Extra-solar planets

M.A.C.Perryman, Astrophysics Division, European Space Agency, ESTEC, Noordwijk 2200AG, The Netherlands; and Leiden Observatory, University of Leiden, The Netherlands

There are hundreds of billions of galaxies in the observable Universe, with each galaxy such as our own containing some 10^{11} stars. Surrounded by this seemingly limitless ocean of stars, mankind has long speculated about planetary systems other than our own, and the development of life elsewhere in the Universe. Only recently has evidence become available to begin to distinguish the extremes of thinking typified by opinions ranging from 'There are infinite worlds both like and unlike this world of ours' (Epicurus, 341–270 BC) to 'There cannot be more worlds than one' (Aristotle, 384–322 BC). Shining only by reflected starlight, extra-solar planets comparable to bodies in our own Solar System should be billions of times fainter than their host stars and, depending on their distances from us, separated from their accompanying star by, at most, a few seconds of arc. This combination makes direct detection extraordinarily demanding. Alternative high-precision detection methods, based on the dynamical perturbation of the star by the orbiting planet, on planetary transits, and on gravitational lensing, have therefore been developed.

Although planets around other 'normal' stars were expected to be common, the astronomical world was nevertheless taken by surprise when the first such discovery was announced by Mayor & Queloz in 1995. High-precision radial velocity (Doppler) measurements revealed a planet of about 0.5 Jupiter masses around a star similar to our own at a distance of about 15 pc. But, contrary to expectations, the planet was at a distance from its parent star of only about 0.05 AU, completing an orbit every 4 days. How such a 'giant' planet could have been formed in, or could have migrated to, such a location, sparked controversy, intensified efforts to detect other systems, and is stimulating much new theoretical and modelling effort in explaining their formation and evolution. Today (December 2000), about 50 extra-solar planets are known, and their numbers are increasing steadily.

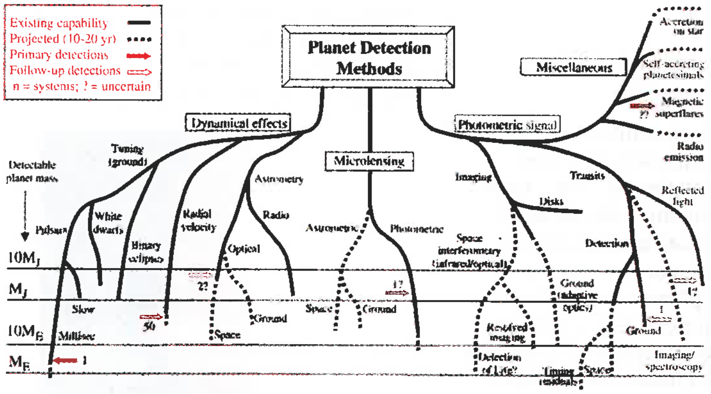

Fig. 1: **Detection methods for extra-solar planets. The lower extent of the lines indicates, roughly, the detectable masses that are in principle within reach of present measurements (solid lines), and those that might be expected within the next 10–20 years (dashed). The (logarithmic) mass scale is shown at left. The miscellaneous signatures to the upper right are less well quantified in mass terms. Solid arrows indicate (original) detections according to approximate mass, while open arrows indicate further measurements of previously-detected systems. '?' indicates uncertain or unconfirmed detections. The figure takes no account of the numbers of planets that may be detectable by each method.**

Stars, Brown Dwarfs and Planets

Stars form from gravitational instabilities in interstellar clouds of gas and dust grains, leading to collapse and fragmentation. In particular, a density perturbation (e.g. due to a shock wave) may cause the gravitational binding energy of a cloud to exceed its thermal energy. In that case it begins to contract, and a star is born. The resulting stars shine by thermonuclear fusion, with stable hydrogen burning occurring for masses above about $0{:}08\ M_\odot$ (about 80 M_J), when the central temperature triggers nucleosynthesis. Brown dwarfs are objects which occupy the mass range of about 12–80 M_J. They have also formed 1 by gravitational instability in a gas, but are not massive enough to ignite stable hydrogen burning. Below about 12 M_J, objects should derive no luminosity from thermonuclear fusion at any stage in their life. Such objects are broadly classified as planets if they formed through the agglomeration of residual protoplanetary disk material, rather than direct gravitational collapse and fragmentation.

