euresis journal

euresis Association for the Promotion of the Scientific Endeavour

Volume 2
Winter 2012

journal homepage
www.euresisjournal.org

Extrasolar Planets, Extraterrestrial Life, and Why it Matters

Dan Maoz

School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel
maoz@astro.tau.ac.il

Abstract

The search for extrasolar planets, and the plans to eventually search for signs of life on those planets, are among the most exciting and fast-developing fields in science. I briefly review some of the historical development of this field, from its early non-scientific beginnings, through the amazingly correct ideas of Giordano Bruno, and up to the modern flood of recent data. I present, in non-technical terms, the four methods being used to detect and study extrasolar planets -- radial velocities, transits, direct imaging, and gravitational microlensing -- and the intriguing results they have produced. I then discuss techniques and prospects for astrobiology, and some recent developments in Earth-bound biology that guide these ideas. I conclude by arguing that, whether life is found or not found on other planets, either result would have profound implications for understanding life on Earth and its emergence, again in the spirit of Bruno.

1. Introduction

In the San Marino 2008 Symposium, concerned with the questions of creativity in science, of how it arises, and of how it is fostered, the organizers have decided (wisely, I believe) to dedicate some of the discussion to specific reviews of a few of the most active and groundbreaking areas of science, where creativity is certainly essential. There is no doubt in my mind that the search for, and discovery of, extrasolar planets (i.e., planets outside our own solar system), and the plans to eventually search for signs of life on those planets, do fall into this category. In this article I will briefly review some of the historical development of the subject, starting from its non-scientific beginnings, up to the latest results of the last few years. I will try to explain in non-technical terms the various simple but ingenious ("creative" is certainly appropriate) methods that have been devised to detect extrasolar planets and to study them, and the intriguing results they have produced. I will then continue to the next logical step, but one that is still at the pre-discovery stage: astrobiology - the search for, and study of, signs of life outside Earth. Again, I will discuss some of the techniques being considered and planned, and some recent developments in Earth-bound biology that guide our ideas of where and how we should look for evidence of extraterrestrial life, and what forms it may assume. Finally, I will argue that a major motivation for the whole pursuit, other than sheer curiosity as to the question of "are we alone?", is the potential for a much deeper understanding of the emergence of life here on Earth. Contrary to the popular perception of extraterrestrial life, as envisioned in countless (often highly entertaining) Hollywood films, and in the delusions of UFO aficionados, the scientific reality, as so often happens, will likely turn out to be much stranger and wondrous than imaginable by any screenwriter or crackpot. Finding extraterrestrial life, even though it will almost surely be of a most primitive form, will shed much light on what life is. No less important, a general absence of life on planets that could potentially harbor life would also have profound implications for us and our place in the cosmos.

2. Early History

Human speculation about the existence of other worlds and other sentient beings must be as old as humanity itself. For example, in the 6th century B.C., the Greek philosopher Anaximander discussed the possibility of a "plurality of worlds". The Assyrian satirist Lucian of Samosota composed in the 2nd century the humorous fantasy "A True Story", which is considered the first work of science fiction, replete with lifeforms and warring civilizations on the moon, the sun, and the planets. The concept naturally finds expression also in most, if not all, religions. To bring an example close to home, in the Babylonian Talmud, Volume "Avoda Zara", Tractate 3b, there is a discussion of what is God's daily routine, followed by the question of what does God then do at night. One possibility, based on interpretation of a verse in Psalms, is that ``he rides on a light angel of his and sails through eighteen thousand worlds".

As one of the very few places in Jewish scripture suggesting the existence of other worlds, this sentence in the Talmud has elicited further debate in the religious Jewish literature over the centuries. The issue is particularly critical in the monotheistic religions, where Earth and Man hold a special position in the Creation. Indeed, in medieval Christian discussions (e.g., in the condemnations of Aristotelian doctrines by Bishop Stephanus Tempier of Orleans in 1270), the concept of human centrality has been used to argue against the existence of other worlds.

The earliest well-known semi-scientific exploration of the subject was by the 16th century Italian philosopher Giordano Bruno, an early champion of the Copernican world view. In his dialogue "De L'Infinito Universo et Mondi" (On the Infinite Universe and Worlds, 1584) and other works, he conjectured that:

- Celestial bodies are composed of the same elements as the Earth (assumed to be earth, water, wind, and fire), rather than of a fifth and purer ``quintessence".
- The stars are immensely distant suns, each orbited by their own planetary systems.
- There is an infinity of other stars and planets, all inhabited.

The fact that these conjectures, made 25 years before Galileo's first use of the telescope, were not based on any scientific evidence, but rather on a pure but amazingly accurate intuition, make them all the more striking. It would take until the early twentieth century, with the development of spectroscopy and atomic physics, to confirm that, indeed, the stars are made of the same elements found on Earth, and until only the last few decades to extend this result to the furthest reaches of the Universe. The understanding that the Sun is a normal star has again developed over the past century, with the availability of increasingly accurate astronomical observations and the development of nuclear physics. The infinity of the Universe, with its implied infinity of stars (or at least their overwhelmingly large numbers) have emerged only over the past decade, with the advent of precision cosmology. Finally, as I will describe in detail below, it is only two decades ago that the first few extrasolar planets were discovered, and only within the last few years that it has become clear that planets are common around other stars. The sole Bruno conjecture that remains unconfirmed is the one about the ubiquity of life. Based only on his success so far, it could be argued that surely he must have gotten that one right as well. As I will explain below, we should know before too long.

