for the motion of a
point particle under the gravitational influence of a large mass. A large mass, M,
located at the origin in $\mathbf{R}^3$ sets up a gravitational acceleration field which is the
gradient of the Newtonian potential function $\Phi(x) = \dfrac{M}{r}$, where $r = |x|$, the
Euclidean distance from the origin. The Newtonian potential function can be
characterized by four properties: it is time-independent, spherically symmetric,
dies out at infinity, and solves the Laplace equation:

$$\nabla^2 \Phi(x) = 0 .$$

The motion of the point particle is described by the differential equation:

$$\ddot{x} = \nabla \Phi(x).$$

The relativistic analogue of the Kepler problem is known as the Schwarzschild problem. Before describing this it is appropriate to present a brief account of some of the principles of general relativity. For more information on relativity and the Schwarzschild problem, good references are [MTW, O'Neill, and Sch]. A moving particle is described not by a curve $x(t) \in \mathbf{R}^3$ but rather by a so-called worldline in spacetime: $\gamma(\tau) = (t(\tau), x(\tau)) \in \mathbf{R}^4$. Here the parameter τ is the proper time of the particle, that is, the time as measured by a clock carried along with the particle. The other time parameter t turns out to be the time as measured by an observer at rest far from the origin. Gravitational effects are described not by a potential function, but rather by a Lorentzian metric on $\mathbf{R}^4$. Even in the absence of massive particles spacetime is equipped with such a metric, namely the Minkowski metric:

$$\mathbf{g} = -dt \otimes dt + dx^1 \otimes dx^1 + dx^2 \otimes dx^2 + dx^3 \otimes dx^3.$$

The presence of a massive particle changes the metric. To describe it, it is convenient to introduce spherical coordinates (r, θ, ϕ). In these coordinates the Minkowski metric is:

$$\mathbf{g} = -dt \otimes dt + dr \otimes dr + r^2 d\phi \otimes d\phi + r^2 \sin^2 \phi \, d\theta \otimes d\theta.$$

A particle of mass, M, at the origin changes this to:

$$\mathbf{g} = -A(r)dt \otimes dt + A^{-1}(r)dr \otimes dr + r^2 d\phi \otimes d\phi + r^2 \sin^2 \phi \, d\theta \otimes d\theta, \qquad (2.1)$$

where $A(r) = 1 - \dfrac{2M}{r}$. This is known as the Schwarzschild metric. It is characterized by four properties analogous to those which determine the Newtonian potential function: it is time-independent, spherically symmetric, converges to the Minkowski metric at infinity, and solves the Einstein equation:

$$\mathrm{Ric}(\mathbf{g}) = 0,$$

where Ric is the Ricci curvature tensor (a symmetric rank-two tensor obtained by contracting the Riemann curvature tensor). The motion of a point particle under the influence of this mass is described by a worldline $\gamma(\tau)$ which is a geodesic of this metric. More precisely, γ is a unit-speed, time-like geodesic, that is, $\mathbf{g}(\gamma'(\tau), \gamma'(\tau)) = -1$. The differential equation for a geodesic is:

$$\nabla_{\gamma'} \gamma' = 0,$$

where $\nabla_{\gamma'}$ denotes covariant differentiation along the curve.

It is interesting to compare the motions predicted by the classical and relativistic models. Only bounded, planar motions will be considered. The Kepler problem in the plane can be formulated as a Hamiltonian dynamical system of two degrees of freedom. Let

$$H = \tfrac{1}{2}|p|^2 - \Phi(x),$$

where $x \in \mathbf{R}^2$ is the position of the point particle and $p \in \mathbf{R}^2$ is its momentum (in this case $p = \dot{x}$). There are two simple constants of motion for this system, the energy H and the angular momentum $\Omega = r^2 \dot{\phi}$, where r and ϕ are polar co-

ordinates in the plane. Since there are only two degrees of freedom, the system is completely integrable. The well-known procedure for finding the orbits is to use the angular momentum to eliminate ϕ and thereby obtain a system with one degree of freedom for r:

$$\frac{1}{2}\left[\dot{r}^2 + \frac{\Omega^2}{r^2}\right] - \Phi(r) = h\,,$$

where h is the energy. It is also traditional to make the substitution $u = \frac{1}{r}$ and use ϕ instead of t as a parameter. The following equation for $u(\phi)$ is the result:

$$\Omega^2 u'(\phi)^2 = 2h + 2Mu - \Omega^2 u^2\,. \tag{2.2}$$

For $h < 0$ this leads to the familiar elliptical orbits with one focus at the origin.

