PROCEEDINGS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The Gravity Field of a Particle

Charles Darwin

Proc. R. Soc. Lond. A 1959 **249**, 180-194
doi: 10.1098/rspa.1959.0015

The gravity field of a particle

By Sir Charles Darwin, F.R.S.

(*Received* 13 *August* 1958)

Einstein's equations for the orbits round an attracting point mass, here called the sun, are examined so as to see whether there are orbits which end in the sun, as there are in the corresponding case of electrical attraction when relativity is allowed for.

With the measure of the radius as usually taken, it is shown that no hyperbolic orbit can have perihelion inside $r = 3m$, and an elliptic orbit cannot have perihelion inside $r = 4m$. Particles going inside these distances will be captured.

Circular orbits are possible for any greater radius. If $3m < r < 4m$ the orbit is unstable; with one disturbance it falls into the sun, with the opposite it escapes in a spiral to infinity. If $4m < r < 6m$, it is also unstable, either falling into the sun, or moving out to some aphelion at a greater radius before returning to its circle. Only if $r > 6m$ is the orbit stable.

A study is made of the travel of light rays. No light ray from infinity can escape capture unless its initial asymptotic distance is greater than $3\sqrt{3}\,m$.

A field of stars surrounds the sun, and is viewed in a telescope pointed at the sun from a distance. If the field as seen is mapped as though in a plane through the sun, each star, in addition to its direct image, will show a series of faint 'ghosts' on both sides of the sun. The ghosts all lie just outside the distance $3\sqrt{3}\,m$.

A few technical details are given about the orbits of the captured particles.

1. Introduction

The stimulus for the present note came from the recollection of a paper I published as long ago as 1913 entitled 'On some orbits of an electron'. In this it was shown that, when relativity effects are included, certain of the orbits of an electron going past a nucleus become spirals and fall into it. It seemed it might be interesting to see whether a similar effect would be given by Einstein's gravitational field.

The present work could easily have been done nearly forty years ago, but it does not appear that it ever was, or anyhow that it has ever become widely known.* The method was indeed to some extent worked out by Forsyth in 1920, but he applied it only to orbits of small eccentricity, and approximate methods suffice for these. No doubt the chief reason why the present work was not done is that the results are quite unpractical. Thus the sun with mass 2×10^{33} g, is represented in Einstein's theory by a length $m = 1{\cdot}5$ km. In that theory the sun is treated as a point-particle, but, forgetting the difficulties of non-Euclidean geometry, if this were replaced by a sphere of matter as large even as $10m$, that is 15 km in radius, this sphere would have a density of about $10^{14}\,\text{g}\,\text{cm}^{-3}$. Thus the present questions of spiral orbits are not practical in the sense that, with a sun composed of any ordinary matter however dense, capture will already have occurred by collision long before the phenomena here worked out could come in evidence.

The problem is made more realistic by casting it in a form that also goes back to those early days. Rutherford discovered the nucleus by shooting α-particles past it, and observing the emergent orbits. Even from the first he hoped to find out something of the size and shape of the nucleus by observing departures from the inverse

* I owe this information to Professor W. H. McCrea, F.R.S., and I wish to thank him for it and for various suggestions arising in the course of the work.

square law, though in fact he did not get any such effects. Let us then suppose that an experimenter wishes to test the gravitational field round a heavy particle—call it the sun—by shooting past it small particles—call them comets—and observing their emergent orbits. If there should be any spiral orbits the comets moving in them would be lost to him, so that part of the field of gravity would not be observable. The problem here then is to apply the method to Einstein's gravitational field for an attracting point mass. As is well known, this field has a 'barrier' at a place conventionally indicated by $r = 2m$, so that no observation could possibly be made of anything inside that; but, as will appear, the experimenter will in fact find that he can get no information about the field inside the radius $r = 3m$. To test the field nearer the sun than that, he would have to go there himself, and then fire a rocket so as to escape.

In the electric case no attempt was made to include the effects of the acceleration-radiation, which would certainly have increased the probability of capture. Here, too, any similar effect will be disregarded for the cogent reason that so little is known about gravity waves.

2. Electric orbits

The electric case can be seen quite briefly, and so is repeated here. The electron moves past a heavy nucleus in an orbit determined by the equations of energy and angular momentum:

$$\frac{mc^2}{\sqrt{(1-v^2/c^2)}} - \frac{e^2}{r} = W, \tag{1}$$

$$\frac{mr^2\dot{\theta}}{\sqrt{(1-v^2/c^2)}} = p. \tag{2}$$

If the time is eliminated, and if u is written as usual for $1/r$, the result is

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2 = \frac{W^2-m^2c^4}{c^2p^2} + 2\frac{e^2W}{c^2p^2}u - u^2\left(1-\frac{e^4}{c^2p^2}\right). \tag{3}$$

If then

$$p < e^2/c \tag{4}$$

the coefficient of u^2 is positive, and since $W > mc^2$ for a hyperbolic orbit, there are only positive terms on the right. It follows that there can be no apse, and so the orbit must move in a spiral ending in the nucleus. It is quite easy to work out the details, but it is unnecessary to do so to see the point of the argument.