A number of theories of the origin of our own Solar System have been advanced, starting with the ideas of Laplace more than 200 years ago. In the most widely considered 'solar nebula theory' planet formation in our Solar System (and, by inference, planetary formation in general) follows on from the process of star formation. Basically, the dust grains first settle into a dense layer in the mid-plane of the disk. They then begin to stick together as they collide and form macroscopic objects with sizes of order 0.01–10 m. These all orbit the protostar in the same direction and in the same plane, analogous to the few hundred metre thick rings around Saturn. Over the next 10^4–10^5 years, further collisions lead to the formation of 'planetesimals', objects up to a km or so in size, driven by graviational interactions. In the presence of an adequate mass of planetesimals, rapid run-away growth occurs as a result of dynamical friction and three-body effects. When the gravitational pull of the largest planetesimals is sufficient, they grow rapidly to the size of small planets by mutual collisions and mergers, with the terrestrial planets growing over time scales of 10^7–10^8 years. Mergers proceed by pairwise accretion until the spacing of planetary orbits becomes large enough that the configuration is stable for the lifetime of the system.

In oru Solar System, formation theories can be confronted with numerous observational constraints: orbital motions, stability, and spacings of the nine planets; planetary masses, rotations,

Article available at http://www.europhysicsnews.org or http://dx.doi.org/10.1051/epn:2001102

and the existence of planetary satellites and rings; angular momentum distribution; bulk and isotopic composition; radio-isotope ages; cratering records; and the occurrence of comets, asteroids and meteorites including the presence of the Oort Cloud and the Edgeworth-Kuiper Belt. The present paradigm of planet formation offers many attractive features, but various problems remain, and it seems wisest to view it as a plausible framework with which future observations are to be confronted.

Detection Methods

Figure 1 summarises the detection possibilities referred to in this section, and Figure 2 illustrates the respective parameter regions probed by some of these methods.

Imaging

Imaging of an extra-solar planet generally refers to the detection of an unresolved point source image of the object seen by the reflected light from the parent star (resolved imaging of an extra-solar planetary surface will require ground- or space-based interferometric arrays of 10–100 km baseline). The ratio of the planet to stellar brightness is given by:

$$\frac{L_p}{L_*} = p(\lambda, \alpha) \left(\frac{R_p}{a} \right)^2$$

where p(λ, α) is a wavelength and phase-dependent 'albedo', and α is the angle between the star and observer as seen from the planet. L_p/L_* is very small, of order 10^{-9} for a Jupiter-type object. Viewed with a ground-based telescope, with a star-planet separation of 1 arcsec (Jupiter viewed from 5 pc) the planet signal is immersed in the photon noise of the telescope's diffraction profile ($\lambda/D \simeq 0.02$ arcsec at 500 nm for a 5-m telescope) and more problematically within the 'seeing' profile (of order 1 arcsec) arising from turbulent atmospheric refraction. Under these conditions elementary signal-to-noise calculations imply that obtaining a direct image of the planet is not feasible.

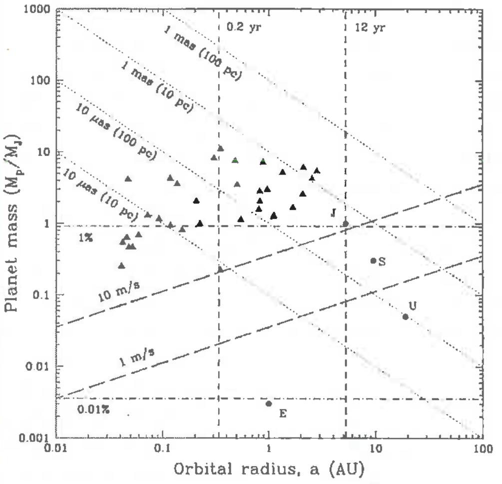

Fig. 2: Detection domains for methods exploiting planet orbital motion, as a function of planet mass and orbital radius, assuming $M_* = M_\odot$. Lines from top left to bottom right show the locus of astrometric signatures of 1 mas and 10 μarcsec at distances of 10 and 100 pc (equation 3). Very short and very long period planets cannot be detected by planned astrometric space missions: vertical lines show limits corresponding to orbital periods of 0.2 and 12 years. Lines from top right to bottom left show radial velocities corresponding to $K = 10$ and $K = 1$ m s^{-1} (equation 2). Horizontal lines indicate photometric detection thresholds for planetary transits, of 1% and 0.01%, corresponding roughly to Jupiter and Earth radius planets respectively (neglecting the effects of orbital inclination, which will diminish the probability of observing a transit as *a* increases). The positions of Earth (E), Jupiter (J), Saturn (S) and Uranus (U) are shown, as are the lower limits on the masses of first 34 known planetary systems (triangles).

