1
Structure of the Solar System

There's not the smallest orb which thou behold'st
But in his motion like an angel sings,
Still quiring to the young-eyed cherubins;
Such harmony is in immortal souls;

William Shakespeare, *Merchant of Venice, V, i*

1.1 Introduction

It is a laudable human pursuit to try to perceive order out of the apparent randomness of nature; science is, after all, an attempt to make sense of the world around us. Moving against the background of the "fixed" stars, the regularity of the Moon and planets demanded a dynamical explanation.

The history of astronomy is the history of a growing awareness of our position (or lack of it) in the universe. Observing, exploring, and ultimately understanding our solar system is the first step towards understanding the rest of the universe. The key discovery in this process was Newton's formulation of the universal law of gravitation; this made sense of the orbits of planets, satellites, and comets, and their future motion could be predicted: The Newtonian universe was a deterministic system. The *Voyager* missions increased our knowledge of the outer solar system by several orders of magnitude, and yet they would not have been possible without knowledge of Newton's laws and their consequences. However, advances in mathematics and computer technology have now revealed that, even though our system is deterministic, it is not necessarily predictable. The study of nonlinear dynamics has revealed a solar system even more intricately structured than Newton could have imagined.

In this chapter we review some of the observations that have motivated the quest for an understanding of the dynamical structure of the solar system.

1

1.2 The Belief in Number

The desire to perceive order in the distribution of objects in the solar system can be traced to early Greece, although it may have had its roots in Babylonian astronomy. Anaximander of Miletus (611–547 B.C.) claimed that the relative distances of the stars, Moon, and Sun from the Earth were in the ratio 1:2:3 (Bernal 1969). The importance of whole numbers to members of the Pythagorean school led them to believe that the distances of the heavenly bodies from the Earth corresponded to a sequence of musical notes and this gave rise to the concept of the "harmony of the spheres". This, in turn, influenced Plato (427–347 B.C.) whose work was to have a great effect on Johannes Kepler nearly two thousand years later (Field 1988).

Kepler was obsessed with the belief that numbers and geometry could be used to explain the spacing of the planetary orbits. He firmly believed in the Copernican rather than the Ptolemaic system, but his views on planetary orbits had foundations in numerology and astrology (Field 1988) rather than scientific method. In the first edition of his book *Mysterium Cosmographicum*, Kepler (1596) described his model of the solar system, which consisted of the six known planets (Mercury, Venus, Earth, Mars, Jupiter, and Saturn) moving within spherical shells whose inner and outer surfaces had precise separations determined by the circumspheres and inspheres of the five regular polyhedra (cube, tetrahedron, dodecahedron, icosahedron, and octahedron). Kepler believed that the widths of these shells were related to the orbital eccentricities. This is illustrated in Fig. 1.1 for the outer solar system. He also developed a similar theory to explain the relative spacings of the newly discovered moons of Jupiter (Kepler 1610). Between the first and second editions of *Mysterium Cosmographicum*, Kepler had empirically deduced the first two of his laws of planetary motion (Kepler 1609), and the notes accompanying the second edition (Kepler 1621) make it clear that his belief in astrology was waning (Field 1988). Although it is unlikely that he had a literal belief in musical notes emanating from the planets, Kepler persisted in his search for harmonic relationships between orbits.

Despite its metaphysical origins, Kepler's geometrical model was a surprisingly good fit to the available data (Field 1988). Although Kepler looked unsuccessfully for simple numerical relationships between the orbital distances of the planets, it was his fascination with numbers that ultimately led to the discovery of his third law of motion, which relates the orbital period of a planet to its average distance from the Sun. On 15 May 1618 he became convinced that *"it is most certain and most exact that the proportion between the periods of any two planets is precisely three halves the proportion of the mean distances"*. Kepler's most important legacy was not his intricate geometrical model of the spacing of the planets, but his empirical derivation of his laws of planetary motion.
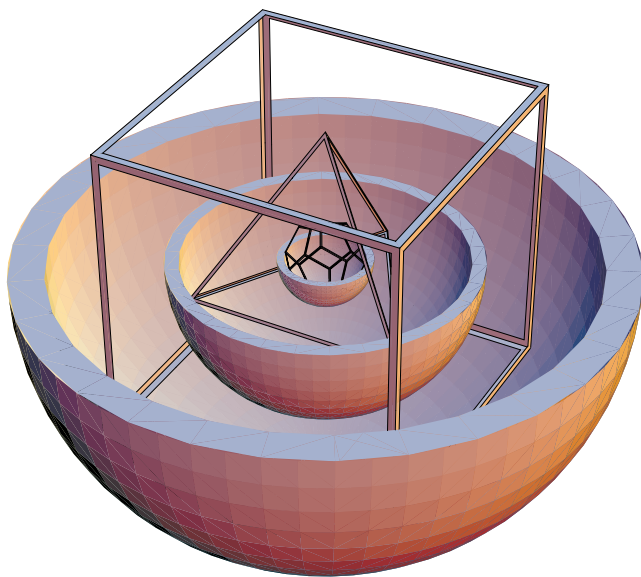
Fig. 1.1. Kepler's geometrical model of the relative distances of the planets. Each planet moved within a spherical shell with inner and outer radii defined by the limiting spheres of the regular polyhedra. For the outer solar system the orbits of Saturn, Jupiter, and Mars are on spheres separated by a cube, a tetrahedron, and a dodecahedron.

1.3 Kepler's Laws of Planetary Motion

Kepler (1609, 1619) derived his three laws of planetary motion using an empirical approach. From observations, including those made by Tycho Brahe, Kepler deduced that:

1) The planets move in ellipses with the Sun at one focus.
2) A radius vector from the Sun to a planet sweeps out equal areas in equal times.
3) The square of the orbital period of a planet is proportional to the cube of its semi-major axis.

The geometry implied by the first two laws is illustrated in Fig. 1.2. An ellipse has two foci and according to the first law the Sun occupies one focus while the other one is empty (Fig. 1.2a). In Fig. 1.2b each shaded region represents the area swept out by a line from the Sun to an orbiting planet in equal time intervals, and the second law states that these areas are equal. The geometry of the ellipse will be considered further in Chapter 2.

Half the length of the long axis of the ellipse is called the semi-major axis, a (see Fig. 1.2a). Kepler's third law relates a to the period T of the planet's orbit. He deduced that $T^2 \propto a^3$, so that if two planets have semi-major axes a_1 and
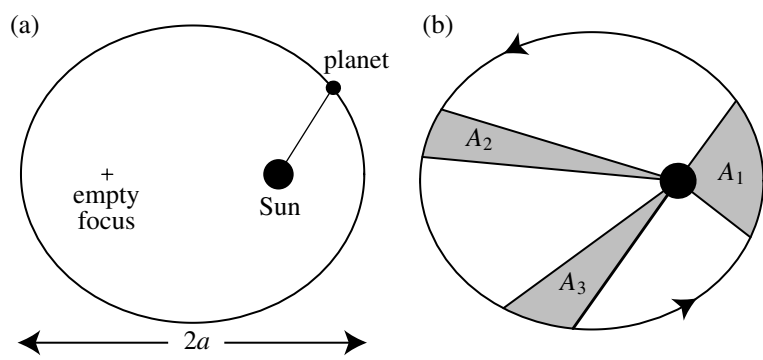
Fig. 1.2. The geometry implied by Kepler's first two laws of planetary motion for an eccentricity of 0.5. (a) The Sun occupies one of the two foci of the elliptical path traced by the planet; the other focus is empty. (b) The regions A_1, A_2, and A_3 denote equal areas swept out in equal times by the radius vector.

a_2 and periods T_1 and T_2, then $T_1/T_2 = (a_1/a_2)^{3/2}$, which is consistent with his original formulation of the law.

It is important to remember that Kepler's laws were purely empirical: He had no physical understanding of why the planets obeyed these laws, although he did propose a "magnetic vortex" to explain planetary orbits.

1.4 Newton's Universal Law of Gravitation

In the seventeenth century Isaac Newton (1687) proved that a simple, inverse square law of force gives rise to all motion in the solar system. There is good evidence that Robert Hooke, a contemporary and rival of Newton, had proposed the inverse square law of force before Newton (Westfall 1980) but Newton's great achievement was to show that Kepler's laws of motion are a natural consequence of this force and that the resulting motion is described by a conic section.

In scalar form, Newton proposed that the magnitude of the force F between any two masses in the universe, m_1 and m_2, separated by a distance d is given by

$$F = \mathcal{G}\frac{m_1 m_2}{d^2}, \tag{1.1}$$

where $\mathcal{G}$ is the universal constant of gravitation.

In his *Principia* Newton (1687) also propounded his three laws of motion:

1) Bodies remain in a state of rest or uniform motion in a straight line unless acted upon by a force.
2) The force experienced by a body is equal to the rate of change of its momentum.
3) To every action there is an equal and opposite reaction.

The combination of these laws with the universal law of gravitation was to have a profound effect on our understanding of the universe. Although Newton "*stood on the shoulders of giants*" such as Copernicus, Kepler, and Galileo, his discoveries revolutionised science in general and dynamical astronomy in particular. By extending Newtonian gravitation to more than two bodies it was shown that the mutual planetary interactions result in ellipses that are no longer fixed. Instead the orbits of the planets slowly rotate or *precess* in space over timescales of $\sim 10^5$ y. For example, calculations based on the Newtonian model have shown that the orbit of Mercury should currently be precessing at a rate of $531''$ century^{-1}.

However, observations show that Mercury's orbit is precessing at a rate that is $43''$ century^{-1} greater than that predicted by the Newtonian model. We now know that Newton's universal law of gravitation is only an approximation, albeit a very good one, and that a better model of gravity is given by Einstein's general theory of relativity. Applied to the precession of Mercury's perihelion this predicts an additional contribution of $43''$ century^{-1}, and the combination of the relativistic contribution to the Newtonian model gives an agreement that is within the current observational limitations (Roseveare 1982).

1.5 The Titius–Bode "Law"

The regularity in the spacing of the planetary orbits led to the formulation of a simple mnemonic by Johann Titius in 1766 (Nieto 1972). Titius pointed out that the mean distance d in astronomical units (AU) from the Sun to each of the six known planets was well approximated by the equation

$$d = 0.4 + 0.3\,(2^i), \quad \text{where } i = -\infty, 0, 1, 2, 4, 5. \tag{1.2}$$

The "law" was soon popularised by Johann Bode and is now commonly referred to as Bode's law. Although the "law" had no physical foundation, Bode claimed that an undiscovered planet orbited at the $i = 3$ location. The subsequent discovery of the planet Uranus in 1781 at 19.18 AU ($i = 6$) and the first asteroid (1) Ceres in 1801 at 2.77 AU ($i = 3$) were considered triumphs of the "law" (see Table 1.1 and Nieto 1972).

Such was the success of the Titius–Bode "law" that both John Couch Adams (1847) and Urbain Le Verrier (1847) used it as a basis for their calculations on the predicted orbit of the eighth planet (Grosser 1979). Using $i = 7$ in Eq. (1.2) the "law" predicts a semi-major axis of 38.8 AU; the planet Neptune was discovered in 1846, but it has a semi-major axis of 30.1 AU. The breakdown of the "law" was complete with the discovery of Pluto in 1929 at 39.4 AU compared with a predicted distance of 77.2 AU ($i = 8$). Of course, it could be argued that Pluto is too small to be considered a planet and should therefore be excluded from the

Table 1.1. *A comparison of the semi-major axes of the planets, including the minor planet Ceres, with the values predicted by the Titius–Bode law.*

Planet	i	Semi-major Axis (AU)	Titius–Bode Law (AU)
Mercury	$-\infty$	0.39	0.4
Venus	0	0.72	0.7
Earth	1	1.00	1.0
Mars	2	1.52	1.6
Ceres	3	2.77	2.8
Jupiter	4	5.20	5.2
Saturn	5	9.54	10.0
Uranus	6	19.18	19.6
Neptune	7	30.06	38.8
Pluto	8	39.44	77.2

calculation. However, if every value of i is filled we should expect an infinite number of planets between Mercury and Venus!

Some of the regular satellite systems of the outer planets appear to have nonrandom distributions because there are a number of simple numerical relationships between their periods (see Sects. 1.6 and 1.7). In an attempt to incorporate orbital resonances into a Titius–Bode "law", Dermott (1972, 1973) discussed the significance of a modified form of the law using a two-parameter, geometric progression of orbital periods rather than orbital distances. Writing

$$T_i = T_0 A^i, \tag{1.3}$$

where the satellites are numbered in order of increasing period, T_i is the predicted orbital period of the ith satellite, and T_0 and A are arbitrary constants, Dermott pointed out that if the possibility of "empty orbitals" is excluded from the system, then such a relationship can only hold for the regular satellites of Uranus. Taking the logarithm of each side of Eq. (1.3), we have

$$\log T_i = \log T_0 + i \log A. \tag{1.4}$$

The observed data can then be fitted to this model using a standard linear regression technique where one measure of the goodness of fit is the root mean square (rms) value, χ, of the residuals. This is given by

$$\chi^2 = \frac{1}{n} \sum_{i=1}^{n} (\log T_i - \log T_0 - i \log A)^2, \tag{1.5}$$

where T_i is now understood to be the observed orbital period. The comparison of the predictions with the observed data shown in Table 1.2 and Fig. 1.3 seem remarkably favourable. Using the fitted parameters of $T_0 = 0.7919$, $A = 1.777$

Table 1.2. *A comparison of the observed orbital period (in days) of the large satellites of
Uranus with values calculated using a geometric progression of the form of Eq. (1.3).*

Satellite	i	T_i (obs.)	T_i (calc.)
Miranda	1	1.413	1.407
Ariel	2	2.520	2.500
Umbriel	3	4.144	4.442
Titania	4	8.706	7.893
Oberon	5	13.46	14.02

we obtain $\chi = 0.0247$. However, it is not enough to be impressed by the
seemingly remarkable fits. We need to subject the data to a statistical test and
to calculate whether or not the value of χ is small enough to be statistically
significant. This can be addressed using a Monte Carlo approach.

Using a technique similar to that of Dermott (1972, 1973) we have generated
a series of sets of five satellites subject to certain restrictions on the distribution
of their orbital periods. The innermost satellite was always chosen to have the
observed orbital period of Miranda. The remaining periods were then generated
using the relationship

$$\frac{T_{i+1}}{T_i} = L + x_i(U - L) \qquad (i = 1, 2, 3, 4), \tag{1.6}$$

where x_i is a random number in the range $0 \leq x_i \leq 1$, and L and U ($> L$) represent
the fixed lower and upper limits on the ratio of successive orbital periods in the
system. For each such system of five satellites the parameters $\log T_0$ and $\log A$
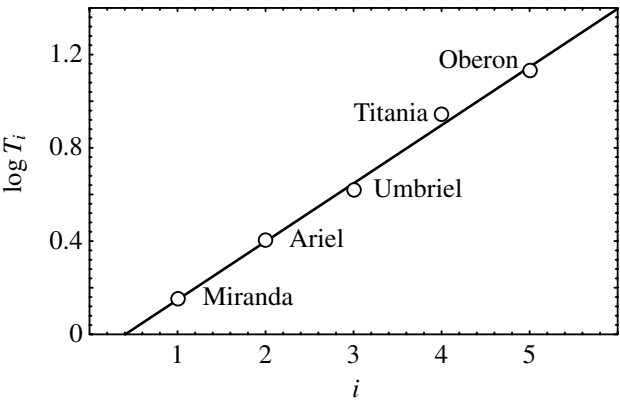and the rms deviation χ are determined.

Fig. 1.3. A linear, least squares fit for the orbital periods, T_i, of the five major uranian
satellites.

The choice of L and U is to some extent arbitrary. For the five uranian satellites the observed values are $L = 1.546$ and $U = 2.101$. By generating 10^5 sets of five satellites with periods given by the formula in Eq. (1.6), we can determine the number that have a value of $\chi < 0.0247$ for these values of L and U. Hence we estimate that the probability that the current configuration has arisen by chance is 0.79. Thus, despite the seemingly impressive fit displayed in Table 1.2 and Fig. 1.3, almost any distribution of periods, subject to the same constraints on L and U, would fit into a Titius–Bode "law" equally well.

We have investigated the sensitivity of this estimate to the values of L and U by repeating the above procedure for $L = 1.0$, 1.2, 1.4, and 1.6 using values of U in the range $1.2 < U < 5$. For every value of L and U we have calculated the value of χ for each of the set of 10^5 systems of five satellites. In each case we calculated the number of systems, N, for which $\chi < 0.0247$ and hence we estimated the probability $P(\chi < 0.0247) = 10^{-5}N$ for those values of L and U. The results are shown in Fig. 1.4. It is clear that $P \rightarrow 1$ as $U \rightarrow L$ when the ratios of successive periods are nearly equal. However, P only becomes small ($P < 0.01$) for large values of U, corresponding to widely spaced satellites.

These results suggest that the apparent regular spacing of the orbital periods shown in Table 1.2 is not significant. There is no compelling evidence that the uranian satellite system is obeying any relation similar to the Titius–Bode "law", beyond what would be expected by chance. This leads us to suggest that the "law" as applied to other systems, including the planets themselves, is also without significance. However, even though there are no grounds for belief in a Titius–
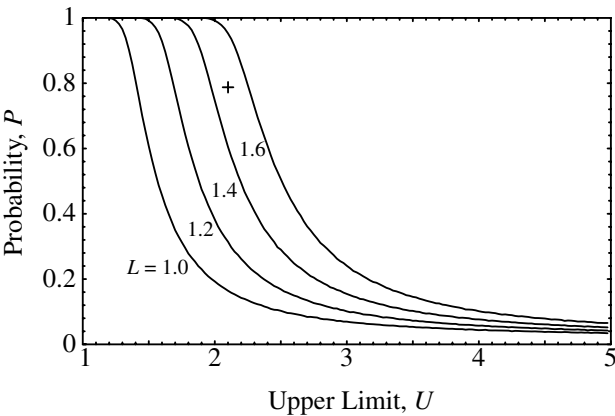

Fig. 1.4. The probability P that the calculated value of the rms deviation χ is less than the observed value of 0.0247 as a function of the upper limit U for values of the lower limit, $L = 1.0$, 1.2, 1.4, and 1.6. The cross marks the value appropriate for the actual uranian system.

Bode "law", the various bodies of the solar system do exhibit some remarkable numerical relationships and these can be shown to be dynamically significant.

1.6 Resonance in the Solar System

Our knowledge of the solar system has increased dramatically in recent years. Although no new planets have been discovered since 1930, there have been a number of advances in the study of the minor members of the solar system. By the start of the twentieth century 22 planetary satellites had been discovered and now there are known to be more than 60 satellites (see Appendix A) with indirect evidence for the existence of others. In addition there are currently more than 10,000 catalogued asteroid orbits and more than 500 reliable orbits for comets. Numerous bodies have been discovered with orbits beyond that of Pluto in the Edgeworth–Kuiper belt. Some estimates suggest that there may be as many as 2×10^8 objects with radii $\sim$10 km in this region (Cochran et al. 1995). Observations by the Infra-Red Astronomical Satellite (IRAS) have revealed the presence of dust bands in the asteroid belt and dust trails associated with comets. The study of planetary rings has also undergone radical changes; prior to 1977 it was believed that Saturn was the only ringed planet, whereas now we know that all the giant planets possess ring systems, each with unique characteristics.

The avalanche of planetary data in recent years has provided striking confirmation that our solar system *is* a highly structured assembly of orbiting bodies, but the structure is not as simple as Kepler's geometrical model nor as crude as that implied by the Titius–Bode "law". It is Newton's laws that are at work and the subtle gravitational effect that determines the dynamical structure of our solar system is the phenomenon of *resonance*.

In basic terms a resonance can arise when there is a simple numerical relationship between frequencies or periods. The periods involved could be the rotational and orbital periods of a single body, as in the case of *spin–orbit coupling*, or perhaps the orbital periods of two or more bodies, as in the case of *orbit–orbit coupling*. Other, more complicated resonant relationships are also possible. We now know that dissipative forces are driving evolutionary processes in the solar system and that these are connected with the origins of some of these resonances.

The most obvious example of a spin–orbit resonance is the Moon, which has an orbital period that is equal to its rotational period, resulting in the Moon keeping one face towards the Earth. Most of the major, natural satellites in the solar system are in a 1:1 or *synchronous* spin–orbit resonance. However, other spin–orbit states are also possible and radar observations of Mercury by Pettengill & Dyce (1965) showed that the planet Mercury is in a 3:2 spin–orbit resonance.

In the following sections we discuss each of the subsystems of the solar system.

1.6.1 The Planetary System

The orbital elements of Jupiter and Saturn are modified on a ~ 900 y timescale by a 5:2 near-resonance between their orbital periods, which French astronomers called *la grande inégalité* (the great inequality). Although the two planets are not actually in a 5:2 resonance they are sufficiently close to it for significant perturbations to be experienced by both bodies.

The planets Neptune and Pluto are in a peculiar 3:2 orbit–orbit resonance that maximises their separation at conjunction with the result that they avoid a close approach. The complexity of the Neptune–Pluto resonance was discovered and has been studied, not by a prolonged series of observations, but by direct numerical integration of the appropriate equations of motion. This is the only possible technique because observations of Pluto span less than a third of its orbital period.

As well as the resonances involving their orbital periods, some of the planets are also involved in long-term or *secular* resonances associated with the precession of the planetary orbits in space.

1.6.2 The Jupiter System

Perhaps the most striking example of orbit–orbit resonance occurs amongst three of the Galilean satellites of Jupiter (see Fig. 1.5). Io is in a 2:1 resonance with Europa, which is itself in a 2:1 resonance with Ganymede, resulting in all three satellites being involved in a configuration known as a *Laplace resonance*. The average orbital angular velocity or *mean motion n* in $^\circ\mathrm{d}^{-1}$ is defined by $n = 360/T$, where T is the orbital period of the body in days. The mean motions of Io, Europa, and Ganymede are

$$n_\mathrm{I} = 203.488992435^\circ\mathrm{d}^{-1}, \qquad (1.7)$$
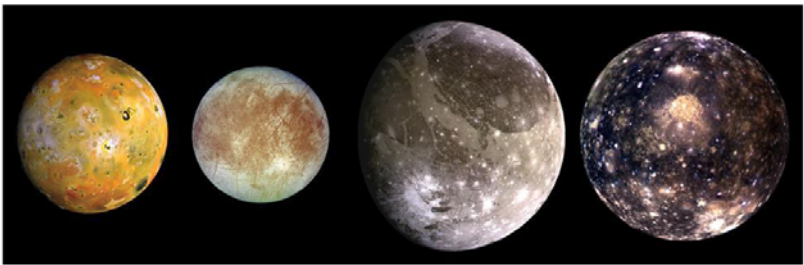

Fig. 1.5. A montage of images of the Galilean satellites of Jupiter shown to the same relative sizes. The satellites are (from left to right) Io, Europa, Ganymede, and Callisto. Ganymede has a mean radius of 2,634 km and is the largest moon in the solar system. The images were taken by the *Galileo* spacecraft. *(Image courtesy of NASA/JPL.)*