Suppose that the electron at infinity has velocity V, and that it is moving on a line at perpendicular distance l from the nucleus, then

$$W = mc^2/\sqrt{(1-V^2/c^2)} \quad \text{and} \quad p = mlV/\sqrt{(1-V^2/c^2)}$$

and the orbit is a spiral if

$$l < \sqrt{(1-V^2/c^2)}\, e^2/mcV. \tag{5}$$

Any electron fired along a line for which l is less than this will not emerge. Thus if the experimenter has only rather slow electrons available, there will be quite a large area that will be useless to him.

All this is pure classical theory, and it was indeed put forward before Bohr had proposed his theory. At that time there seemed little of interest in elliptic orbits, which would also be annihilated if (4) held, but later of course these orbits became of great importance in Sommerfeld's theory of the hydrogen spectrum. The value given by (4) is $\hbar/137$, and of course the quantum theory prohibits such orbits, though the subject does come up in a different form when $p = 0$ in connexion with Dirac's equations for the electron; in that case there is a weak singularity in the wave equations at the origin. It seems unlikely that any corresponding effect can be of importance in gravitation theory, because the dimensionless quantity corresponding to $e^2/c\hbar$ is most naturally taken as $GM^2/c\hbar$, where M is the mass of a neutron, but this instead of being 1/137 is of magnitude 6×10^{-39}, which can hardly have significance. There is therefore little expectation that the quantum will make to gravitation theory the same kind of contribution that it did to atomic theory.

3. The gravity equations

In the case of gravitation, comets are shot past the sun, and the question is whether some of them will be lost by going in spirals into the sun, and so giving no information to the experimenter.

As is usual, take units in which c and G are unity. In these units let m be the sun's mass; the mass of the comet is taken as negligibly small. The Einstein equations then give

$$\left.\begin{aligned} -\frac{r}{r-2m}\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 - r^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}s}\right)^2 + \frac{r-2m}{r}\left(\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 &= 1, \\ r^2\frac{\mathrm{d}\theta}{\mathrm{d}s} &= p, \\ \frac{r-2m}{r}\frac{\mathrm{d}t}{\mathrm{d}s} &= C. \end{aligned}\right\} \tag{6}$$

The other angular co-ordinate ϕ may be taken as zero.

Eliminate s and t, and write u for $1/r$. Also, for short, write

$$C^2 - 1 = A, \tag{7}$$

so that A, loosely speaking, measures the energy. There results the well-known equation (see, for instance, Eddington 1923)

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2 = 2mu^3 - u^2 + \frac{2m}{p^2}u + \frac{A}{p^2}. \tag{8}$$

Near infinity the space becomes Galilean with straightforward geometry, but of course that is not so for small values of r. The formulae can be made there to look quite different, if r is replaced by some other form, such as the 'isotropic form', but this part of space is not occupied by the experimenter, and he is entitled to use the present co-ordinate for the radial measure as being the most convenient form to represent a part of space not directly accessible to him.

At infinity his geometry is normal, and he may say unambiguously that the comet has velocity V, and that it starts on a line at perpendicular distance l from the sun. Then

$$C = 1/\sqrt{(1-V^2)}, \quad p = lV/\sqrt{(1-V^2)}, \tag{9}$$

so that

$$A = V^2/(1-V^2) \quad \text{and} \quad p^2/A = l^2. \tag{10}$$

The main interest is in the hyperbolic orbits, but the elliptic orbits will also be considered. For these A is negative, but on account of (7) it must be greater than -1.

PROCEEDINGS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

4. Escape orbits

The orbits are determined by (8), but the algebra of the cubic on the right is a great deal more troublesome than that of (3), and it is easy to get into very complicated expressions if it is not handled in the best way.

By the 'rule of signs', if A is positive there may be two positive roots of the right-hand side of (8) or else none, while if it is negative there may be three or one. A root signifies an apse, where $\mathrm{d}u/\mathrm{d}\theta$ changes sign, so for a comet that escapes capture, there must be a positive root, and it is therefore the cases where there are two positive roots that are interesting. For the elliptic type of motion with A negative there will always be an apse at aphelion, and the question is whether there is also another for perihelion.

If there are no positive roots it follows that u must increase all the time and so the comet will end in the sun. This is not the interesting case, but for the sake of completeness it will be outlined in § 10. It will be a main point to consider the separation between having three real roots and one, according to the values of the parameters p and A, but the general expressions for the case of three real roots may be looked at first.

In this case the equation can be written

$$(\mathrm{d}u/\mathrm{d}\theta)^2 = 2m(u-a)(u-b)(u-c). \tag{11}$$

Here say $a > b > c$ and a and b are positive, c either positive or negative. Certain of their further properties will be examined later, but first the solution may be given assuming them general. There are only two physically significant fields available to u, first when it is greater than a, which represents a comet that will fall into the sun, and second $b > u > c$, which is the interesting case. For it the solution is given by

$$u = c + (b-c)\,\mathrm{sn}^2\,\zeta \quad (\mathrm{mod}\,k), \tag{12}$$

where

$$k^2 = (b-c)/(a-c). \tag{13}$$

Making the substitution, it is found that

$$(\mathrm{d}\zeta/\mathrm{d}\theta)^2 = \tfrac{1}{2}m(a-c). \tag{14}$$

It may be noted that the variable ζ is analogous not to θ, but to $\frac{1}{2}\theta$. There are some advantages in writing $\zeta = K - \frac{1}{2}\eta$ and using η. This would replace $\mathrm{sn}^2\,\zeta$ by $(1+\mathrm{cn}\,\eta)/(1+\mathrm{dn}\,\eta)$. However, the substitution of (12) in (11) is so much more obvious that I have retained that form throughout the work.

