13.1: Introduction to Calculating Orbital Elements

We have seen in Chapter 10 how to calculate an ephemeris from the orbital elements. This chapter deals with the rather more difficult problem of determining the orbital elements from the observations.

We saw in Chapter 2 how to fit an ellipse (or other conic section) to five points in a plane. In the case of a planetary orbit, we need also to know the orientation of the plane, which will require two further bits of information. Thus we should be able to determine the shape, size and orientation of the ellipse from seven pieces of information.

This, however, is not quite the same problem facing us in the determination of a planetary orbit. Most importantly, we do not know *all* of the coordinates of the planet at the time of *any* of the observations. We know two of the coordinates – namely the right ascension and declination – but we have no idea at all of the distance. All that an observation gives us is the direction to the planet in the sky at a given instant of time. Finding the geocentric distance at the time of a given observation is indeed one of the more difficult tasks; once we have managed to do that, we have broken the back of the problem.

However, although we do not know the geocentric (or heliocentric) distances, we do have some additional information to help us. For one thing, we know where one of the foci of the conic section is. The Sun occupies one of them – though we don't immediately know which one. Also, we know the instant of time of each observation, and we know that the radius vector sweeps out equal areas in equal times. This important keplerian law is of great value in computing an orbit.

To determine an orbit, we have to determine a set of six orbital elements. These are, as previously described, a, e, i, Ω, ω and T for a sensibly elliptic orbit; for an orbit of low eccentricity one generally substitutes an angle such as M_0, the mean anomaly at the epoch, for T. Thus we can calculate the orbit from six pieces of information. We saw in Chapter 10 how to do this if we know the three heliocentric spatial coordinates and the three heliocentric velocity components – but this again is not quite the problem facing us, because we certainly do not know any of these data for a newly-discovered planet.

If, however, we have three suitably-spaced observations, in which we have measured three directions (α, δ) at three instants of time, then we have six data, from which it may be possible to calculate the six orbital elements. It should be mentioned, however, that three observations are *necessary* to obtain a credible solution, but they may not always be *sufficient*. Should all three observations, for example, be on the ecliptic, or near to a stationary point, or if the planet is moving almost directly towards us for a while and consequently hardly appears to move in the sky, it may not be possible to obtain a credible solution. Or again, observations always have some error associated with them, and small observational errors may under some circumstances translate into a wide range of possible solutions, or it may not even be possible to fit a single set of elements to the slightly erroneous observations.

In recent years, the computation of the orbits of near-Earth asteroids has been a matter of interest for the public press, who are likely to pounce on any suggestion that the observations might have been "erroneous" and the orbit "wrong" – as if they were unaware that all scientific measurement always have error associated with them. There is a failure to distinguish *errors* from *mistakes*.

When a new minor planet or asteroid is discovered, as soon as the requisite minimum number of observations have been made that enable an approximate orbit to be computed, the elements and an ephemeris are distributed to observers. The purpose of this *preliminary orbit* is not to tell us whether planet Earth is about to be destroyed by a cataclysmic collision with a near-Earth asteroid, but is simply to supply observers with a good enough ephemeris that will enable them to find the asteroid and hence to supply additional observations. Everyone who is actively involved in the process of observing asteroids or computing their orbits either knows or ought to know this, just as he also knows or ought to know that, as additional observations come in, the orbit will be *revised* and *differential corrections* will be made to the elements. Further, the computed orbit is generally an *osculating orbit*, and the elements are *osculating elements* for a particular *epoch of osculation*. In order to allow for planetary perturbations, the epoch of osculation is changed every 200 days, and new osculating elements are calculated. All of this is routine and is to be expected. And yet there has been an unfortunate tendency in recent years for not only the press but also for a number of persons who would speak for the scientific community, but who may not themselves be experienced in orbital computations, to attribute the various necessary revisions to an orbit to "mistakes" or "incompetence" by experienced orbit computers.

When all the observations for a particular apparition have been amassed, and no more are expected for that apparition, a definitive orbit for that apparition is calculated from all available observations. Even then, there will be small variations in the elements obtained by different computers. This is because, among other things, each observation has to be critically assessed and weighted. Some observations may be photographic; the majority these days will be higher-precision CCD observations, which will receive a higher weight. Observations will have been made with a variety of telescopes with very different focal lengths, and there will be variations in the experience of the observers involved. Some observations will have been made in a great hurry in the night immediately following a new discovery. Such observations are valuable for computing the preliminary orbit, but may merit less weight in the definitive orbit. There is no unique way for dealing with such problems, and if two computers come up with slightly different answers as a result of weighting the observations differently it does not mean than one of them is “right” and that the other has made a “mistake”. All of this should be very obvious, though some words that have been spoken or written in recent years suggest that it bears repeating.

There are a number of small problems involving the original raw observations. One is that the instant of time of an observation is recorded and reported by an observer in Universal Time. This is the correct thing for an observer to do, and is what is expected of him or her. The computer, however, uses as the argument for the orbital calculation the best representation of a uniformly-flowing dynamical time, which at present is TT, or Terrestrial Time (see chapter 7). The difference for the current year is never known exactly, but has to be estimated. Another difficulty is that observations are not made from the centre of Earth, but from some point on the surface of Earth – a point that is moving as Earth rotates. Thus a small parallactic correction has to be made to the observations – but we do not know how large this correction is until we know the distance of the planet. Or again, the computer needs to know the position of the planet when the sunlight reflected from it left the planet, not when the light eventually arrived at Earth twenty or so minutes later – but we do not know how large the light travel-time correction is until we know the distance of the planet.

There is evidently a good deal involved in computing orbits, and this could be a very long chapter indeed, and never written to perfection to cover all contingencies. In order to get started, however, I shall initially restrict the scope of this chapter to the basic problem of computing elliptical elements from three observations. If and when the spirit moves me I may at a later date expand the chapter to include parabolic and hyperbolic orbits, although the latter pose special problems. Computing hyperbolic elements is in principle no more difficult than computing elliptic orbits; in practice, however, any solar system orbits that are sensibly hyperbolic have been subject to relatively large planetary perturbations, and so the problem in practice is not at all a simple one. Carrying out differential corrections to a preliminary orbit is also something that will have to be left to a later date.

In the sections that follow, I am much indebted to Carlos Montenegro of Argentina who went line-by-line with me through the numerical calculations, resulting in a number of corrections to the original text. Any remaining mistakes (I hope there are few, if any) are my own responsibility.