With Galileo and the beginning of modern astronomy in 1609, it became clear that the moons and the other planets in the solar system, at least, do qualify as "other worlds", and the search for signs of life on them was on. The field got a boost in 1877 when Italian astronomer Schiaparelli, sketching the surface of Mars based on his visual observations, believed he saw long straight features which he termed "canali". This prompted American astronomer Percival Lowell to build and use an observatory dedicated to Mars observations. Based on his studies, he promoted between 1895 and 1908 the idea that Mars is covered by a network of canals built by an advanced but desperate civilization, in order to channel water from the poles to the arid equatorial regions. While professionals from the start viewed these ideas skeptically, and the spacecraft missions of the 1960s finally demonstrated the canals to be optical illusions, these ideas ignited the public imagination with respect to the subject of extraterrestrial life. Most notable in this sense was H.G. Wells's novel "The War of the Worlds" (1898), about an attempt by the desperate Martians to conquer Earth for the sake of our natural resources (in line with the colonialist thinking of the period), only to be vanquished by local bacteria (in line with the then-recent proof of the germ theory of disease).

The planetary probes that have explored the surface of Mars over the past few decades have found that Mars is probably inhospitable to life. However, Mumma et al. [1] have recently reported telescopic observations in 2003 showing a transient release of methane into the Northern summer hemisphere of Mars from plumes, which could, in principle (but not necessarily) be of subterranean biological origin. Regardless, the possibility remains that Mars hosted life in the past, when it possessed a denser atmosphere and liquid water on its surface.

Other sites in the solar system, such as the moons of Jupiter and Saturn, have not been ruled out yet as sites for some form of life. But, if we are to find more definitive evidence for extraterrestrial life, or the lack thereof, we need to expand the search beyond the confines of the solar system. Particularly relevant are other stars similar to the sun, and having "terrestrial" planets - rocky planets that are similar to Earth in terms of their mass (and hence gravity), temperature (which are set by their distances to their parent stars), and hence permitting the existence of liquid water. Remarkably, until 1992 there was not a single known example of an extrasolar planet.

As we will see, at the large distances of even the nearest stars, detecting a planet is extremely challenging. Following an initiative by Tel-Aviv University astronomer Tsevi Mazeh, his group [2] announced the discovery of a 10-Jupiter-mass companion in a close orbit around another star. Although at the time it was not widely recognized as a planet, many such planets are known today, and in retrospect this was likely the first extrasolar planet found. The first unambiguous discovery of extrasolar planets was made in 1992 by Wolszczan and Frail, who found, by means of radio timing observations, two planets around a pulsar - the extremely dense remnant of an exploded massive star. The first detection of a Jupiter-mass planet around a normal star came only three years later, by Mayor and Queloz [3]. This opened the floodgates on extrasolar planet discoveries, with over 700 found to date, and the number rapidly growing. Let us review the various techniques that have made it possible, after more than 400 years, to confirm Bruno's planetary conjecture.

3. How to find extrasolar planets

3.1. The radial-velocity method

The large majority of known extrasolar planets have been discovered by measuring the "wobble" that they induce on their parent stars. A basic result of Newtonian mechanics is that two masses under the influence of their mutual gravitational attraction will move in elliptical orbits around an imaginary point between them called the center of mass. The ratio of the distances of the two masses to this point equals the inverse of the ratio of the masses, and thus the center of mass is always closer to the more massive object. In the context of stars and their planets, consider, for example, the Sun and the planet Jupiter. With Jupiter's mass being 1/1000 that of the Sun, and the Sun-Jupiter separation being about 5 "astronomical units" (i.e., 5 times the Earth-Sun separation, or 5 x 150,000,000 km), the center of mass is at distance from the center of the sun of 5 x 150,000,000km / 1000 = 7

50,000 km. This happens to be just over the solar radius. Thus, while Jupiter goes through a full, more-or-less circular, 12-year orbit, the sun moves is a corresponding little circle around a point just outside its limb. It is then easy to calculate the velocity at which the sun does this little dance. The circumference of the circle the sun traces is 2π x 750,000 km, or about 5 million km, so the velocity is (5 million km) / 12 years, or 400,000 km/yr. There are 365 x 24 = 8760 hours in a year, so this is equivalent to about 50 km/hr, a typical driving speed. Now, each of the planets in the solar system makes the sun go through its own dance around a point that is always between the sun and the planet, and the actual motion of the sun will be the combination of all those motions. But in practice, Jupiter, because of its large mass and relatively short distance to the sun, is by far the dominant body behind the sun's driving-speed wobbles.

Astronomers are very adept at measuring velocities of celestial bodies using the Doppler effect that velocity induces on the light waves emitted by those bodies. Just as the pitch of the sound from an approaching train whistle is higher than that from a receding one, light from an approaching star gets shifted to blue wavelengths, and from a receding star to red wavelengths. The relative shift in wavelength equals just the ratio of the star's velocity to the velocity of light. The 50 km/hr velocity of the sun thus corresponds to a 50 parts-per-billion Doppler shift in the wavelength of the emitted light. While this sounds challenging, police radar guns that are used to catch speeding vehicles, and which operate on the same principle, reach these and better accuracies. The idea is then simple: monitor over time the velocity of a star, as deduced from the Doppler shifts of its emitted light. If that star has, e.g., a planet orbiting it just like Jupiter orbits the sun, and the orbital plane happens to be inclined "edge-on" to our line of sight, then, over 12-year periods, we will detect its wobble in the form of a periodic variation in the observed velocity. Half the time the star will be approaching us, reaching a maximum of 50 km/s in its observed velocity when its planet is abreast to one side, and 6 years later reaching this velocity in the opposite, receding, direction. When the planet passes exactly before or behind the star, the star is also at the point in its little orbit where it is moving perpendicular to our line of sight, i.e., neither approaching or receding, and hence its velocity is zero. So, if we monitor the Doppler line-of-sight velocity of a nearby star and see this kind of periodic wobbling, we can deduce the presence of a planet around it. From the period (12 years in the above example) and the amplitude of the variation (50 km/hr in the above example) we can deduce the orbital separation and the mass (5 astronomical units and 1 Jupiter mass in the above example). In reality, this is true only if we assume the edge-on inclination; if, as is often the case, we do not know the inclination of the unseen planet, we can only find a lower limit to the planet mass. If the planet is more massive and/or it is in a closer orbit around it's star, its stronger gravitational tug will cause a stronger and faster wobble, and hence the period will shorten and the amplitude will rise. Such planets will therefore be easier to detect (larger Doppler effect), requiring a shorter monitoring period.