For a hyperbolic orbit c is negative, u starts at zero, so ζ starts at ζ_1 where $\mathrm{sn}^2\,\zeta_1 = (-c)/(b-c)$; it increases to K and then goes on to $2K-\zeta_1$, where u again will vanish. For an elliptic orbit u starts at c at aphelion with $\zeta = 0$, increases to b at $\zeta = K$, and then goes back to c at $\zeta = 2K$.

An example may be given. Take

$$\begin{aligned}\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2 &= 2m\left(u-\frac{1}{3m}\right)\left(u-\frac{1}{4m}\right)\left(u+\frac{1}{12m}\right)\\ &= 2mu^3 - u^2 + 2u\frac{5}{144m} + \frac{1}{72m^2}.\end{aligned} \tag{15}$$

Thus $l = m\sqrt{72} = 8{\cdot}49m$; perihelion is at $4m$, and since

$$\frac{m^2}{p^2} = \tfrac{5}{144}, \quad A = \tfrac{2}{5}, \quad \text{so that} \quad V = \sqrt{\tfrac{2}{7}} = 0{\cdot}534.$$

The solution is given by

$$u = -\frac{1}{12m} + \frac{1}{3m}\operatorname{sn}^2\zeta, \quad \text{where} \quad k^2 = \tfrac{4}{5} \quad \text{and} \quad \zeta = \theta\sqrt{\tfrac{5}{24}}. \tag{16}$$

The orbit ranges from ζ_1 to $2K - \zeta_1$, where $\operatorname{sn}^2\zeta_1 = \frac{1}{4}$. Tables of 'elliptic integrals of the first kind' yield ζ from $\operatorname{sn}\zeta$ for the appropriate k, and it would be easy from them to trace the complete orbit, but here it will suffice to give the final value. For this $K - \zeta_1 = 1{\cdot}716$, which yields 215° for the final value of θ as measured from perihelion. Then for this quasi-hyperbolic orbit the entering and emerging paths cross one another and the whole orbit consists of a complete loop with asymptotes going off towards infinity at angle 70° to one another.

5. Limits of escape orbits

The most powerful and simplest attack on the subject is to deal with the cases that separate the cubics having three real roots from those with one real root, and these are those where there are two equal roots. Very complicated algebra comes if (8) is attacked directly by putting down the condition for equal roots, but the matter is made easy by assuming that there is an equal root at $r = R$. This one parameter then determines the whole equation as being

$$\begin{aligned}\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2 &= 2m\left(u - \frac{1}{R}\right)^2\left(u - \frac{R-4m}{2Rm}\right)\\ &= 2mu^3 - u^2 + 2\frac{R-3m}{R^2}u - \frac{R-4m}{R^3}.\end{aligned} \tag{17}$$

Comparing this with (8), if p_0 is the special angular momentum for this orbit,

$$p_0^2 = R^2m/(R-3m) \tag{18}$$

and

$$A = m(4m-R)/R(R-3m). \tag{19}$$

Evidently, to give a real value for p_0, R must be greater than $3m$. This is one of the main results.

Consider now an orbit having the same value of A, but a larger value of p than p_0. Evaluate R from A by (19) and substitute $u = 1/R$ in the right-hand side of (8). The result is

$$2m\frac{1}{R^3} - \frac{1}{R^2} + \frac{2m}{p^2}\frac{1}{R} + \frac{A}{p^2} = \left(\frac{1}{p^2} - \frac{1}{p_0^2}\right)\left[\frac{2m}{R} + A\right] = \left(\frac{1}{p^2} - \frac{1}{p_0^2}\right)\frac{m(R-2m)}{R(R-3m)},$$

which is certainly negative whether A is positive or negative. Therefore, $\mathrm{d}u/\mathrm{d}\theta$ will have vanished already at some smaller value of u than $1/R$, and so the orbit with this p will have had its perihelion at some larger value of r than R. In particular, there can be no perihelion nearer the sun than $3m$. A short algebraic note on this point is given in §12.

6. Three classes of solution

There are three ranges to be considered.

(i) If $3m < R < 4m$, this signifies that the two positive roots a and b in (11) are equal and c is negative. A is positive and the orbit comes from infinity. For the solution in (12) now, $k^2 = 1$, and the solution becomes

$$u = -\frac{4m-R}{2Rm} + \frac{6m-R}{2Rm}\tanh^2\zeta, \tag{20}$$

where

$$\zeta = \theta\sqrt{\frac{6m-R}{4R}}.$$

This means that the orbit ends by approaching asymptotically towards a circle of radius R. At its start

$$l^2 = p_0^2/A = R^3/(4m-R), \tag{21}$$

which determines the length of the perpendicular from the sun on the asymptotic path at infinity. At infinity the velocity is given by

$$V^2 = \frac{A}{A+1} = \frac{m(4m-R)}{(R-2m)^2}.$$

If V is nearly 1, the ending circle is only a little greater than $3m$. The case $V = 1$, corresponding to the propagation of light, will be considered later. If V is a small velocity, l becomes quite large. A numerical example may be given. The actual sun's mass is equivalent to $m = 1{\cdot}5$ km. As an arbitrary choice take the velocity of the earth in its orbit as V for the comets. Then $V = 10^{-4}$. This gives the terminating circle only just less than $4m$, and it makes $l = 4m/V = 6\times10^4$ km, that is to say about nine times the earth's radius. Any comet projected from inside a circle of this radius will be captured.

