I

DYNAMICAL PROPERTIES OF COMETS

Introduction

Until the great advances of very recent years in almost every branch of theoretical astronomy, comets had come to be regarded as perhaps the most puzzling and mysterious of all the many types of heavenly bodies. This was not due to any great difficulty in the way of observing comets, as could be urged with a problem such as the structure of the solar corona, for they are extremely numerous, and some of them at times so extraordinarily bright as to outshine Venus herself and be easily visible in broad daylight with the naked eye. Their various properties are almost the direct antithesis of those of planets, and the two types of object are often referred to as 'the two solar families'. The times and intervals of occurrence of comets in the sky seem quite irregular compared with stars and planets, some remaining visible for months and others only for a few days. Brooks' comet of 1904 was observable for more than twelve months, Comets 1889 and 1936 I were observed for over twenty-four months, Comet 1927 IV for four years, while Grigg's comet of 1901 was observable for only twelve days, and a few very faint comets have been observed perhaps only on a single occasion.

The word 'comet', according to the Oxford Dictionary, comes from the Greek word *κομήτης*, which means a 'long-haired star'—presumably a comet—and is doubtless also connected with *coma*, the Latin word for 'hair'. The name thus fancifully refers to the most obvious feature of many comets, which resemble a star embedded in a mildly luminous fog from which there appears to be carried away a long streaming trail itself faintly luminous. It is this tail that no doubt suggested to the ancients long tresses of feminine hair streaming in the wind. Naturally the brightest comets (as seen from the Earth) and those with the most extensive tails are the ones that have made the strongest impressions, and so appear more typical of comets than they may really be; in fact, most comets are very faint objects detectable only with powerful telescopes. We should perhaps have said 'most observable comets', for there is little room to doubt, from the steady rate of discovery of new comets, that there must be well over a hundred thousand in the solar system at present unobservable but which may all eventually become observable when they arrive nearer the sun, and it may well be that there are in addition a comparable number that move round the sun always at such great distances that they may never become visible at all to us on Earth.

Early notions

Although the Babylonians seem to have suspected that comets moved analogously to planets, the earliest ideas as to the location of comets in space were extremely inaccurate, and the assertions of some of the most learned men among the ancients can now be recognized as little better than unverified guesswork. Anaxagoras and Democritus attributed comets to 'the combined splendour of a concourse of planets'. Even so renowned a personage as Aristotle maintained that they were some kind of exhalations, whatever that may have meant, from the Earth itself that had somehow reached the upper part of the atmosphere and there in some unaccountable way become inflamed to make themselves luminous. Absurd as this idea now seems, it nevertheless may have appeared at the time to have some observational foundation. For comets are generally brightest when nearest the sun, at which time they are of course best visible shortly after sunset or shortly before sunrise, and since the tails necessarily point away from the sun these would usually appear to be more or less upright in the atmosphere, rather like the rising flame of a torch. In justice to Seneca, however, it should perhaps be said that he did not favour this idea, but he was distinctly in the minority, and Aristotle's view seems to have been so widely accepted (perhaps in itself a suspicious feature of any conjectural hypothesis) that Ptolemy for example, according to the Almagest, did not regard comets as among the heavenly bodies at all.

Though Cardan had already concluded that comets must lie far beyond the moon, the first definite demonstration of their truly celestial character was due to the famous Danish astronomer Tycho Brahe who found by careful measurements that the apparent position of the daylight comet of 1577 as seen from his observatory at Hven in the Baltic sea was indistinguishable from its direction as seen from Prague, some 400 miles or so to the south. Tycho had no difficulty in perceiving the implication of this. At such separated stations a body like the moon, for instance, though very much farther away than any part of the Earth's upper atmosphere, nevertheless shows a considerable apparent angular difference of position (at times as much as five minutes of arc relative to the background stars, which are so distant as to show negligible parallax shift, as the effect is termed). From this it could be inferred with certainty that the comet was far more distant than the moon, and hence that comets were truly celestial objects.