Imaging efforts are directed at ways of reducing the angular size of the stellar image, suppressing scattered light, minimising the effects of atmospheric turbulence, and enhancing the contrast between the planet and the star by observing at longer wavelengths. Many ambitious ground- and space-based efforts related to extra-solar planetary imaging are on-going, but none of these imaging techniques has been successfully applied to extra-solar planetary detection so far. But they represent important long-term efforts since they provide the basis for attempts to measure the spectral features of planets, and therefore the possibility of detecting signatures of life in their atmospheres.

Dynamical Perturbation of the Star

The circular motion of a planet and star around their common barycentre causes the star to undergo a reflex motion, with orbital radius $a_* = a \cdot (M_p = M_*)$ and period *P*. This results in the periodic perturbation of three observables, all of which have been detected (albeit in different systems): in radial velocity, in angular (or astrometric) position, and in time of arrival of some periodic reference signal.

Radial velocity

The velocity amplitude K of a star of mass M_* due to a companion with mass $M_p \sin i$ with orbital period *P* and eccentricity *e* is:

$$K = \left(\frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_p + M_*)^{2/3}} \frac{1}{(1 - e^2)^{1/2}}$$

The effect is about K = 12.5 m s^{-1} with a period of 11.9 yr in the case of Jupiter orbiting the Sun, and about 0.1 m s^{-1} for the Earth. The sin *i* dependence means that the radial velocity measurements can de-

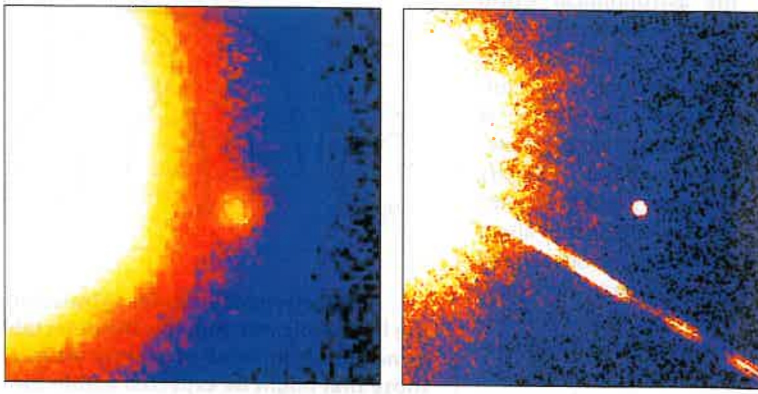

Fig. 3: Image of the brown dwarf Gliese 229 b, obtained with an adaptive optics system at the Palomar Observatory 60-inch telescope (left) and wit the Hubble Space Telescope (right). The brown dwarf is 7 arcsec to the lower right of the companion star, Gliese 229. The star/brown dwarf brightness ratio is ~5000, and the distance between the two objects corresponds roughly to the Sun-Pluto separation. A Jupiter mass planet a a distance of 10 pc would be 14 times closer to its parent star, and roughl 200 000 times dimmer than Gliese 229 b (courtesy of Tadashi Nakajima).

termine only $M_p \sin i$ rather than M_p, and hence provide only a lower limit to the mass of a specific planet.

All of the known extra-solar planets around ordinary stars have been discovered, starting with the first in 1995, using radial velocity techniques. A very significant effort is now directed at radial velocity detection by many groups and at many telescopes around the world, and some 50 extra-solar planets are presently known. Current measurements reach accuracies of around 3 m s^{-1}. Because the stars are only faint sources of light, large telescopes, long integration times, and very precise instrumental calibrations are required.