Table 1 gives a few examples of other limiting orbits.

Table 1

R/m	3·0	3·2	3·4	3·6	3·8	4·0
l/m	5·19	6·40	8·09	10·80	16·57	∞
V	1	0·745	0·553	0·395	0·248	0

(ii) $4m < R < 6m$. Here again it is the two roots a and b that are equal, but now c is positive. The orbit starts at aphelion at $2Rm/(R-4m)$, and again goes in a spiral, ending asymptotically at the circle R. The solution (20) again holds. A is now negative, vanishing at $R = 4m$ and reaching the value $-\frac{1}{9}$ at $R = 6m$. At $R = 6m$ all three roots a, b, c are equal.

Table 2 gives a few values.

Table 2

R/m	4	4·5	5	5·5	6
aph./m	∞	18	10	7·33	6
A	0	$-0{\cdot}074$	$-0{\cdot}100$	$-0{\cdot}109$	$-0{\cdot}111$

(iii) $R > 6m$. Here the inequalities are $a > b = c$. One solution has aphelion at $u = a$ and falls into the sun. There is also the singular solution $u = b = 1/R$. The quantity A, which measures the energy, is at a minimum at $R = 6m$, being there $-\frac{1}{9}$, and it increases towards zero for larger radii. No orbit with $A < -\frac{1}{9}$ can escape capture.

7. Circular orbits

It is interesting to consider the circular orbits in a little more detail, because they show the remarkable differences between Newton's and Einstein's laws.

The angular velocity in any circular orbit can be evaluated by the use of (6), (7), (18) and (19). Thus $C = (R-2m)/\sqrt{R(R-3m)}$ and so

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{p}{r^2}\frac{r-2m}{rC} = m^{\frac{1}{2}}R^{-\frac{3}{2}}.$$

This is exactly Kepler's third law, and it holds for any circle for which $R > 3m$. This seems a remarkable result, but it must be remembered that the significance of R near the sun is to some extent conventional, so that no deep meaning need be assigned to it.

Suppose now that the planet's motion is observed from a distant observatory, chosen somewhere on the line of the pole of the ecliptic, since with this choice all complications of perspective and of varying radiation times of mutual signals are avoided. An astronomer in this observatory and one on the planet itself may differ in their views about r and t, but they will agree about the meaning of the angle θ. Thus the distant astronomer will say that the planet has moved through angle $\Delta\theta$ in time $\Delta t = \Delta\theta R^{\frac{3}{2}}m^{-\frac{1}{2}}$. On the other hand, the planet's clocks will record the *proper time* there, as measured by ds, and

$$\frac{\mathrm{d}\theta}{\mathrm{d}s} = \frac{p}{r^2} = \frac{m^{\frac{1}{2}}}{R\sqrt{(R-3m)}}$$

so that $\Delta s = \Delta\theta R\sqrt{(R-3m)}\,m^{-\frac{1}{2}}$. Thus if flashes are sent out from the planet at a rate of one a minute according to its own clocks, they will arrive at the distant observatory at intervals of $(1-3m/R)^{-\frac{1}{2}}$ minutes according to the local clocks there. The distant observer will say that the planet's clocks are running slow by a factor $\sqrt{(1-3m/R)}$. Conversely, the planet's observer will say the distant clocks are running fast as judged by flashing signals. This result has been given by McCrea (1956). There is no analogy to the case of two bodies in relative motion in outer space, where each judges the other's clock to be running slow.

It is to be noted that this is not the same result as is given by the reddening of spectral lines in a gravitational field, since that is $\sqrt{(1-2m/R)}$. In that case the emitting gas is supposed to be at rest, being held up by unspecified forces, and those forces are not included in relativity theory and so call for careful consideration. Here, put crudely, there is superposed a further slowing on account of the orbital motion, but it would be hard to present the result convincingly as a superposition of the two effects. However, that is quite unnecessary because the result $\sqrt{(1-3m/R)}$ follows out of pure relativity theory without recourse to the idea of extraneous forces.

For Newton's gravity, all circular orbits are stable, and they have negative energies tending towards $-\infty$ as R tends to zero. For Einstein's gravity it has been seen that A has a minimum value of $-\frac{1}{9}$ when $R = 6m$, and circles inside that are unstable. When $3m < R < 4m$, a small displacement in one direction will cause the planet to fall into the sun, while one in the opposite direction will cause it to escape to infinity. For the mass of our own sun the shortest year a planet could have, as seen from a distance, would be about $1{\cdot}6 \times 10^{-4}$ s.

When $4m < R < 6m$, one displacement will again cause it to fall into the sun, while the opposite displacement will start it into a spiral with increasing radius, and it will move out a finite distance towards an aphelion, as given in table 2. The outer part of this orbit will be approximately an ellipse, and then from aphelion it will return towards the circle of radius R.

When $R > 6m$ the circle is stable; this is the case illustrated by the perihelion of Mercury, and the accurate solution of the perihelion's motion is easily given. Take $b = (1/R)$, $c = 1/(R+\epsilon)$, so that $a = (1/2m) - (1/R) - 1/(R+\epsilon)$. By (13) k^2 tends to zero with ϵ. Thus $2K = \pi$, and by (12) approximately

$$u = \frac{1}{R} + \frac{\epsilon}{R^2}\sin^2\zeta. \tag{22}$$

Also by (14) in the limit θ becomes $2\zeta/\surd(1-6m/R)$. Thus successive perihelia occur at angular distances

$$2\pi/\surd(1-6m/R) \tag{23}$$

from one another. When R is much greater than m this gives the well-known result $2\pi(1+3m/R)$.