There yet remained the problem of deciding how the comets moved with reference to the sun, for move they certainly did, but their apparent motion as seen from the Earth in most cases little resembled that of the planets, the elongated cometary paths lying partly inside and partly outside the Earth's orbit producing very mysterious differences. Tycho himself suggested that this particular comet of 1577 might move in a circle somewhere outside the orbit of Venus, but Kepler on the other hand was of the opinion that all comets move in straight lines. It had not yet come to be recognized that comets could return—identity of comets involved in widely separated apparitions rests on comparison of orbits—so it did not occur to Kepler to extend his laws of planetary motion to comets. The suggestion that comets might describe more or less parabolic orbits was put forward quite conjecturally by the German astronomer Hevelius, who attributed the necessary departures from rectilinear motion to resistance by the ether, and the hypothesis was shown to hold, at any rate for the comet of 1681 in particular, by Doerfel, one of his pupils. The theory of gravitation was soon afterwards announced by Newton, and by its means the problem of the cometary orbits was finally solved by Halley, who showed from a discussion of twenty-four comets observed between 1337 and 1698 that within the limits of accuracy of the measurements all these appeared to move in strict accordance with the law of gravitation. Apart from certain slight exceptions (to be discussed later) that have since been established, there is no reason to doubt the general applicability of gravitational theory to cometary orbits, and nowadays all orbits are computed on this basis without the smallest doubt as to its validity for such purposes.

Where the causes of cometary tails are concerned, the early ideas were almost of necessity purely conjectural. Tycho and other contemporaries believed they were mere optical appearances, free of any material nature, formed somehow by the passage of the sun's light through the comet itself. Hooke supposed that the impulsion of solar rays on the comet drove off imponderable material not subject to attraction by the sun but to repulsion, thereby anticipating the general nature of the correct explanation by means of light pressure. Electric and magnetic actions were suggested by Bessel and Olbers, while the combined effects of gravitation and a hypothetical repulsion were examined by Roche on the assumption that a comet consisted of a homogeneous gaseous atmosphere retained by a gravitating nucleus.

Possible forms of orbits

If the attractive forces of the planets are ignored in comparison with the dominating influence of the sun, the path of any small body moving near the sun must have the form of a fixed 'conic section' (that is, one of the various curves in which the surface of a circular cone is cut by a plane) and moreover the sun must always lie at its focus. Although not fundamentally different analytically, the types of curve so obtained fall into three kinds, namely ellipses, parabolas, and hyperbolas, and there are important dynamical distinctions to be made between these. The ellipses are closed curves of finite extent, so that a body describing an ellipse about the sun must always remain in its neighbourhood (see Fig. 1). The parabolas are, as it were, closed at one end but just extend to infinity at the other, and a particle moving about the sun in a parabola has

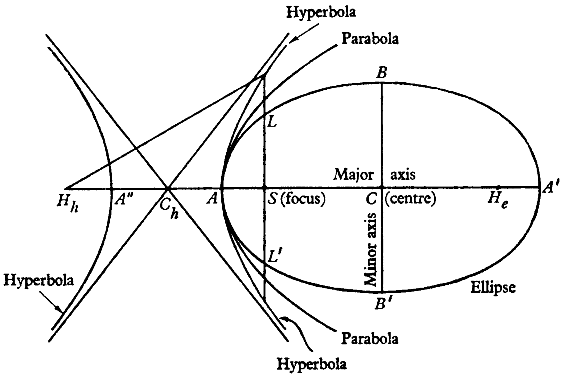

Fig. 1. Diagram of the three possible types of orbit, ellipse, parabola, and hyperbola.

AA' = major axis of the ellipse, S and H_e are its foci. SH_e/AA' = eccentricity (about 3/4 for ellipse shown). BB' = minor axis, LSL' = latus rectum.

AA'' = major axis of the hyperbola, S and H_h are its foci. SH_h/AA'' = eccentricity (about 7/4 for hyperbola shown).

Each of the three paths is a possible orbit under an attraction *towards* S (the sun).

Under a repulsive force from the sun S is the outer focus and the outer branch of the hyperbola is the only possible orbit.

a, e and T (the instant at which the comet passes through A) are the elements determining the path and time within the orbital plane.

just sufficient energy to escape to infinity, though it would arrive there with zero speed and take an infinite time to do so. The hyperbolas may be said to extend beyond infinity (in a complete hyperbola a disconnected branch having precisely the same form lies elsewhere in the plane), and a body moving in a hyperbolic path can escape completely from the sun and retain a finite velocity. Conversely, a particle moving towards the sun with finite velocity when at great distance will eventually sweep round it in a hyperbolic orbit, while a particle falling from rest at very great distance will have a parabolic orbit. The least decrease or increase of speed of a parabolic comet renders its orbit an ellipse or a hyperbola respectively. It is extremely important to realize, however, that the type of orbit of a body within the solar system, while possibly affording some indication of its earlier history when it can be properly interpreted, by no means provides any immediate guide in so far as the origin of the object is concerned.