In the early 1990s, astronomers refined the stability and accuracy of optical spectrographs on telescopes so that precisions of order 50 km/hr could be obtained when monitoring the light from the nearest (and hence brightest) stars. Planet discoveries around some of these stars soon began to flow in, with two main groups of researchers contributing, the Geneva group led by Mayor, and the California group led by Marcy and Butler.

The first planets discovered were, naturally, those that are easiest to find – massive objects like Jupiter. However, the big surprise was that these planets were orbiting their stars at tiny separations, smaller than Mercury's orbit in the solar system, and hence with orbital periods of only a few days. At these small separations, the temperatures of these planets due to the irradiation by their host stars must be quite high, and hence they have been dubbed "hot Jupiters". Their discovery was completely unexpected based on the only planetary system known previously, the solar system, where giant planets exist only in the outer regions - Jupiter and beyond. It was also unexpected theoretically. It was, and still is, thought that giant planets can only form at large distances, beyond the "snow line" where water can exist as a solid (more on this later). Although debate about the nature of hot Jupiters continues, it is generally believed that these planets indeed initially form far from their stars, but then "migrate" to their present close orbits. As the radial-velocity surveys continued and accumulated data, they were able to discover also planets of somewhat lower masses (of order Neptune), and on longer orbits, approaching that of Jupiter in the solar system. However, the very nature of the technique is biased toward finding hot Jupiters, which therefore constitute the large majority of the extrasolar planets discovered so far in this way (over 600 planets). In no way does this imply that such planets are typical. To find other types of planets, which are more similar to ours, and in particular planets that could sustain life, we must turn to additional techniques.

3.2. The transit method

The orbital planes of extrasolar planetary systems are inclined at random angles to our line of sight. Some fraction of them will be seen nearly edge on. A planet in such a system will transit the face of its parent star once per orbit. This "mini-eclipse" will cause a small reduction in the amount of light arriving from the star, in proportion to the ratio of the areas of the disks of the planet and the star. Jupiter, for example, has 1/10 the radius of the sun. Transiting across the face of the sun as viewed from outside the solar system, it would cause an approximately 1/100 shadowing of the sun's output during the transit. Obviously, detecting this requires high photometric (i.e., light measuring) accuracies, of better than 1%, in order to discover Jupiters, and even higher in order to discover smaller planets. As in the radial-velocity method, this requirement limits the search to nearby (and hence bright) stars, although the demands are not as stringent as in the radial-velocity case. Again, large planets on close orbits are the most likely to be found: the larger the planet, the larger its "silhouette"; the smaller the orbit, the greater the range of inclinations around exactly edge-on that will yield a transit; and the smaller the orbit, the shorter the period, and hence the less time required to observe many transits and thus to obtain a significant detection. The first extrasolar planet transit was detected in 2000, (independently by Charbonneau et al. and by Henry et al.) by monitoring the light from a sample of stars that were already known to have orbiting hot Jupiters (but with unknown orbital plane inclinations). The transit occurred exactly when expected based on the radial velocity data, i.e., when the radial velocity is zero, and between the phase when the planet is approaching us and when it is receding from us. Another 50 or so transit-based planets were found in the following decade.

Transit-detected planets have a rich variety of possibilities for interesting follow-up studies. First, the fact that they transit means that their orbital inclination to our line of sight is basically determined (edge-on), and therefore their masses are known accurately. The depth of the eclipses reveal their radii, and hence their mean densities can be calculated. From the densities one can learn about their internal compositions. Perhaps more dramatically, during the transit, light from the star will be partially absorbed by the semi-transparent atmosphere of the planet. By comparing the spectrum of the system in and out of transit one can then find spectral signatures of atoms and molecules in the planet's atmosphere, from which one can learn about its chemical composition, temperature, and more. Similarly, one can compare the spectrum of the system when the planet is out of transit to when it is hidden behind the star, and thus isolate the reflected light of the planet. Again, spectral analysis can then reveal a wealth of detail about the planet surface and atmosphere. Such analyses will figure prominently in future searches for "biomarkers" – molecular spectral signatures of biological processes on other planets (but more on that later).

After the first transit discovery (which was quickly followed by the additional observations that are possible, outlined above), many surveys to search for transiting planets among nearby stars were initiated. This included two space-based missions, CoRoT and Kepler. Thanks to the photometric stability possible above our constantly changing atmosphere, these can detect transit amplitudes down to a part in 10,000. This is the Earth-size domain; an earth transiting a sun will cause a reduction of 1/10,000 in the observed light (Earth has 1/100 the radius of the sun). In spring 2011, the Kepler team announced the discovery of 1235 planet transit candidates, orbiting 997 host stars, based only on the first four months of data from the mission. This included many multiple-planet systems (including a 6-planet one), and 68 roughly Earth-size objects. Fifty-four of the planets are within the habitable zones of their stars, including 5 of the Earth-mass ones. With the caveat that these candidates still require confirmation via radial-velocity measurements, Kepler has already almost quadrupled the number of known planets, and has made the transit method the most productive one.

3.3. The direct imaging method

Paradoxically, the method of detecting extrasolar planets that is conceptually the simplest, getting a picture of a star and looking for little planets near it, is also the most challenging technologically. Nevertheless, the tally of planets discovered in this way has recently risen to 25. The challenge lies in the huge contrast, at small angular separation, between the brightness of a star and the very faint planet seen mainly or entirely by the star's reflected light. For example, the sun and Jupiter, as viewed in visible light from one of the stars nearest to us, would have a brightness contrast of about 1 billion, but at an angular separation of about 5 arcseconds (1 arcsecond is 1/3600 of a degree). For the sun and Earth, the contrast is about 10 billion, with a separation of only 1 arcsecond.