The geodesic flow of the Schwarzschild metric can also be formulated as a Hamiltonian system, this time with three degrees of freedom. First, it is well-known that geodesic flows are Lagrangian dynamical systems with Lagrangian equal to the kinetic energy:

$$L = \tfrac{1}{2}\mathbf{g}(\gamma', \gamma') = \tfrac{1}{2}[-A(r)t'^2 + A^{-1}(r)r'^2 + r^2\phi'^2]\,.$$

If one introduces the conjugate momenta

$$E = A(r)t'\,, \qquad p = A^{-1}(r)r'\,, \qquad \Omega = r^2\phi'\,,$$

then the Hamiltonian is:

$$H = \frac{1}{2}\left[-A^{-1}(r)E^2 + A(r)p^2 + \frac{\Omega^2}{r^2}\right]\,. \tag{2.3}$$

There are three simple constants of motion: the energy $H = -\tfrac{1}{2}$ (recall the unit speed requirement), the angular momentum Ω and the conjugate momentum of t, E. The equation for $u(\phi)$ is now:

$$\Omega^2 u'(\phi)^2 = E^2 - A(u)(1 + \Omega^2 u^2) = (E^2 - 1) + 2Mu - \Omega^2 u^2 + 2M\Omega^2 u^3\,. \tag{2.4}$$

Comparison with (2.2) shows that, in a sense, the transition from the classical model to the relativistic one is accomplished simply by adding a cubic term to the orbit equation. For small values of u, that is, for large values of r, the two models will yield similar orbits, however, near the mass M, the effects of the change are profound.

The coefficient $A(r)$ in the Schwarzschild metric (2.1) vanishes at $r = 2M$, the Schwarzschild radius. This is the edge of a black hole. By means of clever coordinate changes it is possible to remove the singularity at $r = 2M$ and it is possible for geodesics to cross the edge smoothly. Once inside, however, they are forced to the real singularity at $r = 0$. This can be seen intuitively by noting that for $r < 2M$ the signs of the coefficients of $dt \otimes dt$ and $dr \otimes dr$ are reversed. The radial coordinate behaves in certain ways like a time coordinate and in particular, it is monotonic (decreasing in this case) along geodesics. For the purposes of this paper it is enough to remember that once a geodesic crosses the Schwarzschild radius it goes to meet its maker.

More insight into the differences between the classical and relativistic flows can be obtained by comparing the functions on the right-hand sides of Eqs.

(2.2) and '(2.4). In each case these are of the form: a constant minus a polynomial $P(u)$. The u coordinate is confined to the region where $P(u)$ is less than the constant. Different kinds of orbits are obtained by choosing different values of the constant. For simplicity, it will be assumed without loss of generality that $M = \frac{1}{2}$.

Figure 1 shows $P(u) = -u + \Omega^2 u^2$ for the Kepler problem. For a negative h, the orbit will oscillate between two fixed radii, namely the semiminor and semimajor axes of the ellipse. At the critical point of $P(u)$, these two radii are equal and one obtains a circular periodic orbit.

Figure 2 shows $P(u) = -u + \Omega^2 u^2 - \Omega^2 u^3$ for several choices of the constant Ω. Given the choice of M, the Schwarzschild radius is $r = u = 1$. Any orbit for which u increases beyond 1 disappears into the black hole.

For low angular momentum, $\Omega^2 < 3$, every bounded orbit suffers this fate. At $\Omega^2 = 3$, a bifurcation occurs. For $\Omega^2 > 3$, there are two circular orbits at the critical points of $P(u)$. The minimum is stable while the maximum is hyperbolic. The phase portrait in the (r, p) plane is shown in Fig. 3. One branch of the unstable manifold of the hyperbolic point meets a branch of the stable manifold to form a homoclinic orbit. The other branches lead to the black hole. The
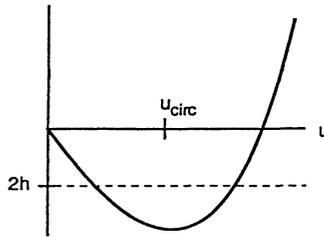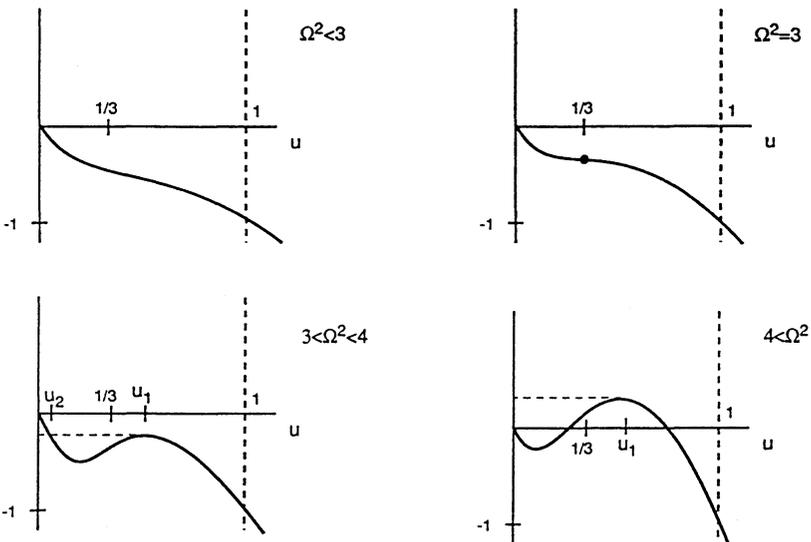

Fig. 1

Fig. 2

homoclinic connection will be exploited later to prove the nonintegrability of a certain perturbation of the Schwarzschild problem. The critical value corresponding to the homoclinic orbit is $E^2 = \frac{8}{9}$ when $\Omega^2 = 3$ and it increases with Ω thereafter. For $\Omega^2 > 4$ the homoclinic orbit disappears because the relevant branches of the stable and unstable manifolds become unbounded.