When allowance is made for the presence of the planets, the paths immediately cease to be accurate ellipses, parabolas, or hyperbolas, and become highly complicated three-dimensional curves, because the force system now contains numerous additional contributions above the simple inverse square attraction to the sun. At any instant, however, a so-called osculating orbit can be defined for a body with given position and velocity as the path that the body would continue in if the planetary attraction suddenly ceased. These osculating orbits are of course perfect conic sections, and are found to bear close resemblance to the real path (because planetary action is usually small), but even so they gradually change and can represent the actual path only for a limited time, of less or greater extent, according to the degree of accuracy required and the particular disturbing effects involved.

It is possible also to arrange all the various shapes of orbits in a single series in which each one has a definite degree of elongation from the circular form which lies at one end of the set. This quality is termed the eccentricity of the curve and its value is usually denoted by the letter e. It increases all the way from 0 for a circle up to precisely 1 for a parabola and thence through all values greater than unity as the series of hyperbolas is described. What in fact Halley found for the several comets he investigated was that e was always quite near to 1, so that their paths were very close to parabolas, a characteristic feature of almost all cometary orbits. There are comets with orbits of moderate eccentricities, but they are comparatively few in number and, as we shall see, have acquired such orbits through special causes. If we except such comets, the standard shape of cometary orbits may be regarded as always nearly parabolic.

The size of the orbit depends on the quantity $2a = AA'$ (Fig. 1) measuring its major axis, so that between them a and e settle the size and shape of the path. Now for a parabola the major axis is obviously not a suitable measure of its size, since it is infinite, and even for an ellipse approaching parabolic shape

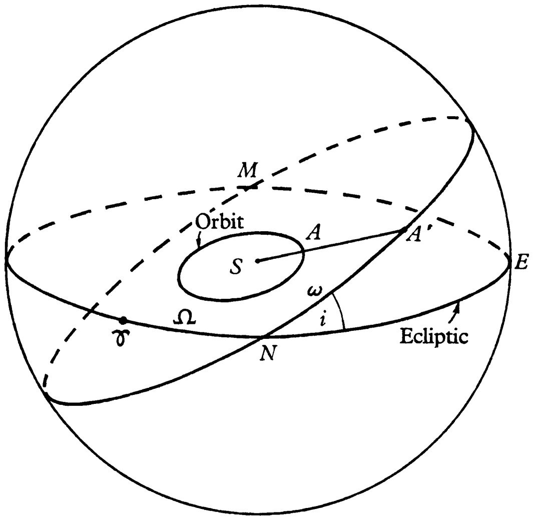

Fig. 2. The centre of the celestial sphere is S, the sun. ♈ is the vernal equinox. N is the ascending node. SA' is the direction of perihelion, A. $NA'M$ is the great circle defined by the orbital plane. $A'NE = i$ = inclination of orbit to the ecliptic. ♈$N = \Omega$ = longitude of the node N. $NA' = \omega$ = angular distance of A' from the node. Ω, ω, and i are the three elements fixing the position of the orbit in space.

it is also unsuitable because it is nearly always very large compared, say, with the size of the Earth's orbit (which provides a standard within the solar system), and is also difficult to measure accurately. It is found more convenient to use the quantity $q = a(1-e)$ which in all cases measures the least distance, SA, of the body from the sun. To fix the position of the comet completely it is necessary also to settle the time at which it passes, or will eventually pass, some specified point of its path, normally the perihelion point A. The place it reaches at any other time, and the time it takes to reach any other point, are then accurately calculable entirely by means of dynamical theory. To relate the position of the comet at any instant to that of the Earth, which is what is involved in predicting its observable future path, the precise position in space of its elliptic or parabolic orbit has also to be determined. For this purpose the plane of the Earth's motion round the sun—the ecliptic—is adopted as a reference plane; and then the inclination to this of the plane of the comet's orbit, and the angular extent to which this latter is slewed round from some standard position (the direction of the sun on 21 March, the so-called vernal equinox point) are measured, and fix the orbit completely (Fig. 2). These quantities, six in all, termed the elements of its orbit, can be tabulated for each comet, and as already mentioned, a knowledge of them suffices for the position of the comet at any future date to be calculated. The following are a few instances of typical elements of cometary orbits.