The contrast ratio can be lowered by a few orders of magnitude by observing in the infrared, taking advantage of the fact that planets are much cooler than stars, and therefore emit more of their light at those wavelengths. Nevertheless, direct imaging remains very difficult because, even given perfect telescope optics (which is an unachievable idealization), the wave nature of light dictates that light from a source, no matter how compact, when imaged through an aperture, is spread out over an extended region in the focal plane in a "diffraction pattern". The angular size of the diffraction pattern is set by the wavelength λ and the size *D* of the entrance aperture of the instrument (e.g. the diameter of the telescope) roughly as $\alpha = \lambda/D$. For the largest telescopes, with D=10 meters, imaging in near-infrared light, e.g., $\lambda = 2$ microns, we get (after converting to suitable units) $\alpha = 0.04$ arcseconds, i.e., about half of the light from a nearby star in concentrated in a spot having a radius 1/100 of the projected separation between that star and a Jupiter-like planet. Sounds good. Unfortunately, there is the other half of the light, which is spread further out in the diffraction pattern. With a factor of 1 billion in contrast in the total light between the star and the planet, the outer parts of the star's diffraction pattern still constitute a huge background that drowns out the planet's light, even at a separation of 5 arcseconds. To make things worse, any slight imperfections in any of the optics will further enlarge and complicate the shape of the diffraction pattern.

Despite these challenges, projects are underway to successfully image extrasolar planets. One approach involves specially designed "apodizing aperture masks". When placed on the aperture of a telescope, they will produce an azimuthally asymmetric diffraction pattern, in which light is concentrated more along one axis than along the perperdicular axis. One can then search for the faint planets along the darker axis with its lower background. Another approach is infrared interferometry, where light is combined from several widely spaced telescopes. The telescope separation B, which can be of order hundreds of meters, now replaces D in the diffraction limit equation above, and the diffraction pattern can be correspondingly more concentrated. Furthermore, using a variant of this technique called nulling interferometry, one can search for the planet in the regions where the combined light of the star from the various telescopes interferes destructively, producing a relatively dark background. By changing the separations among the telescopes, one can null and scan for planets in different regions around the star.

The most ambitious missions of this type conceived for the next decades are concepts like Darwin and Terrestrial Planet Finder, which consist of a flotilla of space telescopes flying in formation. They aim to not only obtain images of terrestrial extrasolar planets, but to follow up with spectroscopy in search of biomarkers. Unfortunately, in the current global financial atmosphere, these projects have been frozen indefinitely.

3.4. The gravitational microlensing method

All of the methods outlined above can be applied only to the nearest stars, at distances of tens of light years. In the first two methods, we need large quantities of light in order to obtain high accuracies, whether spectroscopic or photometric. In the direct imaging method, the same applies, and we also need to maximize the angular separation of the planet from its parent star and its glare. There is one planet-hunting method, however, that is particularly suited for finding planets around stars at distances of 10 to 30 thousand light-years, typical distance scales across the Milky Way galaxy. That method is gravitational lensing. Before addressing its application to planets, let us understand the basics of this effect.

Gravitational lensing refers to the phenomenon whereby the gravitational field in the region around a mass concentration causes light rays propagating through the region to be deflected. “Lensing” (for short), was the first prediction of Einstein’s 1915 general theory of relativity to be verified experimentally, during the 1919 total eclipse of the sun. In the theory, Einstein predicted that stars that happened to be projected near the limb of the sun would appear displaced away from the solar limb by 1.8 arcseconds. A total eclipse, during which the moon hides the glare of the sun and permits seeing stars during daylight, would be an opportunity to measure the effect. Two separate expeditions traveled to two locations in the path of totality, in South America and Africa. The effect to be measured is small, and had to be observed in field conditions, in remote locations, during the brief (few minute) duration of the event, and with the limited technology of the time. In view of this, it is not surprising that the results were ambiguous, with one experiment reporting agreement with Einstein’s prediction, and the other not. Nevertheless, Arthur Eddington, the prominent physicist of the time and a champion of Einstein’s work, after analyzing the results declared that the theory had been vindicated. Although Einstein, by then, was well known among physicists, he was not a public figure. However, the eclipse story reached the headlines of several major newspapers who turned Einstein, literally overnight, into the cult figure he remains today. Thus, lensing is actually what made Einstein famous.

The light rays from any source of light will be deflected (i.e. “lensed”) by any intervening mass lying close to the line of sight of an observer to that source. In particluar, the mass of a star can serve as a lens that deflects the light of another star that happens to lie behind it, if they are at suitable distances from each other and from an observer. When the source star is exactly behind the lens star, the light of the source star, as viewed by the observer, will be distorted into a perfect ring shape - an “Einstein ring” - around the lens star. If the lens star is distant enough, and hence subtends a small enough angle to the observer to avoid hiding the ring, the ring will, in principle, be observable.

Suppose the following: the lens star has the mass of the sun (which is a common type of star, as Bruno guessed); the source star is at a distance of 30,000 light years (the distance to the center of our galaxy); and the lens star is halfway in between. Then, the angular radius of the Einstein ring turns out to be about one milli-arcsecond. If the alignment between source, lens, and observer is not quite perfect, the symmetry of the problem is broken, and the ring breaks up into two distinct arcs straddling the lens. As the alignment is further worsened, both arcs become progressively shorter, with one becoming very faint and eventually disappearing, with the other approaching the size and location of the actual source. (You can see all of this by looking at a light source through the base of a wine glass, which has optics similar to that of a gravitational lens.)

The regime of lensing of stars by other stars is coined "microlensing". At any given moment, from our vantage point in our Milky Way galaxy, such stellar alignments that are good enough to produce perfect or near-perfect Einstein rings are very rare, with about one in a million stars lensing another star at a given moment. However, due to the orbits of the stars (including the sun) around the Milky Way's center, it is a different rare star every time that crosses close enough to the line of sight to another star for the effect to occur. If we had visual-band telescopes with milli-arcsecond resolution (which we do not, yet), and we monitored such a source star, we would see its image gradually getting tangentially stretched around the point where the lens mass is (we need not necessarily see the lens star itself). At the same time, a counter-image would appear and gradually grow on the opposite side of the lens. If, at the moment of closest projected approach, there were near perfect alignment, then the two arcs would merge into a full, or nearly full Einstein ring. Then, as the source and lens continued on their relative trajectories on the sky, the lens would split again into two images, and the whole movie would play itself in reverse as the source gets further and further away on the sky from the lens.