3. Hill's Problem and a Relativistic Analogue

Hill's problem is a perturbation of the Kepler problem invented by G. W. Hill to describe the motion of the moon around the earth. The idea is to view the earth-moon system as a two-body problem which is perturbed by the gravity of the sun. Since the distance to the sun is large compared to the distance between the earth and the moon, the gravitational effect of the sun is almost the same for the earth as it is for the moon. In other words, the relative acceleration of the two produced by the sun is small. Hill's idea was to expand this relative acceleration as a power series and keep only the dominant terms as the perturbation [Hill].

To derive the form of this perturbation, imagine that the earth and the sun are moving in a circular orbit of the two-body problem. Introduce coordinates centered at the earth which rotate in such a way that the sun remains on the positive x_1-axis at position $X = (R, 0)$. Suppose the moon is at position $x \in \mathbf{R}^2$, distance r from the earth and distance ϱ from the sun (see Fig. 4). Furthermore, let the mass of the earth be $\frac{1}{2}$ and the mass of the sun be M. The acceleration produced by the sun on the moon is $\dfrac{M(X-x)}{\varrho^3}$, while the acceleration produced
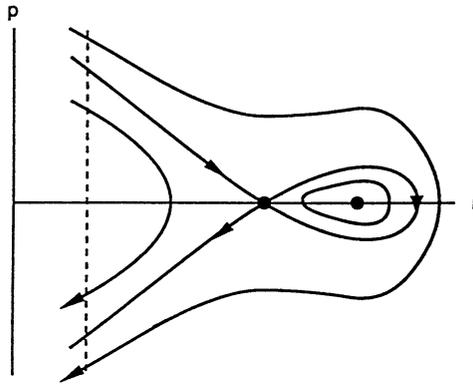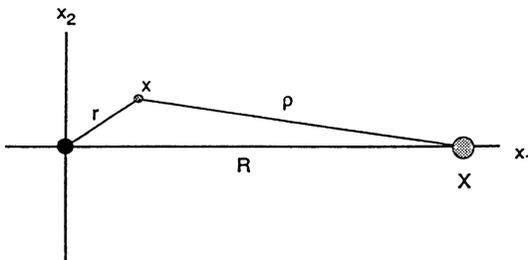

Fig. 3

Fig. 4

by the sun on the earth is $\dfrac{MX}{R^3}$. The relative acceleration is the difference of the two.

Hill's approximation is obtained as follows. First, introduce polar coordinates (r, ϕ). Then $\varrho^2 = R^2 + r^2 - 2Rr \cos\phi$. Expanding $\dfrac{1}{\varrho^3}$ in powers of $\dfrac{r}{R}$ and substituting only the constant and linear terms into the relative acceleration gives the approximation (back in rectangular coordinates):

$$\frac{M}{R^3}\left[\begin{pmatrix} 3x_1 \\ 0 \end{pmatrix} - \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right]$$

for the relative acceleration. This is the gradient of the potential function:

$$\Psi = \tfrac{1}{2}\kappa r^2 [3\cos^2\phi - 1] = \tfrac{1}{2}\kappa r^2 P(\phi), \tag{3.1}$$

where $\kappa = \dfrac{M}{R^3}$ measures the size of the perturbation and $P(\phi) = 3\cos^2\phi - 1$ (a Legendre polynomial in $\cos\phi$). Hill's problem is obtained by adding the perturbing potential Ψ to the Hamiltonian of the Kepler problem in the rotating coordinate system. The effect of the rotation of the coordinate system can also be viewed as a perturbation of the Kepler Hamiltonian. If the frequency of the rotation is ω, the perturbation is simply $\omega\Omega$ in polar coordinates.

It is possible to construct a perturbation of the Schwarzschild problem analogous to Hill's perturbation of the Kepler problem. Since gravity is modelled by the Lorentzian metric of spacetime, what is required is a perturbing metric, **h**, rather than a perturbing potential function. The construction of **h** is complicated by the fact that the Einstein equation is nonlinear. Thus one does not want **h** itself to satisfy the equation, but rather the total metric **g + h**, where **g** is the Schwarzschild metric (2.1). Now it is notoriously difficult to solve the Einstein equation. Very few explicit solutions are known. In particular, it is not known how to construct a metric which models the gravitational effects of two large masses. Since the problem at hand concerns a small perturbation of **g** and is intended only as an approximate description of the effects of a distant mass it is appropriate to replace the Einstein equation by its linearization about the solution **g**. More precisely, let **h** depend on a small parameter κ. Then **h** will be required to satisfy:

$$\frac{d}{d\kappa}\operatorname{Ric}(\mathbf{g}+\mathbf{h})|_{\kappa=0} = 0. \tag{3.2}$$

Since these equations are linear in **h**, it is possible to study the perturbing effects of the third body and those of the rotating coordinates separately. In fact, the Melnikov integral (to be discussed in Sect. 4) is also linear in the perturbation. Furthermore, the perturbation due to rotating coordinates has no intrinsic effect on the dynamics and will therefore contribute nothing to this integral. Thus it is possible to concentrate exclusively on the effects of the third mass.