TABLE I. ORBITAL ELEMENTS OF COMETS

	1942 VII P/Oterma	1947 i P/Encke	1922 II Baade (hyperbolic)	Halley
T (perihelion passage)	1942 Aug. 21·8085	1947 Nov. 26·3295	1922 Oct.26·03015	1910 April 19·65
$q = a(1-e)$	3·389625	0·341130	2·25877	0·587200
e	0·144425	0·846197	1·000865	0·967281
Period (years)	7·89	3·284	—	76·03
ω	354°·8058	185°·1985	118°·3072	111°·6989
Ω	155°·1708	334°·7223	220°·4825	57°·27
i (Equinox 1950)	3°·9899	12°·3505	51°·4639	162°·2117

ω = angular distance of perihelion from node,
Ω = longitude of node ΥN.

Determination of orbits

When a comet is first discovered, the solution of the converse problem is the computer's initial task, and it is necessary to determine its six elements from a limited number of more or less accurate angular observations of the comet at different instants, known with high precision, and the wider spaced these can be the better for the final accuracy and reliability of the results. When a comet is newly discovered wide spacing is obviously not possible at first, since only observations covering a few days' motion will be available. For this reason a provisional orbit of only moderate accuracy is usually calculated to begin with, simply to help keep track of the comet for the purpose of securing more observations; the computation of the final definitive orbit is not made until practically all useful observations are available, and these may cover an interval of a year or even far longer. It must be remembered, however, that all computed orbits are necessarily based on observations over a limited time, and as such can represent only an approximation, even if the observations themselves were perfect, to the temporary osculating orbit—that is to the accurate conic section that the comet would describe (if it were a point mass) were all planetary disturbances suddenly to cease. Residuals between past observations and the computed orbit of anything up to a few minutes of arc are known for some comets, and where predictions of subsequent returns are concerned errors measured in days are by no means unknown especially if planetary influences have meanwhile been large on the comet. Up to the present time the orbits of something approaching 1000 comets have been investigated in this way; the results provide sufficient material to enable a completely adequate general idea of the distribution of the orbits in space to be formed.

From this evidence it appears that about three-quarters of all observed comets move in approximately parabolic orbits ($e = 1$) while about one-quarter move in paths that are definitely elliptic ($e < 1$) though the majority of these nevertheless have high eccentricities. The remaining few per cent have slightly hyperbolic orbits (that is, e just greater than 1) in so far as can be ascertained from their observable motion near the sun. But there is good reason to suppose that the actual complete paths even of these comets do not in reality extend to infinity at all but, while certainly receding to great distance, remain always within the sun's influence. The perihelion distances (q) of known comets range from values just exceeding the sun's radius up to values greater than 4 a.u., the largest value to date being 5·50 a.u. for 1925 II, but comets with large q are inevitably very faint objects likely to be unobserved. All known comets with very small q (of the order of a few solar radii) are of long period.

The reasons why in some cases the path is found to be hyperbolic arise first from the extremely limited arc over which observations are sometimes only possible, and second from the additional attractive effects of the planets which may accelerate a comet that would otherwise be moving in elliptic motion into an apparently hyperbolic path. A few such orbits, apparently hyperbolic near the sun, have been extended further outwards from the observable part and backwards in time, by calculations making due allowance for the influence of the planets (Jupiter is usually the main perturbing agency), and in every case it has been found that the comet has in fact come in from a finite distance, and is therefore to be regarded as a reappearance of a permanent member of the solar system as far as its orbital motion is concerned. According to investigations by Strömgren of about twenty orbits, which near the sun were apparently hyperbolic, at great distances the speeds were always less than the velocity of escape, so that the orbits were on the whole definitely elliptical (see p. 105). Thus it can be concluded beyond any question from all this in regard to the origin of comets, that they do not simply enter the solar system from outside under free gravitational motion.

Apart from such dynamical effects of the planets, it is found that near perihelion, the point of the path closest to the sun, a given set of positions measured with even fairly high accuracy may turn out to be equally well representable by any one of a range of curves differing slightly in eccentricity. This feature of cometary orbits is illustrated in Fig. 1 which shows three different curves, an ellipse, a parabola, and a hyperbola, for each of which the arc near A is practically the same. The calculated orbit of a comet whose position happened to be capable