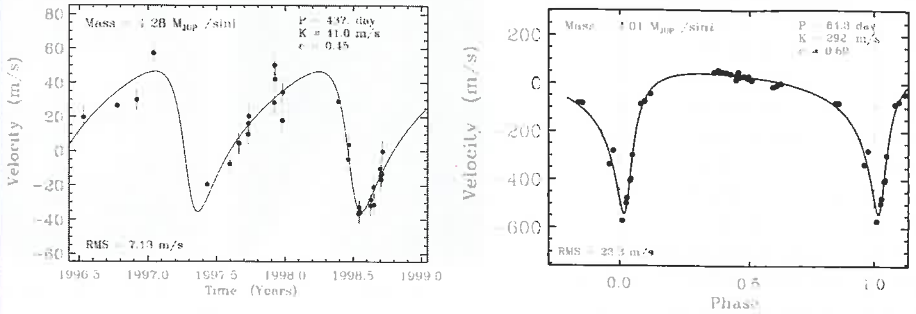

Fig. 4: **Examples of radial velocity measurements: HD 210277 (left) and HD 168443 (right), obtained with the HIRES spectrometer on the Keck telescope. The solid lines show the best-fit Keplerian models. The non-sinusoidal variations result from the eccentric orbits, and the derived $M \sin i$ values are 1.28 and 4.01 M_J respectively. The fit for HD 168443 is improved further by a linear velocity trend, suggestive of an additional nearby, long-period stellar or brown dwarf companion (courtesy of Geoffrey Marcy).**

Astrometric position

The path of a star orbiting th star-planet barcentre appears projected on the plane of the sky as an ellipse with angular semi-major axis α given by:

$$\alpha = \frac{M_p}{M_*} \cdot \frac{a}{d}$$

where α is in arcsec when a is in AU and d is in pc. Astrometric techniques aim to measure this transverse component of the photocentric displacement. Jupiter orbiting the Sun viewed from a distance of 5 pc would result in an amplitude of 1 milliarcsec, while the effect of the Earth at 5 pc is a one-year period with 0.6 μarcsec amplitude.

Measurement of sub-milliarcsec displacements in the optical is impossible so far because of atmospheric effects, although measurements at the tens of microarcsec or better are projected using interferometric techniques planned for the world's largest ground-based telescopes, such as the Keck Interferometer, and from the European Southern Observatory's VLTI.

Astrometric measurements can be made more accurately from above the Earth's atmosphere, from where ESA's Hipparcos satellite provided ~ 1 milliarcsec accuracy for about ~120 000 stars. Future ambitious space experiments at the microarcsec-arcsec level include NASA's SIM space interferometer mission (launch 2009), and GAIA, ESA's stereoscopic mission to map the 3-d positions and space motions of more than a billion stars in our Galaxy, to be launched around 2010. In the process, GAIA should detect many thousands of extra-solar planets out to 100--200 pc, due to the wobble of their photocentre caused by orbiting planets.

Timing and the pulsar planets

Although all orbital systems are affected by changes in light travel time across the orbit, in general there is no timing reference on which to base such measurements. A notable exception are radio pulsars, rapidly spinning highly-magnetised neutron stars. Accurate timing of the pulse arrival can in principle reveal objects as small as an Earth mass in orbit around them. This technique led to the first planetary system detected around the 6.2-ms pulsar PSR 1257+12, with at least two plausible companions having masses of 2.8 and 3.4 Earth masses, and almost circular orbits with $P = 98.22$ and 66.54 days respectively. A couple of other much higher mass companions around other pulsars have been discovered in the last few years, but planetary systems around these old neutron stars seem to rare.

Periodic Photometry: Transits and Reflections

Detection of extra-solar planets, by observing their eclipse of the host star, was first considered almost 50 years ago. The method is conceptually simple: star light is attenuated by the transit of the orbiting planet across its disk, with the effect repeating at the orbital period of the planet. The method is presently considered as one of the most promising means of detecting planets with masses significantly below that of Jupiter, with the detection of Earth-class (and hence habitable) planets within its capabilities. Extrapolation of the method down to masses of planetary satellite may even be feasible.

Detection probabilities depend on the transit geometry and on the luminosity drop produced by an object on the line of sight to the star, which approximates to:

$$\frac{\Delta L}{L_*} \simeq \left(\frac{R_p}{R_*}\right)^2$$

under the assumption of a uniform surface brightness of the star. Values of $\Delta L/L_*$ for the Earth and Jupiter transiting the Sun are 8.4×10^{-5} and 1.1×10^{-2} respectively. The duration of the transit is about 25 hr for a Jupiter-type planet and 13 hr for an