Equation (3.2) was studied by Regge and Wheeler, who were attempting to analyze the stability of the Schwarzschild metric under small perturbations [Reg-Wh]. They show that by making use of the freedom to choose new coordinates, one can restrict attention to perturbations of the form:

$$\mathbf{h} = \kappa[fA(r)dt \otimes dt + fA^{-1}(r)dr \otimes dr + gr^2 d\phi \otimes d\phi + gr^2 \sin^2\phi\, d\theta \otimes d\theta], \tag{3.3}$$

where $f(r, \phi, \theta)$ and $g(r, \phi, \theta)$ are to be determined. Furthermore, they envoke separation of variables to expand f and g in spherical harmonics. Such a separation is fitting for the problem at hand because the Hill potential (3.1) is of this form. Assume that

$$f = F(r)P(\phi), \qquad g = G(r)P(\phi),$$

where $P(\phi) = 3\cos^2\phi - 1$. Then substitution of (3.3) into (3.2) leads, after some extremely laborious computations to a system of ordinary differential equation for the functions $F(r)$, and $G(r)$. Although the Ricci tensor in (3.2) has 10 independent components, many turn out to be zero simply because of the form of $\mathbf{h}$. In the end, (3.2) reduces to the two equations:

$$G'(r) = F'(r) + \frac{F(r)}{r^2 A(r)},$$

$$F'(r) + \frac{2F(r)}{r} + \frac{F(r)}{r^2 A(r)} + 4(F(r) - G(r)) = 0.$$

These are as in [Reg-Wh] except for the sign attached to $G(r)$.

It is easy to verify that the quadratic polynomials

$$F(r) = r(r-1), \qquad G(r) = r^2 - \tfrac{1}{2},$$

or any constant multiples of them satisfy the required equations. In particular, the metric

$$
\begin{aligned}
\mathbf{h} = {} & \kappa P(\phi)[F(r)A(r)dt \otimes dt + F(r)A^{-1}(r)dr \otimes dr \\
& + r^2 G(r)d\phi \otimes d\phi + r^2 G(r)\sin^2\phi d\theta \otimes d\theta] \\
= {} & \kappa P(\phi)[(r-1)^2 dt \otimes dt + r^2 dr \otimes dr \\
& + r^2(r^2 - \tfrac{1}{2})d\phi \otimes d\phi + r^2(r^2 - \tfrac{1}{2})\sin^2\phi d\theta \otimes d\theta]
\end{aligned}
\tag{3.4}
$$

is a solution of the linearized Einstein equations. There is a good reason to view $\mathbf{h}$ as the relativistic analogue of Hill's potential.

It is a basic tenet of general relativity that its predictions should agree with those of Newtonian gravitational theory in situations where gravity is weak. In this case, the metric of spacetime should be a small perturbation of the Minkowski metric. It can be shown that the geodesic equations for such a perturbed Minkowski space reduce in the appropriate limit to Newton's laws for motion in potential Φ provided that the coefficient of $dt \otimes dt$ in the perturbed metric is $-1 + 2\Phi$ [Sch]. As an example of this principle one can look at the Schwarzschild metric (2.1) in the region where r is large: the relevant coefficient of the perturbing metric is just $-1 + \dfrac{2M}{r}$. Now the metric $\mathbf{h}$ of (3.4) is proposed as a perturbation of the Schwarzschild metric, $\mathbf{g}$. The coefficient of $dt \otimes dt$ in $\mathbf{g} + \mathbf{h}$ is $-1 + \dfrac{2M}{r} + 2\tilde{\Psi}$, where $\tilde{\Psi} = \tfrac{1}{2}\kappa(r-1)^2 P(\phi)$. In the limit of large r $\left(\text{but still small } \dfrac{r}{R}\right)$ this reduces to Newtonian motion in a Kepler problem with perturbing potential $\tilde{\Psi}$. However, in this limit, the difference between $\tilde{\Psi}$ and Hill's perturbation Ψ in (3.1) is insignificant. Thus the metric $\mathbf{g} + \mathbf{h}$ can be viewed as the relativistic Hill problem.

4. Nonintegrability of the Relativistic Hill Problem

In this section, the Melnikov integral technique will be used to prove the existence of transverse homoclinic orbits in the geodesic flow of the perturbed Schwarzschild problem. Let $\mathbf{g}$ denote the Schwarzschild metric (2.1) and let $\mathbf{h}$ be the perturbing metric (3.4). The geodesic flow of the total metric $\mathbf{g}+\mathbf{h}$ will be analyzed by Hamiltonian methods such as were used in Sect. 2. Once again, only the planar case will be studied.