8. Light rays

The special case of the propagation of light-rays may be considered not only for its own interest, but also because certain peculiar results will come out of it in the next section. Light-rays may be regarded simply as particles with velocity $V = 1$. Alternatively in (6) $\mathrm{d}s$ may be equated to zero, and the infinite constants p and C may be replaced by their ratio which gives a finite constant. Then if θ is taken as independent variable

$$-\frac{r}{r-2m}\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2 - r^2 + \frac{r-2m}{r}\left(\frac{\mathrm{d}t}{\mathrm{d}\theta}\right)^2 = 0, \tag{24}$$
$$\frac{r-2m}{r^3}\left(\frac{\mathrm{d}t}{\mathrm{d}\theta}\right) = \frac{C}{p} = \frac{1}{l},$$

From these the elimination of t gives

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2 = 2mu^3 - u^2 + \frac{1}{l^2}. \tag{25}$$

The cubic here has one negative root, and may have either two positive roots or none. The case of equal roots has already been considered in § 6. It gives $R = 3m$ for the terminating spiral, and the starting point for this ray has $l_0 = 3\surd 3\ m$. The equation of this orbit is

$$u = -\frac{1}{6m} + \frac{1}{2m}\tanh^2 \tfrac{1}{2}\theta. \tag{26}$$

For orbits outside this the general solution can be expressed in terms of elliptic integrals, and since these are numerically neither familiar nor obvious, the results are given here, even though little use will be made of them. The solution is expressed in terms of the perihelion distance P, and to make the equation conform to (25) it is not hard to see that the three roots must be

$$\frac{Q+P-2m}{4mP}, \quad \frac{1}{P}, \quad -\frac{Q-P+2m}{4mP},$$

where

$$Q^2 = (P-2m)(P+6m).$$

The equation becomes

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\theta}\right)^2 = 2mu^3 - u^2 + \frac{P-2m}{P^3},$$

so that for the initial asymptotic line

$$l^2 = P^3/(P-2m). \tag{27}$$

The solution is given by

$$\left.\begin{aligned} u &= -\frac{Q-P+2m}{4mP} + \frac{Q-P+6m}{4mP}\,\mathrm{sn}^2\,\zeta, \\ \text{with} \quad k^2 &= \frac{Q-P+6m}{2Q} \quad \text{and} \quad \zeta = \frac{1}{2}\sqrt{\frac{Q}{P}}\,(\theta + \text{const.}) \end{aligned}\right\} \tag{28}$$

The orbit starts where $\mathrm{sn}^2\,\zeta_1 = (Q-P+2m)/(Q-P+6m)$ and ζ runs through K to $2K-\zeta_1$. Then $\mathrm{sn}^2\,(K-\zeta_1) = 2Q/(3P-6m+Q)$. It is more convenient to use the deflexion of the ray than the final value of θ. If the total deflexion is μ, the final value of θ as measured from the apse is $\frac{1}{2}\pi + \frac{1}{2}\mu$. If then $\mathrm{sn}\,(K-\zeta_1) = \sin\phi_1$, this is given by

$$\tfrac{1}{2}\pi + \tfrac{1}{2}\mu = 2\sqrt{(P/Q)}\,F(k, \phi_1), \tag{29}$$

where F is the elliptic integral of the first kind. Table 3 gives roughly some of the values of μ expressed in degrees.

TABLE 3

P/m	3·2	3·4	3·6	3·8	4	5	6	7	8	9	10	11	12
μ°	273	205	162	143	125	79	58	46	38	32	28	25	23
l/m	5·23	5·30	5·40	5·53	5·66	6·46	7·35	8·28	9·24	10·22	11·20	12·17	13·15

For large values of P the results to a first approximation are

$$\mu = 4m/P \quad \text{and} \quad l = P. \tag{30}$$

This corresponds to the well-known result that was verified by the eclipse expeditions.

For values of P near $3m$ the integral can be evaluated by the device of expressing it in terms of the dn function instead of the sn, and the leading term is

$$\tfrac{1}{2}\pi + \tfrac{1}{2}\mu = \ln\frac{36(2-\sqrt{3})\,m}{P-3m}. \tag{31}$$

The corresponding l is
$$3\sqrt{3}\ m+\frac{\sqrt{3}}{2}\frac{(P-3m)^2}{m},$$
and from this it follows that approximately
$$l = (5{\cdot}19+3{\cdot}48\,e^{-\mu})\,m. \tag{32}$$

9. Occultations

These results were worked out partly in order to study the effects of occultation; for example, what will be observed in a telescope when a star passes slowly behind the sun. First consider what will be seen in the case of a star exactly in the centre behind the sun. The star is at a great distance L_0 from the sun, and the telescope is at a distance L on the other side. There will be a ray going from star to telescope through an orbit with perihelion distance P. This P will determine both μ and l. The orbit can be regarded as composed of three parts, its two asymptotes and a short curved arc between them. The asymptotes are in the Galilean part of space, so that it is quite unambiguous to say that the ray starts from the star at angle $\chi_0 = l/L_0$ from the direct line joining star and sun, and that it approaches the telescope at angle $\chi = l/L$ to the direct line from the sun. Between these two lines it has been deflected through angle μ, and so $\chi_0+\chi = \mu$. All these angles are small, and it may be assumed that χ_0 is negligible on account of the great distance of the star. Hence, using (30) $\chi = P/L = 4m/P = 2\sqrt{(m/L)}$. In this special case there will be rays reaching the telescope in any plane through the sun, so that in the field of the telescope there will be seen a circle of this radius round the sun.