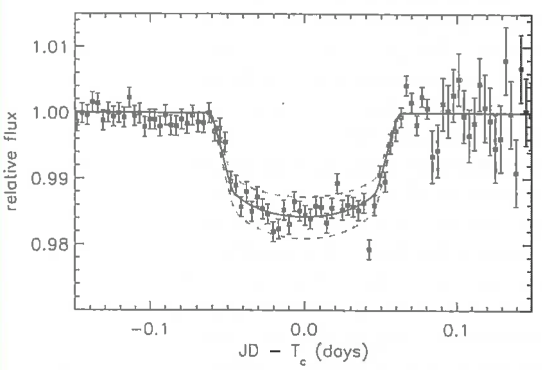

Fig. 5: **The first detected transit of an extra-solar planet, HD 209458. The figure shows the measured relative intensity versus time. Measurement noise increases to the right due to increasing atmospheric air mass. From the detailed shape of the transit, some of the physical characteristics of the planet can be inferred (courtesy of David Charbonneau).**

Earth-type system. If the radius of the star is known, then R_p is also determined. The greatest disadvantage of the method is that it requires configurations in which the viewing direction (to the Earth) happens to lie in the orbital plane of the planet.

Ground-based photometry beyond 0.1% accuracy is complicated by variable atmospheric extinction and scintillation, and extension of the transit method to space experiments, where very long uninterrupted observations can be made above the Earth's atmosphere, therefore holds particular promise. Small 'precursor' space experiments are underway, while an ESA mission, Eddington, devoted to both planetary detection and asteroseismology studies, has recently been approved.

An important result in extra-solar planetary studies has been the detection of the first transit event, for the system HD 209458 (Figure 5). The precise shape of the light curve yields estimates of the mass and radius of the orbiting planet. The estimated density of $\rho \sim 380$ kg m^{-3} is consistent with a hydrogen gas giant, and yields a surface gravity of $g \sim 9.7$ m s^{-2}. Other experiments to survey large numbers of stars from ground, from the Hubble Space Telescope, and to detect reflected light from orbiting planets, are also in progress.

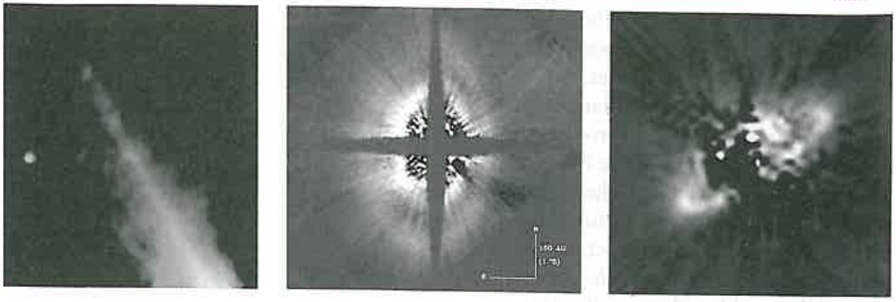

Fig. 6: Images of disks from the NASA/ESA Hubble Space Telescope. Left: part of the β Pic disk imaged by the WFPC2 instrument. Clumps of dust might represent elliptical rings viewed edge-on, possibly created by the gravitational force of a stellar encounter ~ 10^5 years ago (courtesy of Paul Kalas). Centre: the circumstellar disk of HD 141569 imaged in reflected light at 1.1 μm with the NICMOS instrument. The central star and the dark vertical and horizontal bands are regions obscured by a coronographic mask (courtesy of Alycia Weinberger). Outwards from the centre, a bright inner region is separated from a fainter outer region by a dark band, superficially resembling the Cassini division (the largest gap) in Saturn's rings. The disk extends to about 400 AU, about 13 times the diameter of Neptune's orbit, with the gap at 250 AU. An unseen planet may have carved out the gap, in which case its mass can be estimated at ~ 1.3 M_J, and its orbital period as 2600 years. If it takes ~ 300 orbital periods to clear such material then the gap could be opened in ~ 8×10^5 years, consistent with the age of the star. Right: the circumstellar disk around HR 4796A, also observed by coronographic imaging by NICMOS. The ring is almost completely visible, although the clumpy structure arises from the coronographic system and is not real (courtesy of Glenn Schneider, Brad Smith, and the NICMOS IDT/EONS team). The colour of the material (from infrared measurements at 1.1 and 1.6 μm) implies particles with a size of a few μm, larger than typical interstellar grains. The implied confinement of material also argues for dynamical constraints on the particles by one or more as yet unseen bodies.