Write $\mathbf{g}+\mathbf{h}$ in the form:

$$\mathbf{g}+\mathbf{h}=-Bdt\otimes dt+Cdr\otimes dr+Dd\phi\otimes d\phi,$$

where $B=A(1-\kappa f)$, $C=A^{-1}(1+\kappa f)$, and $D=r^2(1+\kappa g)$. Here $f=r(r-1)P(\phi)$ and $g=(r^2-\frac{1}{2})P(\phi)$ with $P(\phi)=3\cos^2\phi-1$. As in Sect. 2, the geodesic flow has Lagrangian:

$$L=\tfrac{1}{2}\mathbf{g}(\gamma',\gamma')=\tfrac{1}{2}[-Bt'^2+Cr'^2+D\phi'^2].$$

The conjugate momenta are

$$E=Bt',\quad p=Cr',\quad \Omega=D\phi',$$

and the Hamiltonian is:

$$H=\frac{1}{2}\left[-\frac{E^2}{B}+\frac{p^2}{C}+\frac{\Omega^2}{D}\right].$$

H is expanded as a series in the parameter κ:

$$H=H_0+\kappa H_1+\dots,$$

where H_0 is the Hamiltonian (2.3) of the Schwarzschild problem and

$$H_1=-\frac{1}{2}\left[\frac{fE^2}{A}+fAp^2+\frac{g\Omega^2}{r^2}\right]. \tag{4.1}$$

The Hamiltonian H has only two constants of motion: E and the energy $H=-\frac{1}{2}$. The angular momentum Ω is no longer constant on account of the asymmetry of the perturbation. Consequently, it is no longer possible to reduce the godesic flow to a flow with one degree of freedom. Once E has been fixed, H is a Hamiltonian of two degrees of freedom with conjugate variables
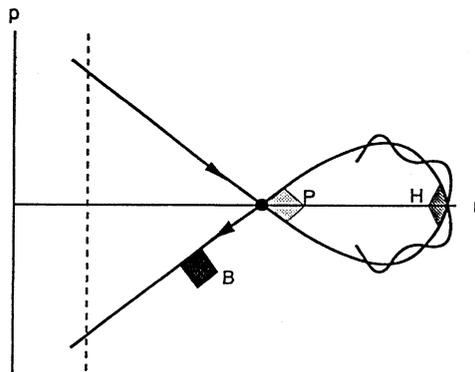

Fig. 5

(r, p, ϕ, Ω). Using the energy equation, Ω can be expressed as a function of the other three variables, although for any given E the requirement that $\Omega^2 > 0$ is satisfied only on an open subset of (r, p, ϕ) space.

The flow of the Schwarzschild problem can also be viewed in this way. In Sect. 2 the value of Ω was fixed and the corresponding phase portraits in the (r, p) plane were drawn (Fig. 3). The different curves represented different values of E. Instead one could fix E at the value of the homoclinic orbit in the figure and let Ω vary. Locally near the homoclinic orbit, the pictures would be much the same with the values of Ω distinguishing the different curves. As E^2 decreases towards the critical value of $\frac{8}{9}$, the homoclinic orbit shrinks and disappears. In the full three-dimensional flow, orbits leave the plane and flow around the space to return to the plane when ϕ increases by 2π. The motion of the (r, p) variables in the meantime is as shown in the phase portrait. The simplifying feature in this case is that the dynamics of (r, p) is independent of ϕ. For the perturbed problem, one can still consider the return mappings of these planes $\phi = \phi_0$. However, since the dynamics of (r, p) now depends on ϕ the mapping will not be obtained from a flow in the plane. If the perturbation parameter κ is small, the return mappings will still have a hyperbolic fixed point with one-dimensional stable and unstable manifolds, but in contrast to Fig. 3, one expects these curves to intersect transversely as in Fig. 5. If this is indeed the case then the points of intersection will be transverse homoclinic points and nearby orbits will display chaotic behavior.

To prove that this is indeed the case one can apply the version of the Melnikov integral method found in [Rob]. Consider a plane $\phi = \phi_0$ as in Fig. 5. At the point where the stable and unstable manifolds cross the r axis they will be separated by a certain distance. If this distance is zero for a certain value of ϕ_0 then the manifolds have a point of intersection along the r axis in that plane. One can measure the distance in many ways, but it is convenient to use the difference of the values of Ω at the two points of crossing [recall that Ω is being viewed as a function on (r, p, ϕ) space]. For the unperturbed system, the separation is zero because the stable and unstable manifolds coincide. But if κ is nonzero, the manifolds generally split apart. It is shown in [Rob] that the value of the separation is $\kappa M(\phi_0) + O(\kappa^2)$, where:

$$M(\phi_0) = \lim_{j \to \infty} \int_{T^*}^{T_j} D\Omega \cdot X_1 |_\gamma dt.$$

Here γ is the parametrization of the homoclinic orbit of the unperturbed system which crosses the r axis in the plane $\phi = \phi_0$ at $t = 0$, T_j and T_j^* are respectively the times when this orbit crosses the same plane again for the jth time in the future and the past, and where X_1 is the Hamiltonian vectorfield of the perturbing Hamiltonian H_1. The significance of $M(\phi_0)$ is as follows: if the function $M(\phi_0)$ has simple zeroes, then for all sufficiently small κ the stable and unstable manifolds have a transverse intersection.