While the splitting and Einstein rings of microlensing are currently unobservable, a secondary effect actually is. What the lens is actually doing to the source is magnifying it, albeit, in a rather peculiar way. This means that light that was intended for someone else is reaching you. As a result, even if you do not see the ring and the image splitting, but just monitor the total amount of light from your source, you will see it brighten, reach a maximum corresponding to the time of best alignment, and then return symmetrically to its normal brightness. The shape of the rise and fall as a function of time is very particular, and can be used to distinguish such a "microlensing event" from other variable astronomical phenomena. The timescale for such an event depends on the several parameters in the problem, but is typically of order weeks. Thus, a microlensing event can be identified by monitoring the light from many stars (of order several million need to be followed in order to have a fair chance of observing the phenomenon), and looking for the specific brightening and fading behavior described above.

Einstein was aware of all of this from the start. His notes from 1912, 3 years before he published general relativity, show the sketches and the basic equations for the lensing of a star by another star. However, for over two decades these results remained unpublished. In 1936, when already living in Princeton, Einstein was approached by an engineer and amateur physicist, Rudy Mandl who, reading about Einstein's theories, had conceived independently of the possibility of microlensing. Einstein confirmed to Mandl that, in principle, such an event could occur, but that in practice it was unobservable, and hence there was no point in publication. Apparently, Mandl persisted in pushing Einstein to publish a paper on the effect, until Einstein reluctantly agreed. In the paper, Einstein emphasized that "There is little chance of observing this phenomenon." And, in a private note to the editor of the journal, he condescendingly wrote, "Let me thank you again for your help with the small publication that Mister Mandl has squeezed out of me. It is of little value, but it makes the poor fellow happy."

It would take over four more decades until lensing became an active observational field, but Einstein's "Mandl-driven" 1936 paper launched a considerable body of theoretical work on the many possibile manifestations that lensing could take, and the astrophysical and cosmological information that could be revealed by observing it.

Lensing by the sun, mentioned at the start, has by now been confirmed by many experiments to obey general relativity's prediction to great precision. The first additional astronomical gravitational lensing phenomena were discovered starting in 1979, with hundreds of more examples turning up in the subsequent decades. These cases involved galaxies, or their sometimes-active central regions, called quasars, serving as light sources and being lensed by the masses of entire intervening galaxies or clusters of galaxies. For the masses and distances involved in such cases (10^{10} to 10^{12} solar masses, and billions of light years, respectively) Einstein rings and related phenomena occur on angular scales of one to a few tens of arcseconds, resolvable by telescopes on Earth, and even better by telescopes (such as Hubble) above the Earth's distorting atmosphere. The first microlensing (i.e., lensing of stars by other stars) events were announced in 1993. By now, thousands of microlensing events have been detected and measured based on the particular symmetric brightening and fading behavior of a source star.

Where, then, did Einstein go wrong in his assessment of the observability of gravitational lensing? In 1936, it was indeed impossible to monitor many millions of stars for periods of years, in order to find the handful undergoing transient microlensing magnification[1] Einstein could hardly have foreseen the development of large digital imaging arrays combined with

1 Furthermore, it may be that he did not even think of the time-variable aspect of the problem, and was considering only stationary configurations with constant magnification.

modern computing power, which permit searching automatically for these rare events. In fact, even when the idea of microlensing surveys was first proposed by Paczynski in 1982, it was considered unfeasible, but Moore's Law of exponentially increasing computing power turned it into a reality within less than a decade.

Returning to the issue of planets, imagine now that a star is lensing the light of another star that is behind it into a complete, or nearly complete, Einstein ring. If the lens star has a planet near it, and the light rays producing the ring image happen to pass near that planet, the planet's gravitational field will cause an additional deflection of the rays. When we monitor the light from the source star as it passes behind the lens, we will see deviations from the simple symmetric brightening and fading produced by a single, isolated star. This perturbation in the brightness as a function of time, caused by the planet, can assume a rich variety of forms, depending exactly on the mass ratio of the lens star and its planet, and on the location of the planet relative to the path of the source star in the background. But in general, these perturbations signaling the presence of a planet or planets will be brief compared to the entire lensing event, often lasting only a few hours. Thus they will be "caught" only if the event is monitored around the clock, with few gaps. Large magnification events, in which the main lens attains near-perfect alignment with the source, will be particularly sensitive to planets, because the Einstein ring encompasses a large region around the lens star, and hence planets lying over a large region will cause a conspicuous perturbation to the ring. Even Earth-mass and lighter planets, if they lie close enough to the Einstein ring, will cause a significant perturbation and can be detected.

Over the past years, two projects, with the acronyms OGLE (Optical Gravitational Lens Experiment) and MOA (Microlensing Observations in Astrophysics), have been monitoring the brightnesses of tens of millions of stars in the direction of the center of the Milky Way, in search of microlensing events. Because of the large density of stars in this direction, there is the highest probability on the sky of close line-of-sight alignment between two passing stars, and indeed almost all of the thousands of events that have been discovered have been found in this direction in the sky. Considering the facts above, several years ago a collaboration of astronomers with the acronym MicroFUN (Microlensing FollowUp Network, of which I am a member) set out to find planets by means of microlensing using the following strategy. Wait for OGLE or MOA to alert that a particular lensing event may have a large magnification (and thus may be highly sensitive to the presence of planets). Track the brightness changes of that event over its peak, using a network of telescopes around the globe, in order to get the most complete time coverage, with the fewest possible gaps (in which the signature of a planet might get lost). That strategy has proved to be effective. Over the past 7 years, a dozen extrasolar planets have been found through microlensing, almost all involving observations by MicroFUN. A "second generation" of microlensing experiments that has just begun should discover of order 10 new planets per year [4].

Contrary to the "strange" planets found by other techniques, the planets turning up by the microlensing searches so far seem to be quite "normal" planets - mostly Neptune-mass to Jupiter-mass planets on Jupiter-like orbits. More specifically, the planets being discovered by microlensing are generally in the region of the "snowline" of their parent stars. The snowline is the distance from a star beyond which the temperature is low enough for water vapor to condense into ice (this depends also on the pressure of the water vapor). According to the most popular scenario for planet formation, the availability of water ice in large quantities just outside the snowline allows the formation of relatively large agglomerations of planetesimals composed of rock and ice. These, in turn, serve as cores that are massive enough to accrete, and hold on to, a large mass of gas, leading to the formation of gas giants. As one goes to large distances, less raw material is available, and progressively smaller gas and ice giants are formed. This explains the mass sequence seen in the solar system, with the first and most massive gas giant, Jupiter, just outside the snowline, and progressively smaller gas giants at larger distances from the sun - Saturn, Uranus, and Neptune. Inside the snowline, one finds only the small rocky planets: Mercury, Venus, Earth, and Mars.

Stars that are less massive than the sun are cooler and less luminous, and hence have snowlines at smaller radii than the sun, and vice versa for more massive stars. The fact that microlensing-dicovered planets have been found largely near the snowlines of their stars is the result of a fortunate coincidence: The Einstein-ring radius of a solar-mass star serving as a lens at a typical distance in the Milky Way happens to be similar to the radius of the snowline of a solar-mass star. Since the Einstein-ring radius is proportional to mass (to its square root, actually), lower-mass stars will have both smaller Einstein rings and smaller snowline radii. And since microlensing is most sensitive to planets in the region of the Einstein ring, it is no surprise then that most of the microlensing-discovered planets are turning up near the interesting region of their parent stars' snowlines.

But do these "normal" extrasolar planetary systems resemble our own in other respects as well? Microlensing has provided a first, tentative, "yes" to this question. In April 2006, the MicroFUN collaboration monitored a microlensing event (with the uninteresting name OGLE-2006-BLG-109; this was the 109th event discovered in 2006 by the OGLE network in the direction of the "bulge" of the Milky Way), an event that promised to rise to large magnification and therefore to be sensitive to planets around the lens star. Early on, indeed, perturbations in its brightening pattern indicated the presence of a Saturn-mass planet. However, once all the data were collected and analyzed, it became clear that they could not be explained solely with that single planet. The collaboration's Science journal article by Gaudi et al. [5] showed that another, Jupiter-mass planet on a closer orbit was required by the data. The signature of the "Jupiter" in this first discovery via microlensing of a planetary *system* (i.e., a system with more than one planet) was visible only for a few hours, at a time when all but one of the 12 telescopes in the network were in daylight, and hence could not observe. The information on this second planet came solely from the Wise Observatory

1-meter telescope in Israel. The event was overall observed well enough that it permitted determination of the sytems's parameters better that any previous microlensing planet discovery. The system's masses, separations, and distance to Earth, can all be measured to an accuracy of about 10%.

The picture that emerges is of a planetary system very reminiscent of the solar sytem. The mass ratio (0.37) of the "saturn" and the "jupiter" in the system is like the mass ratio (0.30) of Saturn and Jupiter in the solar system. The distance ratio of the two planets from their star (0.50) is similar to that of Saturn and Jupiter (0.55). But in terms of absolute values, everything is roughly scaled down by one half: the star has one-half the mass of the sun; the planetary distances are close to one half the distances of Jupiter and Saturn to the sun; and the masses of the planets are smaller than Jupiter and Saturn. So, effectively, this is a scaled-down solar system. This is exactly what one would expect from the standard theory of planetary formation described above: a lower-mass star would result in a closer-in snowline, and therefore the same descending sequence of gas giants, but closer in. With just one example so far, it is much too soon to jump to conclusions. But finding (as soon as we had available a technique than *could* find) a system that resembles the solar system so nicely, with the expected scaling, hints that the Copernican principle - we are not in a privileged or special position (which in the end, was all that Giordano Bruno was invoking) - has been successful yet again: yes, it seems quite possible that many or most stars have planetary systems very similar to that of the sun.

We cannot say whether or not the OGLE-2006-BLG-109 system includes additional planets, and specifically an "earth", perhaps also scaled down in mass and orbit. This specific event did not have the sensitivity to discover such an inner planet. However, earth-mass planets are being and will be discovered, whether by microlensing or by the other methods. The progress of the past few years and the near-term future exoplanet projects tell us that a full picture of the frequency of occurrence, the characteristics, and the variety of types, of planetary systems is just around the corner.

4. The Next Step: Astrobiology

After finding extrasolar terrestrial planets that seem to be, in principle, capable of sustaining life, the next obvious step will be to actually search for signs of life. The fairly new science of studying life outside the Earth is called Astrobiology. How will we go about it? Beyond the bodies in the solar system, for which life-searching experiments can be done by robots and space probes, searches for life on extrasolar planets will always involve remote sensing. The simplest way to find evidence of life may be by life's indirect effect on the environment. On Earth, the oxygen in the atmosphere is of biological origin, having been first released by cyanobacteria 2-3 billion years ago, and later boosted by blue-green algae and plants. Without life, oxygen in the atmosphere would decrease to very low levels within a few million years, similar to the situation in Mars today. Thus the first biomarkers that will be searched for are oxygen, both as diatomic (O_2) and ozone (O_3) molecules.

Finding such molecules in an exoplanet's atmosphere together with water vapor and CO_2, in proportions similar to Earth, would be strong evidence for the presence of life. The spectral signatures of these atoms in an atmosphere can be detected either in the transmitted light or the reflected light of a transiting planet (see above). Such measurements should be possible within a few years with the James Webb Space Telescope, NASA's replacement for Hubble. The same spectral biomarkers can be found by analyzing the light of planets isolated from the glare of their parent stars by means of interferometric imaging.

In terms of more direct signs of life, optical imaging of the Earth's surface from artificial satellites easily reveals vegetation by means of a spectral signature called the "red edge", due to chlorophyll, at 700 nm. This feature is strong enough that it could be detected on Earth-like planets among those that, e.g., the Darwin mission might find. In principle, one could do much more. Remote sensing of the Earth is a mature science, and Earth-imaging satellites can routinely identify and measure the density of specific plant species based on their spectral signatures. Doing the same for an extrasolar planet at a huge distance is simply a question of having a sufficiently strong signal. In practice, going to such levels of detail in analyzing life on other planets is unfeasible with currently conceivable technology. The telescope sizes (and costs) required to gather enough of the feeble light are impractical. We should, however, remember the lessons of the past; what is impractical today may become fairly easy with the advent of new, yet unimagined, technology.

A possibly more serious problem is the non-Copernican, anthropocentric aspect of the strategies outlined above for searching for signs of life. Of course, it would be amazing if we do find the signature of chlorophyll in a distant planet. But should we expect it? And if we do not see it, will that exclude the presence of life there? Probably not. This is a lesson that has been learned in recent years, actually by studying life here on Earth. A huge variety of previously unknown organisms, mostly microbial, is being discovered, not only in normal environments such as the oceans, but also in the most unexpected of places - near undersea oceanic vents, in deep underground aquifers, within rocks, in ice, and in hot acidic lakes. Each of these "extremophile" life forms has adapted to exploit a different energy source that is available in its particular niche.

For example, Chivian et al. [6] have discovered a self-sustaining community of bacteria in the Mponeng gold mine in South Africa at a depth of 3 km. These organisms, which have been isolated from the Earth's surface for millions of years, derive their metabolic energy indirectly from the natural radioactivity of the surrounding rock. The radioactive radiation dissociates molecules of ground water to hydrogen and oxygen. The hydrogen combines with sulphur to form compounds that the bacteria feed on. It seems that wherever there is liquid water (at least some of the time), a source of energy, and some common chemical compounds, an organism has developed to exploit it. On the one hand, this bodes well for finding life in remote environments that are different from those we traditionally consider as hospitable to life. On the other hand, it raises the problem of how to recognize remotely the signs of activity by such lifeforms, which can be so different from those we know on Earth.

5. Why Finding Life (or not) Matters

The popular concept of extraterrestrial life has not changed much since Lucian's second-century fantasies. It almost always consists of a Universe densely populated by sentient beings, more or less technologically sophisticated, but remarkably similar to humans in terms of body plan and all aspects of behavior. This concept is constantly reinforced by a huge entertainment industry, but also by governments (mainly the US), who are aggressively promoting and pursuing the idea that humanity's "destiny" is to colonize space. The largest and most expensive space project of recent years is the International Space Station. Although the word "science" is often heard in the context of the space station, there is actually little science that is carried out there, and the main work of the astronauts is home maintenance. The express purpose of the project is to prepare for the colonization of the moon and Mars. Such colonization, again, has little to do with science. Scientific exploration of the solar system can be done much more effectively, safely, and inexpensively with robotic probes than with astronauts. But NASA's great reputation and the mix of science with manned space flight certainly end up promoting in the public perception the plausibility of Star Wars imagery.

In reality, extraterrestrial life, if it exists and we find it, will certainly not be of the "Mars Attacks" film sort. We can look at the evolution of life on Earth for guidance. Life on Earth emerged perhaps as soon as the young planet cooled enough, after a billion years or so, and hence Earth has been inhabited for most of its 4.5 billion year history. However, land plants and insects have been here for only the last 10% of the time. Reptiles have existed for only 7% of the time, and mammals for only 4%. Humans have been around for only of order 100,000 years. Civilizations have existed for less than 10,000 years, of order one millionth the age of the Earth. And the capacity for radio communication and space travel have been here only 50 - 100 years, of order one part in one-hundred-million of Earth's history.

We do not know, of course, for how much longer there will be humans on Earth, doing the things we do today. The possibilities for human extinction through war, artificial, or natural global catastrophes (climate change, disease, asteroid impacts) are numerous. Looking at the history of other species and at the extinction record, it is hard to imagine that humans in their present form will still exist on geological timescales into the future. If so, this would mean that, to find extraterrestrial creatures of the sort we see in the movies (i.e., very similar to us), we would have to search among one-hundred million earths. Even if earths are abundant around stars, and evolution always proceeds as it did on Earth (which is contrary to the whole principle of randomness behind evolution itself, and therefore is highly unlikely), there are only about one-million stars within 300 light years of the sun. It is only within such a distance that there is any hope, in the foreseeable future, of remotely sensing any evidence for life, using the various techniques outlined above. So, just from these simple considerations, it is almost certain that extraterrestrial life, should we find it, will be microbial, or not much more complex than that.

If human extinction of the "back-to-square-one" kind is avoided and evolution continues at the pace it has, then our descendants, and those of similar species on other planets, are likely to be so highly advanced compared to us that we face a new problem: next to them, we would be like microbes. Considering mutual visits and communication between galactic civilizations, it is unlikely that a species as advanced as we might be a billion years hence would have any interest in communicating with "microbes".

Despite these sobering facts, the search for extraterrestrial life is among the most exciting human endeavors ever, and I believe it must be pursued. The reason has to do with the very existence of life on Earth. The emergence of the first life forms on Earth is a complete mystery. Although some ideas are emerging, creating the first structure that could be called a living organism, one that then reproduces, multiplies, and evolves, still seems like an insurmountable challenge. Was "abiogenesis" a wildly improbable accident that occurred only once, here on Earth? Or is the appearance of life unavoidable whenever the appropriate, but altogether common, conditions exist? There are hints pointing in both directions.