Gravitational Microlensing

Gravitational lensing is the focusing and amplification of light rays from a distant source by an intervening object, first considered by Einstein in 1936. Precise alignment between the observer, some intervening object, and the more distant planetary system, can lead to significant brightening of the background object due to this lensing effect, with a duration depending on the relative velocity of the lens and source. The phenomenon offers a powerful but experimentally challenging route to the detection and characterisation of planetary systems. However, the chance of substantial microlensing magnification is extremely small, being ~ 10^{-6} for background stars in the Galactic bulge. Since 1993 several hundred photometric microlensing events have now been observed. Events with durations ranging from hours to months are detected and reported while still in progress, allowing concerted follow-up observations. Detailed light-curve structure probes the lens kinematics, the frequency and nature of binary systems, stellar atmospheres, and the presence of planetary systems. Unconfirmed planetary candidates have been identified in these data.

Miscellaneous Signatures and Planetary Disks

Other possible signatures of planetary presence include searching for evidence of giant impacts during the late stages of planetary formation, 'super-flares' caused by magnetic reconnection between fields of the primary star and a close-in Jovian planet, coherent cyclotron radio emission driven by the stellar wind/magnetospheric interaction, or signatures of a planet's final accretion onto the central star.

It is now relatively easy to observe the precursors of planetary systems, protoplanetary disks. Not only are they extremely large – a typical disk extends to of order 1000 AU from the star – but the surface area of the small particles which make up the disk is many orders of magnitude larger than that of a planet. Disks appear to be long-lived, ~ 10^6 – 3×10^7 years, ar ecommon throughout our Galaxy, and appear quite similar to the picture of our primitive solar nebular. A number of disk systems have been images using the Hubble Space Telescope (Figure 6).

Properties and Formation Theories

The extra-solar planets discovered so far are all more massive than Saturn, and most either orbit very close to their stars or travel on much more eccentric paths than any of the major planets in our Solar System. While Doppler surveys are most sensitive to small orbits, what had not been generally anticipated was the small orbital radii and large eccentricities of many of the first systems discovered. This trend has continued: 14 of the 50 known planets reside in orbits with $a < 0.1$ AU, and 24 have $a < 0.3$ AU. Close-in planets are generally in rather circular orbits, while the 26 planets beyond 0.3 AU are all in non-circular orbits with $e \geq 0.1$, with 22 having $e \geq 0.2$.

A robust formation mechanism must explain the mass distribution, the large eccentricities, and the large number of orbits with $a < 0.2$ AU, where both high temperature and relatively small amount of protostellar matter available for their agglomeration would inhibit formation *in situ*. The existence of the close-in planets has focussed attention on some earlier predictions that Jupiter-mass gas giants could be formed further from the star, followed by non-destructive migration inwards, driven by tidal interactions with the protoplanetary disk. Increasingly realistic simulations of evolving planetary systems have been made possible through improved physical models and developments

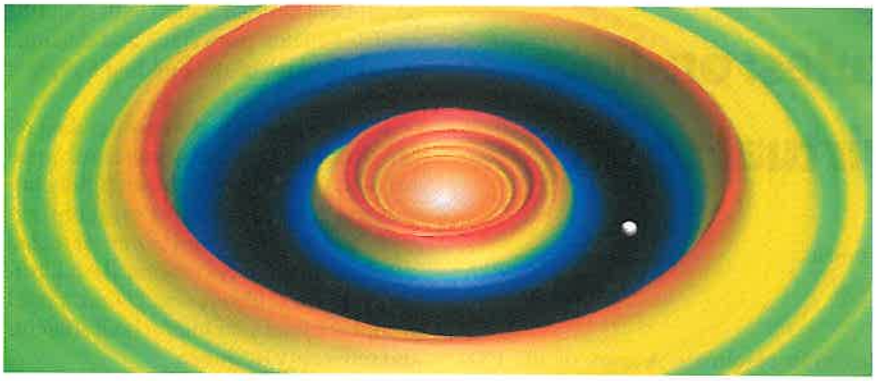

Fig. 7: Simulation of the formation of a planetary system. The surface density of disk material has been reduced by four orders of magnitude near to the planet of mass $M_p/M_* = 10^{-3}$, and waves are clearly seen propagating both inward and outward away from it (courtesy of Douglas Lin).

in computer speed and computational algorithms. Figure 7 shows results of simulations of disk-planet interactions and tidal interactions with the central star.

Apart from the pulsar example, only two multiple extra-solar planetary systems are presently known. Their relative orbital inclinations are still unknown. This information is required for stability analyses, and for further investigation of the formation theories. As the temporal baseline of the radial velocity measurements increases, it is likely that many other multiple systems will be discovered, and it is not yet known what fraction of planetary systems will turn out to be multiple.

Predicted spectral properties of extra-solar planets were made in advance of their discovery, giving brightness versus mass and age, and including sensitivity to parameters such as deuterium and helium abundance, rotation rate, and presence of a rock-ice core. Models have been updated subsequently, including effects of the outer radiative zones caused by the strong external heating which inhibits atmospheric convection. Hydrostatic evolution calculations and atmospheric models, including radiative effects, are being used to determine structures, radii, equilibrium temperatures, luminosities, colours, and spectra of objects with temperatures from 1300 K down to 100 K. These will help to classify and characterise the planets as more information about them becomes available.

1 AU = 1 astronomical unit (mean Sun-Earth distance) $\simeq 1.5 \times 10^{11}$ m. Solar mass = 1.99×10^{30} kg. Jupiter mass = 1.90×10^{27} kg. Earth mass = 5.98×10^{24} kg. For Jupiter, $a = 5.2$ AU and $P = 11.9$ yr. Stellar distances are conveniently given in parsec (pc), defined as the distance at which 1 AU subtends an angle of 1 second of arc (or arcsec); 1 pc $\simeq 3.1 \times 10^{16}$ m $\simeq 3.26$ light-years. For reference, distances to the nearest stars are of order 1 pc; there are about 2000 known stars within a radius of 25 pc of our Sun, and the distance to the Galactic centre is about 8.5 kpc.
Parameters used are mass M, radius R, and luminosity L, with subscripts $*$ and p referring to star and planet respectively. Systems are characterised by their orbital period P, semi-major axis a, eccentricity e, orbital inclination with respect to the plane of the sky i ($i = 0°$ face-on, $i = 90°$ edge-on), and distance from the Solar System d.

The Future

The search for other planets is motivated by efforts to understand their formation mechanism, and to gain an improved understanding of the origins of our own Solar System. Search accuracies will progressively improve to the point that the detection of telluric planets in the 'habitable zone' will become feasible. Improvements in spectroscopic measurements, whether from Earth or space, and in atmospheric modelling, will lead to searches for planets which are progressively habitable, inhabited by micro-organisms, and ultimately by intelligent life. Search strategies will be assisted by improved understanding of the conditions required for development of life on Earth. This will be a cross-disciplinary effort, with the participation of astronomers, chemists and biologists. The last few years has seen a number of exo-biology initiatives, and numerous conferences on the search for life, beginning finally to quantify a philosophical debate that has been ongoing for centuries.

Based on present knowledge, about 5% of solar-type stars may harbour massive planets, and an even higher percentage may have planets of lower mass or with larger orbital radii. If these numbers can be extrapolated, the number of planets in our Galaxy alone would be of order 1 billion. Global astrometric and photometric measurements from space should lead to the detection and characterisation of many thousands of significantly lower mass planets by the year 2020. Together, such data sets will provide a vast statistical description of planetary masses, orbits, and eccentricities, allowing important constraints to be placed on the complex processes believed to be involved in planetary formation. Improved knowledge of individual systems, in particular from transit measurements, combined with improved atmospheric modelling and theories of habitability, will narrow down the range of identified planets on which life may have developed. Nearby, Earth-mass planets should exist, and would be the natural targets for infrared space interferometers which, by 2020, may succeed in imaging and providing evidence for life on them.

We now know that other worlds – large ones at least – are common. Developments have been so rapid over the last few years that many significant developments, and many new surprises, can be predicted with confidence.

This article is based on the author's review article 'Extra-Solar Planets', which appeared in 'Reports on Progress in Physics', August 2000, Vol. 63(8), pp1209-1272, by kind permission of the Institute of Physics, to which the reader is referred for further details and references.

Bibliographical Details

Michael Perryman holds a degree in theoretical physics and a PhD in radio astronomy from Cambridge University. He has been employed in the Space Science Department of the European Space Agency since 1980, and is professor of astrophysics at Leiden University. For his involvement in the Hipparcos space astrometry mission, he was awarded the Prix Janssen of the Société Astronomique de France in 1996, and the Academic Medal of the Royal Netherlands Academy of Arts and Sciences in 1999.