The integrand of the Melnikov integral is just the rate of change of Ω along orbits of X_1; since X_1 is Hamiltonian:

$$D\Omega \cdot X_1 = -\frac{\partial H_1}{\partial \phi} = -\frac{P'(\phi)}{2} \left[\frac{F(r)E^2}{A(r)} + F(r)A(r)p^2 + \frac{G(r)\Omega^2}{r^2} \right].$$

Using the definitions of $P(\phi)$, $F(r)$, $G(r)$, and $A(r)$ and the energy equation for H_0, this can be reduced to:

$$D\Omega \cdot X_1 = \frac{3}{4}\sin(2\phi)\left[2(2E^2-1)r^2+2r+\frac{(2r-1)\Omega^2}{r^2}\right].$$

Because the parametrization of the homoclinic orbit is easier when ϕ is the independent variable, it is convenient to make a change of variables in the Melnikov integral. From the definition of Ω one finds $dt=\frac{r^2}{\Omega}d\phi$. Substituting this into the integral and ignoring some constant factors which are irrelevant to the question of simple zeroes, one finds:

$$M(\phi_0)= \lim_{j\to\infty} \int_{-2\pi j}^{2\pi j} [2(2E^2-1)r(\phi)^4+2r(\phi)^3+\Omega^2(2r(\phi)-1)]\sin(2\phi)d\phi . \quad (4.2)$$

Here $r(\phi)$ is the parametrization of the homoclinic orbit of the unperturbed problem which crosses the r axis in the plane $\phi=\phi_0$. An explicit formula for $r(\phi)$ can be obtained from Eq. (2.4). This equation can be written in the form:

$$u'(\phi)^2 = \frac{E^2}{\Omega^2} - P(u),$$

where $P(u)$ is the cubic polynomial $(1-u)\left(u^2+\frac{1}{\Omega^2}\right)$. For the value of Ω corresponding to the homoclinic orbit, $\frac{E^2}{\Omega^2}$ is the critical value $P(u_1)$, where u_1 is the critical point representing the hyperbolic circular periodic orbit being studied (see Fig. 2). Thus we can factorize the right-hand side of the last equation to get:

$$u'(\phi)^2 = (u-u_1)^2(u-u_2) .$$

This can be solved explicitly with a bit of effort to obtain:

$$u(\phi)=u_1-\alpha^2 \operatorname{sech}^2(\alpha(\phi-\phi_0)/2), \quad (4.3)$$

where $\alpha^2=u_1-u_2$. As usual, $r(\phi)=\frac{1}{u(\phi)}$. Note that as the parameter E^2 decreases towards the critical level of $\frac{8}{9}$, α tends to zero (the width of the homoclinic orbit shrinks). In fact, it is not hard to show that near the bifurcation, α behaves like $(E^2-\frac{8}{9})^{1/4}$.

As it stands, the improper integral in (4.2) is not absolutely convergent. Note, however, that one could add any constant to the expression in square brackets without changing the value of the integral. Choosing this constant so that the square bracket vanishes at $r=\frac{1}{u_1}$ gives an integral which is absolutely convergent. This follows from the fact that $r(\phi)$ converges exponentially to $\frac{1}{u_1}$ as $|\phi|\to\infty$. Let $Q(r)$ denote the polynomial in square brackets in (4.2) with the added constant. Then

$$M(\phi_0)= \int_{-\infty}^{\infty} Q(r(\phi))\sin(2\phi)d\phi . \quad (4.4)$$

Although (4.3) and (4.4) bring the question of nonintegrability into an explicitly computable form the computation of the relevant integral is still rather difficult owing to the complexity of the function $r(\phi)$. It turns out that this function becomes much simpler in the limit $\alpha \to 0$, that is as $E^2 \to \frac{8}{9}$. So now some effort will be devoted to putting $M(\phi_0)$ into a form where it is possible to take this limit.

First, $M(\phi_0)$ can be written::

$$M(\phi_0) = \frac{1}{2i}(M_+(\phi_0) - M_-(\phi_0)),$$

where

$$M_\pm(\phi_0) = \int_{-\infty}^{\infty} Q(r(\phi))e^{\pm 2i\phi}d\phi.$$

To see the dependence on ϕ_0 more clearly, define a new integration variable $s = \phi - \phi_0$. Then

$$M_\pm(\phi_0) = e^{\pm 2i\phi_0} \int_{-\infty}^{\infty} Q(r(s))e^{\pm 2is}ds,$$

where

$$\frac{1}{r(s)} = u(s) = u_1 - \alpha^2 \operatorname{sech}^2(\alpha s/2).$$

Now the integrals are independent of ϕ_0. Furthermore, the two integrals are equal because $r(s)$ is an even function. Thus

$$M(\phi_0) = \sin(2\phi_0)I,$$

where

$$I = \int_{-\infty}^{\infty} Q(r(s))e^{2is}ds.$$

The next step is to reduce I to a contour integral over a compact contour. From now on, all variables will be viewed as complex. Introduce another new integration variable $z = s - \dfrac{i\pi}{\alpha}$. Then

$$I = e^{-\frac{2\pi}{\alpha}} \int_L Q(r(z))e^{2iz}dz,$$

where L is the line $\left\{z : \operatorname{im}(z) = -\dfrac{i\pi}{\alpha}\right\}$ oriented from left to right and

$$\frac{1}{r(z)} = u(z) = u_1 + \alpha^2 \operatorname{csch}^2(\alpha z/2). \tag{4.6}$$