To give a numerical example, suppose that L represents the distance from the earth to the sun, that is $1{\cdot}5\times 10^8$ km. Then $m/L = 10^{-8}$ and the angular aperture of the circle would be 40″. Alternatively, so as to avoid considering very small angles, it may be supposed that the astronomer maps what he sees as though everything were in a plane through the sun. Then the star will be mapped as a circle of radius $2\sqrt{mL} = 3\times 10^4$ km.

Now consider the general case of a star somewhere near the direction of the sun and moving slowly so as to pass behind it. Let the angle between the radii drawn from sun to star and to telescope be $180°-\psi$. The ray advancing from the star is again at angle χ_0 to the line through the sun, and the ray approaching the telescope is again at angle χ to the line from the sun, and they are at angle μ to one another. So now
$$\psi = \chi_0+\chi-\mu. \tag{33}$$
As it approaches the line of the sun the star will appear to lag behind its real position, and this image will continue to be visible even after the star has got right beyond the sun, because there is always an orbit corresponding to any value of μ. Thus even after the star has gone well to the left of the sun there will still be an image to the right of it, which will continue there indefinitely. It may be called a 'ghost'. Its perihelion will be only a little greater than $3m$, and so its apparent position will be only a little greater than $3\sqrt{3}\ m$.

Conversely, while the star is still well to the right of the sun there will already be a ghost on its left just outside $3\sqrt{3}\,m$, and this will start slowly to move outwards. It will gradually increase in brilliance, and its motion will accelerate, and it will finally turn into the main star image. Furthermore, there will be secondary ghosts on both sides still closer to $3\sqrt{3}\,m$, which come from rays that have gone a complete circle round the sun before escaping to the telescope, and there will be still further ghosts that have done two or more rounds before escaping. It is to be expected that all these images will be very faint, and it is interesting to find their intensities as well as their positions.

Consider two adjacent rays emerging from the star at angular difference $\delta\chi_0$. These are to enter between the two edges of the telescope tube which are at angular distance $\delta\psi$ apart. The intensity of the light received is thus proportional to $\delta\chi_0/\delta\psi$. In the case of the direct image P is large and μ negligible, so that (33) gives for this a factor $\delta\chi_0/(\delta\chi_0+\delta\chi) = L/(L_0+L)$. In the case of the ghost, μ is much greater than either χ_0 or χ, so that the intensity is $-\delta\chi_0/\delta\mu = -\delta l/L_0\delta\mu$. Then since L_0 is much greater than L, if the direct image is taken to have unit brightness, the ghost will have brightness

$$I(\mu) = -\delta l/L\delta\mu. \tag{34}$$

Consider first a star rather near to the line of the sun and to its right. Then the μ corresponding to the first ghost on the left will be a small angle and (30) can be used. The result is $I(\mu) = 4m/L\mu^2$. To give a numerical value take m/L again as 10^{-8}. Then if $\mu = 1^\circ = 0{\cdot}0174$, $I(1^\circ) = 1{\cdot}3\times10^{-4}$. Star magnitudes are defined by the rule that the magnitude is $-\frac{5}{2}\log I$. So for a star of zero magnitude 1° to the right of the sun its first ghost on the left will appear as a star a little brighter than the tenth magnitude. It will be mapped where $l = 4m/\mu = 230m$.

For large values of μ, by (31) and (32), $I(\mu) = \frac{1}{2}\sqrt{3}\,(P-3m)^2/mL$ which is $(m/L)\,3{\cdot}48\,\mathrm{e}^{-\mu}$. It will be mapped at $l = (5{\cdot}19+3{\cdot}48\,\mathrm{e}^{-\mu})\,m$. Intermediate values can be roughly estimated by the use of table 3. Thus for the columns headed 9 and 10 the average of μ is 30°, while $\delta\mu = 4^\circ = 0{\cdot}070$ and $\delta l = 0{\cdot}98m$. Then $I(30^\circ) = 14m/L$. A star 30° to the right of the sun will give a ghost to its left of about the seventeenth magnitude, which will be mapped at $10{\cdot}71m$.

The brightness of a secondary ghost is simply $\mathrm{e}^{-2\pi} = 0{\cdot}0018$ of that of its primary, so that it is about seven magnitudes fainter. Its position is nearer to $3\sqrt{3}\,m$ by a factor also of $\mathrm{e}^{-2\pi}$. This has the curious consequence that the successive ghosts of a sequence crowd together more and more closely to such an extent as exactly to counterbalance their increasing feebleness.