Amino acids are the components of proteins, which, in turn, are the building blocks of living organisms. When synthesized in the laboratory, amino acid molecules are created in similar numbers in two mirror forms that are chemically equivalent, called right-handed and left-handed. Left-handed amino acids can combine only with other left-handed amino acids to build proteins, and the same is true for right-handed ones. A creature composed of right-handed amino acids would be able to eat only organisms that are right-hand based, and to use their amino acids in order to grow and reproduce[2]. If life emerged quickly and spontaneously in the young Earth, wherever conditions were suitable, we would expect to see both right-hand-amino-acid and left-hand-amino-acid life forms.

All of the above applies also to sugars. However, all known organisms on Earth are based on left-handed amino acids and right-handed sugars. This suggests that all organisms on Earth descended from one single "mother" cell, that happened to have this handedness of amino acids and sugars. It would seem, then, that the emergence of such a viable cell is something exceedingly rare and improbable, that occurred just one single time in Earth's history.

2 In fact, opposite-handed organisms would likely be poisonous to it.

Another hint in this direction comes from the fact that all living organisms, without exception, use the adenosine triphosphate (ATP) molecule as an energy "currency" for managing and transferring energy in the cell. One could imagine many different molecular mechanisms for energy manipulation that would have developed in organisms that had had an independent start. The universality of ATP again hints at a single ancestor.

On the other hand, the speed with which life appeared on Earth argues for the "inevitable life" option. Perhaps the specific handedness of amino acids and sugars was somehow enhanced in the "primordial soup" out of which the first organisms formed. And, perhaps ATP gave our microbial ancestors some evolutionary advantage over other species using different energy currencies, and those species all became extinct.

The search for extraterrestrial life offers a way to address this mystery, which is at the heart of understanding life itself. If, over the next years, we discover a Universe teeming with life, albeit primitive, it will be a clear verdict that nature forms life "easily" under suitable conditions, even if we do not understand the process at present. Alternatively, a complete absence of what we could call life on other planets, after excluding also the possibility of exotic and difficult-to-identify life forms, different from those most familiar from Earth, would also have profound implications. It would tell us that, at least in our local neighborhood of stars, we are truly alone, and that we are the result of a highly improbable chain of events, a "miracle".

6. Some Final Thoughts

We have seen that Giordano Bruno, with an Olympian intution, foresaw in the late 16th century many of the facts that have been established only in the last few decades and years. However, during his times, Bruno's ideas, coupled with a tendency to get into trouble, led him to the life of a fugitive, drifting across the capitals of Europe and constantly making new enemies. The Inquisition eventually caught up with him, imprisoned him for 8 years, convicted him of heresy, and burned him at the stake in Rome in 1600. There is debate among historians about the weight of Bruno's astronomical ideas in sealing his fate, compared to that of his heretical opinions on other matters of Church dogma. But we can probably be confident that his astronomical ideas did not help him.

Returning to ``De L'Infinito Universo et Mondi", we find in the treatise, along with the dialogue exposing the ideas that have been discussed above, three sonnets by the author, as sometimes found in Rennaissance essays. Written in 1584 in London, the third sonnet evokes an eerie feeling that, among all that Bruno foresaw, he foresaw also his own final fate. One cannot help imagining Bruno reciting the poem 16 years later in his dungeon cell in Rome, as he awaits his execution. In beautiful verse, he expresses the humanistic Rennaissance idea of the power of thought, observation, and reasoning to transcend all physical obstacles and distances, an idea that I feel is still at the heart of basic science, and of astronomy in particular. With no humility whatsoever, Bruno concludes by notifying his contemporaries that he has seen much further than they ever will. Indeed!

I reproduce below the Italian text, and provide my humble attempt at an English translation.

E chi mi impenna, e chi mi scalda il core?	*And who delights me, and who warms my heart?*
Chi non mi fa temer fortuna o morte?	*Who makes me fear neither fortune nor death?*
Chi le catene ruppe e quelle porte,	*Who breaks the chains and those doors,*
Onde rari son sciolti ed escon fore?	*through which few are released and exit?*
L'etadi, gli anni, i mesi, i giorni e l'ore,	*The seasons, the years, the months, the days, and the hours,*
Figlie ed armi del tempo, e quella corte	*Daughters and weapons of time, and that court*
A cui ne ferro, ne diamante son forte,	*Against which neither iron nor diamond is strong,*
Assicurato m'han dal suo furore.	*They have safeguarded me from the fury.*
Quindi l'ali sicure a l'aria porgo;	*Therefore, confident wings to the air I spread,*
Ne temo intoppo di cristallo o vetro,	*I fear not obstacles of crystal or glass,*
Ma fendo i cieli e a l'infinito m'ergo.	*But cleave the heavens and toward the infinite I rise.*
E mentre dal mio globo a gli altri sorgo,	*And while from my sphere to the others I surge,*
E per l'eterio campo oltre penetro:	*And through the ethereal field I further penetrate,*
Quel ch'altri lungi vede, lascio al tergo.	*That which others see far away, I leave behind me.*

Acknowledgements

I thank Giorgio Dieci for pointing out Giordano Bruno's beautiful sonnet, Omer Bromberg for Lucian's "True Story", and Marco Bersanelli and Charles Harper for organizing an extremely stimulating and enjoyable meeting.

References

1 Mumma, M. J., Villanueva, G. L., Novak, R. E. et al. 2009, Science, 323, 1041
2 Latham, D.W, Stefanik, R.P., Mazeh, T., Mayor, M., Burki, G. 1989, Nature, 339
3 Mayor, M. & Queloz, D. 1995, Nature, 378, 355
4 Shvartzvald, Y. & Maoz, D. 2011, MNRAS, 1939
5 Gaudí, B. S., Bennett, D. P., Udalski, A. et al. 2008, Science, 319, 927
6 Chivian, D. et al. 2008, Science, 322, 275

Volume 2
Winter 2012

euresis
Association for the Promotion of the Scientific Endeavour