Now the integral over the line $\tilde{L} = \left\{z : \operatorname{im}(z) = +\dfrac{i\pi}{\alpha}\right\}$ differs from I only by a factor of $e^{-\frac{4\pi}{\pi}}$ so I can be written as

$$I = C(\alpha)J \tag{4.7}$$

with

$$C(\alpha) = \frac{e^{-\frac{2\pi}{\alpha}}}{\left(1 - e^{-\frac{4\pi}{\alpha}}\right)}, \qquad J = \oint_{\Gamma} Q(r(z)) e^{2iz} dz. \tag{4.8}$$

Here Γ is initially the contour $L - \tilde{L}$ but by Cauchy's theorem, it could be replaced by any convenient closed contour lying in $\left\{z: |\text{im}(z)| < \frac{2\pi}{\alpha}\right\}$ which encircles all of the poles of the integrand. It will be shown that it is possible to choose a contour satisfying these conditions which is independent of α. This will facilitate the computation of the limit of the Melnikov integral as $\alpha \to 0$. Note also that the constant which was added to $Q(r)$ earlier does not contribute to J, so in the discussion to follow it will be dropped.

Since $Q(r)$ is a polynomial, the poles of the integrand of J are just the poles of $r(z)$, that is, the zeroes of $u(z)$. In (4.6), the constants u_1 and α depend continuously on the parameter E and as $E^2 \to \frac{8}{9}$, $u_1 \to \frac{1}{3}$, and $\alpha \to 0$. Let $\lambda = E^2 - \frac{8}{9}$. Then

$$\lim_{\lambda \to 0} u(z) = u_0(z) = \frac{1}{3} + \frac{4}{z^2}$$

uniformly on compact subsets of the complex plane with the origin deleted. Therefore, the zeroes of $u(z)$ converge as $\lambda \to 0$ to those of $u_0(z)$, namely to $z = \pm i\sqrt{12}$. So if Γ is any simple closed curve encircling 0, $\pm i\sqrt{12}$, it will satisfy the required conditions for all sufficiently small λ.

Fixing such a contour gives:

$$J = J_0 + o(1) \tag{4.9}$$

as $\lambda \to 0$, where

$$J_0 = \oint_{\Gamma} Q(r_0(z)) e^{2iz} dz$$

and

$$r_0(z) = \frac{1}{u_0(z)} = \frac{3z^2}{z^2 + 12}. \tag{4.10}$$

From (4.5), (4.7), (4.8), and (4.9) one finds:

$$M(\phi_0) = \sin(2\phi_0) C(\alpha) (J_0 + o(1)).$$

Note that the constant $C(\alpha)$ is nonzero if $\alpha \neq 0$. It follows that if $J_0 \neq 0$, then for all sufficiently small $\lambda > 0$ the Melnikov function has the desired property of having nondegenerate zeroes and so the corresponding geodesic flows have transverse homoclinic points. Therefore, the proof of nonintegrability for these parameter values depends only on the value of J_0. Note that the simplicity of $r_0(z)$ makes the computation of J_0 much more tractable than that of J itself. Furthermore, this analysis brings out an interesting feature of the problem: the

constant $C(\alpha)$ is exponentially small. Since α behaves like the fourth root of λ near the bifurcation,

$$C(\alpha(\lambda)) \sim e^{-K/\lambda^{1/4}}$$

for some constant K. This means that size of the splitting of the stable and unstable manifolds is also exponentially small in λ.

To compute J_0, note that the residue of $Q(r_0(z))e^{+2iz}$ at $-i\sqrt{12}$ is just minus the residue of $Q(r_0(z))e^{-2iz}$ at $+i\sqrt{12}$. It follows that

$$\begin{aligned}
J_0 &= 2\pi i[\mathrm{Res}_{+i\sqrt{12}}(Q(r_0(z))e^{2iz}) + \mathrm{Res}_{-\sqrt{12}}(Q(r_0(z))e^{2iz})] \\
&= -4\pi\,\mathrm{Res}_{+i\sqrt{12}}(Q(r_0(z))\sin(2z))\,.
\end{aligned}$$

Substituting (4.10) into $Q(r)$ gives (ignoring the constant term):

$$Q(r_0(z)) = 18z^2\frac{11z^6 + 72z^4 + 432z^2 + 1728}{(z^2+12)^4}\,.$$

The residue of $Q(r_0(z))\sin(2z)$ at $z = i\sqrt{12}$ is therefore:

$$-18\left(\frac{1107}{4}\cosh(2\sqrt{12}) + \frac{2655}{16}\sqrt{3}\sinh(2\sqrt{12})\right) \neq 0\,.$$

So, in fact, the transversal homoclinic points do exist.