This result seems to have no practical importance at all, but it is sufficiently curious to be examined a little further. Take a short range of finite length x outwards from $3\sqrt{3}\,m$. Let the first ghost to fall inside x have had deflexion μ, which for a small value of x may imply a path which has done several turns round the sun before escaping to the telescope. The brightness of this ghost is $(m/L)\,3{\cdot}48\,\mathrm{e}^{-\mu}$. Inside it there will be a second ghost of brightness $(m/L)\,3{\cdot}48\,\mathrm{e}^{-\mu-2\pi}$, a third of brightness $(m/L)\,3{\cdot}48\,\mathrm{e}^{-\mu-4\pi}$, and so on. The total light emerging from the range x will thus be $(m/L)\,3{\cdot}48\,\mathrm{e}^{-\mu}/(1-\mathrm{e}^{-2\pi})$. It remains to find the *average* relation of μ to x. The position of the first ghost is $3{\cdot}48\,\mathrm{e}^{-\mu}m$, which falls inside x if $\mu>\mu_0$, where $x = 3{\cdot}48\,\mathrm{e}^{-\mu_0}m$. On the other hand, μ must be less than $\mu_0+2\pi$, for otherwise it would count as the second ghost, not as the first. Now for a star equally likely to be anywhere round the sky the probability of its first ghost having deflexion between μ and $\mu+\mathrm{d}\mu$ is $\mathrm{d}\mu/2\pi$. Therefore the average value of $\mathrm{e}^{-\mu}$ is

$$\frac{1}{2\pi}\int_{\mu_0}^{\mu_0+2\pi}\mathrm{e}^{-\mu}\mathrm{d}\mu = \frac{1}{2\pi}\mathrm{e}^{-\mu_0}(1-\mathrm{e}^{-2\pi}).$$

Substituting this average for $\mathrm{e}^{-\mu}$ above, the total light coming from the range x is

$$\frac{m}{L}\frac{3{\cdot}48}{2\pi}\mathrm{e}^{-\mu_0} = \frac{x}{2\pi L}.$$

Thus the ghosts together form a line of light of average constant luminosity $1/2\pi L$ per unit length in the neighbourhood outside of $3\sqrt{3}\,m$. This curious result has been mentioned for its own interest, but it should be observed that averaging can hardly be justified for a series that is so strongly convergent, because the total brilliance will depend mostly on whether the first ghost falls just inside or just outside of x. Furthermore, this is all of course only ray theory, and no doubt it would be very different if wave theory were applied.

Finally, suppose that there is a roughly uniform star-field all round the sky, and consider what will be seen and mapped from the telescope. In the parts of the sky well away from the sun the field will be uniform in density. Nearing the sun the density will gradually increase a little, since each star will appear as though so to speak it had been repelled by the sun to an increasing degree. Round a circle of radius $3\sqrt{3}\,m$ there will be a faint glow due to all the ghosts, and this glow will have a uniform density in the immediate neighbourhood of this circle. To the area inside this circle only the sun itself can contribute light. The most obvious assumption is that its rays would emerge in straight lines, so that there would be a brilliant point of light surrounded by blackness. However, in general dynamics rays of light are reversible, and those captured by the sun have mostly entered in spiral orbits, and so it seems likely that there will also be emergent rays moving in spiral orbits. To decide whether this is so would call for deeper assumptions about the structure of the sun than any attempted here, but it seems likely that the sun will be seen as a blaze of light entirely filling the circle of radius $3\sqrt{3}\,m$.

10. Capture orbits

For the sake of completeness a short account will be given of the orbits that fail to escape from the sun.

The simplest are those where the cubic has three real roots; as in (11). Then the solution is given by the substitution

$$u = a+(a-b)\,(\mathrm{sn}^2\,\zeta/\mathrm{cn}^2\,\zeta)$$

with k and ζ as in (13) and (14). The orbit starts at aphelion at $u = a$, and evidently u is infinite at $\zeta = K$. The formula shows no trace of anything happening at the 'barrier', at $r = 2m$. Thus the example (15) now has solution

$$u = \frac{1}{3m}+\frac{1}{12m}\frac{\mathrm{sn}^2\,\zeta}{\mathrm{cn}^2\,\zeta}$$

with aphelion at $r = 3m$, that is to say well outside the barrier. Though we know from (6) that the whole solution fails at $r = 2m$, this corresponds to $\mathrm{sn}^2\zeta = \frac{2}{3}$, but there is nothing to show the importance of this value of ζ.

When there is only one real root for the cubic, all orbits must go into the sun. Let the equation be

$$(\mathrm{d}u/\mathrm{d}\theta)^2 = 2m(u-a)(u^2-2fu+g) \tag{35}$$

and $g > f^2$, to make the second factor essentially positive. The routine procedure is to select a sum of the two quadratic factors (here one of them is linear) so as to make a perfect square, and construct a transformation from that. Here the substitution is

$$u = a + q\frac{1-t}{1+t}, \tag{36}$$

where

$$q^2 = g - 2fa + a^2. \tag{37}$$

The quadratic factor in (35) now becomes

$$2q[(q+a-f) + t^2(q-a+f)]/(1+t)^2. \tag{38}$$

The equation is therefore solved by substituting

$$t = \mathrm{cn}\,\xi \quad (\mathrm{mod}\,h), \quad \text{where} \quad h^2 = \frac{q-a+f}{2q} \quad \text{and} \quad \xi = \theta\sqrt{2mq}.$$

Here a may be either positive or negative. If positive there is aphelion at $\theta = 0$ and evidently u becomes infinite at $\xi = 2H$, where H is the complete elliptic integral. If a is negative, the orbit starts at infinity where $\mathrm{cn}\,\xi = (q+a)/(q-a)$ and again goes on till $\xi = 2H$.

