Planetary Systems: from Symmetry to Chaos

Caroline Terquem
Professor of Physics, University of Oxford

Abstract
This paper first discusses the definitions of disorder and chaos in physics. Then, Professor Terquem explores the dynamics of planetary systems. In particular, she examines how the seemingly chaotic process of planetary formation yields synchronised systems in their orbits. The transition is an excellent example of how a disordered physical system can generally tend to a state of orderedness.

‘Planet’ comes from the Greek word which means ‘wanderer’, as planets were seen to move relative to the stars. Although the orbits of the planets seemed chaotic viewed from the Earth, the Greek astronomers were convinced that they could be explained by superposing spheres and circles. According to Plato, these were the preferred figures of the natural world. From the geocentric model of Ptolemy to the heliocentric construction of Copernicus, the underlying assumption was that the planets were moving along superpositions of perfect circles. In his *De Revolutionibus*, Copernicus alludes to the symmetry of the universe which is revealed by the arrangement of these orbs. ‘Symmetry’, in that context, is taken to mean well-proportioned.

Modern observations of the solar system have confirmed that it is indeed highly structured. The eight planets are well ordered, with the terrestrial rocky planets closer to the Sun, and the more massive (gas or ice) giant planets further away. Each of the giant planets has a cortege of moons, which makes them look like a reduced-scale version of the solar system. These observations are well explained by the theory of planet formation which, in its modern form, originates in the 19th century from the idea that the Earth is made of meteoritic material. This theory was put on a quantitative basis in the second part of the 20th century, and has been very successful at explaining the over-all order of the planets in the solar system. Planets form in a so-called protoplanetary disc, made of gas and solid particles, which surrounds every newly formed star. The solid particles collide with each other and agglomerate to form bigger and bigger objects, until rocky planets form. Far enough from the Sun, where low temperatures enable ices to condensate, there is enough solid material to produce yet more massive objects, onto which large amount of gas can be captured by gravity, resulting in giant planets. It is likely that in the same way that planets have formed around the Sun, satellites (moons) have assembled around giant planets.

In the solar system, satellite systems display interesting dynamical effects which are not observed among the planets. Indeed, a significant number of satellite pairs are found to be in *mean motion resonances*, which means that the periods of revolution of the two satellites are commensurable: their ratio is that of two relatively small integer numbers. For example, Enceladus and Dione, two satellites of Saturn, are in a 2:1 mean motion resonance, such that Enceladus, which is closer to the planet, completes two revolutions while Dione completes one revolution. Such resonances may even be seen in systems of three satellites. A famous example, which was first studied by Laplace in the late 18th century, involves Jupiter's moons Io, Europa and Ganymede, which periods of revolution are in the ratio 4:2:1. Such resonances lead to repetitive configurations of the system of moons, and hence to an enhancement of their mutual gravitational interaction which builds up over time.

How can such a finely tuned and well-synchronised dance be achieved? The naive expectation would be that the process by which moons form leads to a random distribution of periods of revolution. Therefore, moons must subsequently move with respect to each other to establish resonances. This process is subtle, and not understood until the 1960's.

In a seminal paper published in 1965, it was first proposed by Peter Goldreich that moons can be brought into resonance by tidal interaction with the planet around which they revolve. For example, in the case of Enceladus and Dione, tides raised by Enceladus in Saturn result in Enceladus moving away from the planet, in exactly the same way as our Moon recedes from Earth as a result of the tides it raises on our planet. As Dione is further away, it is much less affected by the tidal forces, as these decrease dramatically with distance. Enceladus then migrates closer to Dione, so that the ratio of the orbital period of Enceladus to that of Dione increases. If this ratio is initially smaller than one-half, then it may reach this value at some point. This corresponds to a resonance being encountered. Once the resonance is reached, the enhanced mutual gravitational interaction between the two moons ensures that commensurability is preserved, even as the tidal forces with Saturn keep pushing Enceladus away. Both Enceladus and Dione then remain locked in resonance as they both migrate away from the planet. The same process may also explain how Io, Europa and Ganymede are in the so-called Laplace resonance described above: the tidal forces of Jupiter push Io away until it "captures" Europa in a resonance. The pair then continues to migrate away as a unit until Ganymede gets caught, at which point the three planets subsequently stay locked in this configuration. A whole system of satellites can be captured that way! This remarkable arrangement would certainly have qualified to be part of the symmetry of the Universe dear to Copernicus. However, as will be shown below, the very same process that can lead to the stable configuration described above can also result in disruption and chaos.

The idea that our own planetary system is not unique in the Universe dates back at least to antiquity. If Nature does not produce truly unique objects, there must indeed be planets around stars similar to our Sun. As stars and planets form together by a mechanism which repeats itself, planets were expected to be found around solar-type stars. The first discovery was not made until 1995… and it came as a surprise! This is because the planet, called 51 Pegasi b, is a giant similar to Jupiter, but located ten times closer to its host star than Mercury is to the Sun! It is so close to the star that its period of revolution is less than five days. Since then, many more such 'hot Jupiters' have been detected. On the basis of the theory of planet formation mentioned above, only smaller, rocky planets would be expected to be found so close to the star. This is not to say that the theory is wrong. Rather, processes which have not played an obvious role in the context of the solar system have been much more significant for at least some extrasolar planetary systems. These processes are responsible for the so-called *migration* of planets in the protoplanetary disc in which they form. Once planetary cores reach a mass large enough, their gravitational interaction with the gas in which they are embedded changes their orbit in such a way that they may move closer to the star.

This is a process which, in many ways, is similar to the tidal interaction between the Moon and the Earth. As the planetary cores accrete gas and become giant planets, the details of their gravitational interaction with the disc change but it may still lead to inwards migration. This process, which can explain why some planets are found very close to their host star, had previously been studied in the context of the solar system already in the late 1970s, as it also applies to moons embedded in the rings which surround the giant planets. It had even been predicted that some extrasolar planets could have migrated significantly that way! However, given that the planets in the solar system do not appear to have been subject to significant migration, this prediction had not attracted much scrutiny from observers.

It has been 25 years since the discovery of 51 Pegasi b. Thousands of planets have now been detected, with masses ranging from a fraction of an Earth mass to several Jupiter masses. In almost all cases, planets are not being observed directly. Instead, their presence is inferred from the effect they have on their host star: their gravitational pull results in a wobble of the star that can be detected, or their transit blocks some of the light we receive from the star. Numerous multiple systems, containing up to seven planets, have been observed. In some systems, the presence of a planet is inferred not from its effect on the star, but from how it perturbs the motion of another planet which has already been detected. This is very similar to the way that Neptune was discovered historically, from its measurable perturbations of Uranus's orbit. It is clear that a lot can be learnt from the disruption of the symmetry of an orbit!

Since planets migrate in their protoplanetary discs and often exist in multiple-planet systems, it is not surprising that there are many examples of mean motion resonances. The migration rate depends on many parameters, such as the mass of the planets and their distance to the star. Therefore, situations where the distance between two planets decreases with time are frequent, and this leads to captures into resonances, just as described above in the context of planetary moons in our solar system. A large number of resonances (or near-resonances) involving two planets have been reported, but *chains* of planets in resonances are also common, with up to seven planets involved, such as the Trappist-1 system. These systems are a beautiful illustration of how symmetry and order may emerge from a rather chaotic planet formation scenario.

But chaos may also emerge from order…

The systems we have mentioned above and which exhibit mean motion resonances are stable, which means that there can persist for billion of years without being disrupted by perturbations. However, in some cases, the gravitational interaction between two objects in a resonance can actually lead to the disruption of the system! Evidence of this process is found in the structure of the asteroid belt, which is populated by small rocky bodies and located between the orbits of Mars and Jupiter. These objects, whose sizes are at most a few percent of that of the Earth, are the leftovers of planet formation, which were prevented from growing further because of their gravitational interaction with Jupiter. The distribution of asteroids has voids at locations in the belt corresponding to commensurabilities with Jupiter: these so-called *Kirkwood gaps* (discovered by Kirkwood in the late 19th century) indicate that such commensurabilities are unstable, and that any asteroid which might have been present at that location has been ejected. By contrast, there is a clump of asteroids, called the Hilda family, at the location of the 3:2 resonance with Jupiter, located between the main asteroid belt and Jupiter's orbit, which indicates that this particular resonance is stable.

In general, a configuration is more likely to be stable when the asteroid has its closest approach with Jupiter (which happens when all three bodies are in conjunction, i.e. in a straight line) at the asteroid's *perihelion*. This configuration maximises the separation at closest approach to Jupiter. The interaction between the two objects is therefore rather small at conjunction, and does not build up to large values as the geometrical configuration is necessarily repeated (because of the resonant orbits). The asteroids of the Hilda family are actually in this configuration. It was originally suggested by Kirkwood himself that unstable configurations were the results of conjunctions arising when the asteroid was at *aphelion*, which minimises the separation at closest approach to Jupiter. In this case, the stronger interaction with Jupiter builds up over time and reaches large values. This leads to high eccentricities and/or inclinations causing the asteroids to collide with nearby objects. While this type of instability may indeed happen, more recent advances in celestial mechanics have shown that the trajectories of asteroids in some specific resonances (3:1 for example) are *inherently* chaotic. This is a fundamental property of the equations that govern the motion of the bodies, rather than been being caused by some specific configuration of the system. If both orbits were circular, there would be no resonant effect. If only the orbit of the asteroid were eccentric, the resonant gravitational force between the two objects would appear as a single dominant term in the equations, acting at a precise location for the asteroid. Such a configuration may be stable along the lines discussed above. However, even though the eccentricity of Jupiter's orbit is small, it does not vanish, and with both orbits being eccentric, there are in fact several dominant resonant terms in the gravitational force, each giving a strong contribution at their own different locations. Although those locations are distinct, they are close enough to each other in space that they all affect the motion of the asteroid. This is called *resonance overlap*. Recall that one single resonance leads to repeating configurations. The presence of additional resonances, however, produces *uncorrelated* forces which make the motion not repeated, but chaotic. Similar effects happen when distinct mean motion resonances overlap with each other, which is the case in the region between the orbits of the asteroids of the Hilda family and Jupiter's orbit. For example, the 4:3 resonance overlaps with the 5:4 resonance, which itself overlaps with the 6:5 resonance, etc. (the ratios being close to each other) and all resonances then contribute simultaneously to the motion of the object. This explains why there are no asteroids in this region of space. When resonances overlap, the motion of the asteroid is unpredictable, depending very sensitively on the initial conditions. This is called chaos.

The physics of mean motion resonances in the solar system and elsewhere is rich and fascinating, and illustrates how processes that can lead to such elegant configurations seen among the moons of giant planets or ensembles of extrasolar planets can also lead to chaos. While a single resonance acts as a Master Clock in the Universe, the superposition of two resonances yields disruption and unpredictability.[1]

---

[1] Reproduced from *The Language of Symmetry* (Eds. Rattigan, Noble & Hatta), 2023