5. The Qualitative Interpretation

This section is devoted to a qualitative discussion of the behavior and physical meaning of the geodesics of the perturbed Schwarzschild problem. Figure 5 shows one of the planes $\phi = \phi_0$ containing a hyperbolic fixed with transversely intersecting stable and unstable manifolds. This is the situation which has been shown to hold for certain values of the parameters λ and κ. More precisely, for each sufficiently small positive λ, there is a constant $\kappa_0(\lambda)$ such that for $|\kappa| < \kappa_0(\lambda)$ the transverse homoclinic orbits exist.

The hyperbolic fixed point is the intersection of the plane $\phi = \phi_0$ with a hyperbolic periodic orbit of the geodesic flow of the perturbed Schwarzschild problem. This orbit is close to the corresponding one for the unperturbed Schwarzschild problem; namely, it is a nearly circular orbit revolving around the
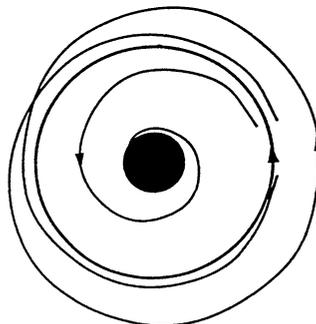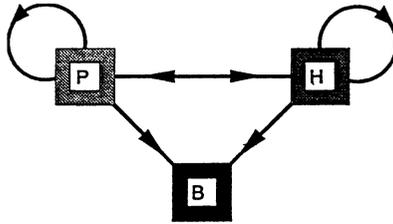

Fig. 6

Fig. 7

black hole approximately at radius $r = 3$, that is, at about three Schwarzschild radii (see Fig. 6). Similarly, the transverse homoclinic orbit behaves like the homoclinic orbit of the unperturbed problem: it gradually spirals out to a somewhat larger radius and then spirals back in towards the periodic orbit.

The new feature in the perturbed problem is the existence of a chaotic invariant set of orbits near the transverse homoclinic orbit. These can be described in the usual way using symbolic dynamics. Figure 5 shows three rectangular "windows" in the plane through which orbits might pass. The first window P is close to the periodic orbit, so any orbit passing through it will behave for a while like the periodic orbit itself. Similarly, the second window H contains points which will behave like the homoclinic orbit. Finally, the last window B consists of points which are headed for oblivion in the black hole. Let T denote the first return map to the plane $\phi = \phi_0$. This will be defined in some open set around the homoclinic orbit. Because of the hyperbolicity of the fixed point, the window P will get stretched out along the unstable manifold as T is iterated. Thus a suitable iterate T^N will stretch P across itself and across both of the windows, H and B. The effect of iteration on H is similar. First, it is mapped along the homoclinic orbit into a neighborhood of the fixed point, where it is stretched until it maps across each of the other windows. A similar analysis applies to the iterates of T^{-1}. Thus under T^N the windows are mapped across one another as indicated schematically in Fig. 7.

It follows from the usual symbolic dynamical arguments that any bi-infinite directed path in the graph of Fig. 7 can be realized by a geodesic of the perturbed Schwarzschild problem in the sense that the geodesic passes through the windows in the order indicated by the path in the graph. For example, the path which always passes through P is realized by the periodic orbit itself. The path which passes through P every time except for one pass through H is realized by the transverse homoclinic orbit. Other paths give rise to motions which were not present in the unperturbed Schwarzschild problem. For example, there are orbits which pass through P and H in many different orders. If the given order of passages is periodic, then there is a hyperbolic periodic geodesic realizing it. These orbits will revolve near $r = 3$ except for erratic episodes of spiralling out to larger radii and back in again. Finally, there are orbits which behave in this way for an arbitrarily long time and then disappear abruptly into the black hole. Which course will be followed by a given geodesic depends sensitively on the initial conditions. In particular, the domain of attraction of the black hole will intersect the windows P and H in an open set containing infinitely many strips running roughly parallel to the stable manifold of the fixed point and it is a delicate matter to decide whether a given geodesic will fall into the black hole or not.

References

[Hill]   Hill, G.W.: Researches in the lunar theory. Am. J. Math. **1**, 5–26, 129–147, 245–260 (1878)

[MTW]   Misner, C., Thorn, K., Wheeler, J.: Gravitation. New York: W. H. Freeman 1973

[O'Neill] O'Neill, B.: Semi-Riemannian geometry. New York: Academic Press 1983

[Reg-Wh] Regge, T., Wheeler, J.: Stability of a Schwarzschild singularity. Phys. Rev. **108** (4), 1063–1069 (1975)

[Rob]   Robinson, C.: Horseshoes for autonomous Hamiltonian systems using the Melnikov integral. Ergodic Theory Dyn. Syst. **8**, 395–409 (1988)

[Sch]   Schutz, B.: A first course in general relativity. Cambridge: Camb. Univ. Press 1984