Just the same process as this can be used when $g < f^2$, so that there are three real roots, but the result looks quite different from the solution given in (12). The two forms are related together by a Landen transformation. If $a > f$ the three roots are now

$$a, f + \sqrt{[(a-f)^2 - q^2]}, \quad f - \sqrt{[(a-f)^2 - q^2]}$$

taken in decreasing order. It follows that $f = \frac{1}{2}(b+c)$ and $q = \sqrt{(a-b)(a-c)}$. The second factor in (38) must now be written as $(a-f+q)-(a-f-q)t^2$, and this demands the substitution $t = \mathrm{sn}\,\xi$ with $h^2 = (a-f-q)/(a-f+q)$. This yields

$$h = \frac{\sqrt{(a-b)} - \sqrt{(a-c)}}{\sqrt{(a-b)} + \sqrt{(a-c)}}$$

and the rule for a Landen transformation is $h = (1-k')/(1+k')$ which shows how the two expressions are related. In the case $f > a$, the order of the three roots is different, and a slightly different Landen transformation must be used, but it will not be given here.

The form (12) is evidently much more convenient than (36). Unfortunately, it does not seem possible to apply a similar transformation for the case where there is only one real root.

11. Negative gravity

In much of present-day physics electric charges can have either sign, and there are hints that something similar may be needed for gravitation, so it is worth briefly considering what the orbits will be like for negative mass. Let $m_1 = -m$. Then (8) becomes

$$\left(\frac{du}{d\theta}\right)^2 = -2m_1u^3 - u^2 - \frac{2m_1}{p^2}u + \frac{A}{p^2}. \tag{39}$$

For any solution to exist with positive u, A must evidently be positive. The equation then always has only one positive root, and may have either two negative roots or none. There is thus always a perihelion, say at $u = a$, and u ranges between this and zero.

If there are three real roots (noting that b and c are now negative) the solution is

$$u = a - (a-b)\,\mathrm{sn}^2\,\zeta \quad (\mathrm{mod}\,k), \tag{40}$$

where $k^2 = (a-b)/(a-c)$ and $\zeta = \theta\sqrt{\tfrac{1}{2}m_1(a-c)}$. Perihelion is at $\theta = 0$, and the asymptote is at

$$\mathrm{sn}^2\,\zeta_1 = a/(a-b).$$

If there is one real root the equation may be written as

$$(du/d\theta)^2 = -2m_1(u-a)(u^2-2fu+g). \tag{41}$$

The substitution
$$u = a - q\frac{1-\mathrm{cn}\,\xi}{1+\mathrm{cn}\,\xi} \tag{42}$$
now gives a quadratic factor

$$2q[(q-a+f)+\mathrm{cn}^2\,\xi(q+a-f)]/(1+\mathrm{cn}\,\xi)^2.$$

This makes $h^2 = (q+a-f)/2q$ and $\xi = \theta\sqrt{2m_1q}$. The apse is again at $\theta = 0$, and now the asymptote is where

$$\mathrm{cn}\,\xi_1 = (q-a)/(q+a).$$

Once again it does not seem possible to simplify the expressions in this case.

An examination of the intermediate case with two equal negative roots does not suggest any features of interest distinguishing the two different types of solution.

12. Algebraic note

The most interesting point in the present work is the fact that there can be no perihelion at less than $r = 3m$. This was shown by a dynamical proof in § 5, and that is probably the simplest, but a direct algebraic proof may be given.

It will suffice here to simplify by taking $m = 1$. Then there are three quantities $a \geqslant b \geqslant c$, of which a and b are positive. They obey the relations

$$a+b+c = \tfrac{1}{2}, \tag{43}$$

$$bc+ca+ab > 0, \tag{44}$$

and the problem is to see what conditions this imposes on b.

It is evident that $a \geqslant \frac{1}{6}$, and also that if c is positive $a < \frac{1}{2}$. In that case the most favourable condition for b will be when $c = 0$, and since b must be less than a its greatest possible value is $\frac{1}{4}$.

It will be shown, however, that a must be less than $\frac{1}{2}$ even when c is negative. The elimination of c from (44) gives

$$\tfrac{1}{2}a+\tfrac{1}{2}b-a^2-ab-b^2>0,$$

from which

$$(2b+a-\tfrac{1}{2})^2<\tfrac{1}{4}(1-2a)(1+6a). \tag{45}$$

From this it follows that $a<\frac{1}{2}$, and therefore also that $b>-c$.

By eliminating a instead of c from (44) an inequality like (45) results, from which it follows that $1+6c$ must be positive. Then, since $-c<\frac{1}{6}$, it follows that $a+b=\frac{1}{2}-c<\frac{2}{3}$. Since $b<a$, it follows that $b<\frac{1}{3}$. The three extreme values are thus $\frac{1}{3}$, $\frac{1}{3}$, $-\frac{1}{6}$ for a, b and c.

This confirms algebraically the proposition that there can be no perihelion at a distance less than $3m$.

References

Darwin, C. G. 1913 *Phil. Mag.* (6) **25**, 201.
Eddington, A. S. 1923 *The mathematical theory of relativity*, p. 86. Cambridge University Press.
Forsyth, A. R. 1920 *Proc. Roy. Soc.* A, **97**, 145.
McCrea, W. H. 1956 *Helv. Phys. Acta*, Suppl. 4, p. 121.

PROCEEDINGS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES