Saas-Fee Advanced Course 31

P. Cassen T. Guillot A. Quirrenbach

Extrasolar Planets

Saas-Fee Advanced Course 31

Swiss Society for Astrophysics and Astronomy
Edited by D. Queloz, S. Udry, M. Mayor and W. Benz

With 140 Figures, 10 in Color

Springer

Patrick Cassen
SETI Institute
515 N. Whisman Rd.
Mountain View CA 94043 USA
pcassen@mail.arc.nasa.gov

Andreas Quirrenbach
Sterrewacht Leiden
PO Box 9513
2300 RA Leiden, The Netherlands
quirrenbach@strw.leidenuniv.nl

Tristan Guillot
Observatoire de la Côte d'Azur
BP 4229
06304 Nice, France
guillot@obs-nice.fr

Volume Editors:
Michel Mayor
Didier Queloz
Stephane Udry
Observatoire de Genève,
Chemin des Maillettes 51,
1290 Sauverny, Switzerland
michel.mayor@obs.uige.ch,
didier.queloz@obs.uige.ch,
stephane.udry@obs.uinge.ch

Willy Benz
Universität Bern,
Physikalisches Institut,
Sidlerstrasse 5,
3012 Bern, Switzerland

This series is edited on behalf of the Swiss Society for Astrophysics and Astronomy:
Société Suisse d'Astrophysique et d'Astronomie
Observatoire de Genève, ch. des Maillettes 51, 1290 Sauverny, Switzerland

---

Cover picture: Artwork of "51 Pegasus" made by L.R. Cook shortly after the announcement of the discovery of the planet. With permission of the artist.

Library of Congress Control Number: 2005938665

ISBN-10 3-540-29216-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-29216-6 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Printed in The Netherlands

Typesetting by authors and SPI Publisher Services using a Springer LaTeX macro package
Cover design: *design & production*, Heidelberg

Printed on acid-free paper SPIN: 11563556 55/SPI 5 4 3 2 1 0

Preface

In April 2001 the Swiss Society of Astrophysics and Astronomy (SSAA) organized its 31st winter "Saas-Fee" course on "Brown Dwarfs and Planets" in a picturesque resort at Grimentz on the Swiss Alps. The range of topics mainly focused on extrasolar planets' science. We entitled these lecture notes "Extrasolar Planets."

Research on extrasolar planets is one of the most exciting fields of activity in astrophysics. In just a decade a huge step has been made from the early speculations on the existence of planets orbiting "other stars" to the first discoveries and the characterization of extrasolar planets. This breakthrough is the result of the growing interest of a large community of researchers as well as the development of a wide range of new observation techniques and facilities. We organized the 31st winter course to cover all relevant aspects of this new field: observation and detection techniques, physics of their interior, and physics of their formation. We were very happy to have three senior lecturers, Andreas Quirrenbach, Tristan Guillot, and Patrick Cassen, cover these three subjects. They provided information to more than 100 participants and also gave updated comprehensive course materials, which is a challenging task considering the rapid development of this field of research. We hope that the level of details and the comprehensive view offered by authors will be appreciated as a comprehensive detailed introduction to this exciting subject.

We would like to warmly thank our three speakers for the high standard of their lectures and notes, as well as their discussion with students. We also thank all participants for their participation, kindness, and enthusiasm in taking part in the events organized. We thank Dominique Briguet of "A La Marena" for his hospitality and his help with the local organization. We would also like to warmly thank Elisabeth Teichamann, our course secretary, who gave us immense support during the preparation of the meeting as well as during the course. This course has been made possible thanks to a grant from the Swiss Academy of Sciences.

Geneva
June 2005

Didier Queloz
Stéphane Udry
Willy Benz
Michel Mayor

Our three lecturers: (*left*) Pat Cassen, (*middle*) Andreas Quirrenbach, (*right*) Tristan Guillot

Contents

Detection and Characterization of Extrasolar Planets

List of Previous Saas-Fee Advanced Courses

!! 2004 The Sun, Solar Analogs and the Climate
J.D. Haigh, M. Lockwood, M.S. Giampapa

!! 2003 GravitationalLensing: Strong,Weak and Micro
P. Schneider, C. Kochanek, J. Wambsganss

!! 2002 The Cold Universe
A.W. Blain, F. Combes, B.T. Draine

!! 2001 Extrasolar Planets
T. Guillot, P. Cassen, A. Quirrenbach

!! 2000 High-Energy Spectroscopic Astrophysics
S.M. Kahn, P. von Ballmoos, R.A. Sunyaev

!! 1999 Physics of Star Formation in Galaxies
F. Palla, H. Zinnecker

!! 1998 Star Clusters
B.W. Carney, W.E. Harris

!! 1997 Computational Methods for Astrophysical Fluid Flow
R.J. LeVeque, D. Mihalas, E.A. Dorfi, E. Müller

!! 1996 Galaxies Interactions and Induced Star Formation
R.C. Kennicutt, F. Schweizer, J.E. Barnes

!! 1995 Stellar Remnants
S.D. Kawaler, I. Novikov, G. Srinivasan

* 1994 Plasma Astrophysics
J.G. Kirk, D.B. Melrose, E.R. Priest

* 1993 The Deep Universe
A.R. Sandage, R.G. Kron, M.S. Longair

* 1992 Interacting Binaries
S.N. Shore, M. Livio, E.J.P. van den Heuvel

* 1991 The Galactic Interstellar Medium
W.B. Burton, B.G. Elmegreen, R. Genzel

* 1990 Active Galactic Nuclei
R. Blandford, H. Netzer, L. Woltjer

! 1989 The Milky Way as a Galaxy
G. Gilmore, I. King, P. van der Kruit

! 1988 Radiation in Moving Gaseous Media
H. Frisch, R.P. Kudritzki, H.W. Yorke

! 1987 Large Scale Structures in the Universe
A.C. Fabian, M. Geller, A. Szalay

! 1986 Nucleosynthesis and Chemical Evolution
J. Audouze, C. Chiosi, S.E. Woosley

! 1985 High Resolution in Astronomy
R.S. Booth, J.W. Brault, A. Labeyrie

! 1984 Planets, Their Origin, Interior and Atmosphere
D. Gautier, W.B. Hubbard, H. Reeves
! 1983 Astrophysical Processes in Upper Main Sequence Stars
A.N. Cox, S. Vauclair, J.P. Zahn

* 1982 Morphology and Dynamics of Galaxies
J. Binney, J. Kormendy, S.D.M. White

! 1981 Activity and Outer Atmospheres of the Sun and Stars
F. Praderie, D.S. Spicer, G.L. Withbroe

* 1980 Star Formation
J. Appenzeller, J. Lequeux, J. Silk

* 1979 Extragalactic High Energy Physics
F. Pacini, C. Ryter, P.A. Strittmatter

* 1978 Observational Cosmology
J.E. Gunn, M.S. Longair, M.J. Rees

* 1977 Advanced Stages in Stellar Evolution
I. Iben Jr., A. Renzini, D.N. Schramm

* 1976 Galaxies
K. Freeman, R.C. Larson, B. Tinsley

* 1975 Atomic and Molecular Processes in Astrophysics
A. Dalgarno, F. Masnou-Seeuws, R.V.P. McWhirter

* 1974 Magnetohydrodynamics
L. Mestel, N.O. Weiss

* 1973 Dynamical Structure and Evolution of Stellar Systems
G. Contopoulos, M. Hénon, D. Lynden-Bell

* 1972 Interstellar Matter
N.C. Wickramasinghe, F.D. Kahn, P.G. Metzger

* 1971 Theory of the Stellar Atmospheres
D. Mihalas, B. Pagel, P. Souffrin

---

* Out of print
! May be ordered from Geneva Observatory
Saas-Fee Courses
Geneva Observatory
CH-1290 Sauverny
Switzerland
!! May be ordered from Springer-Verlag

Detection and Characterization of Extrasolar Planets

A. Quirrenbach

1 Methods of Planet Detection

1.1 The Quest for Planets Around Other Stars

The realization that our Sun is just one "average" star amongst billions and billions in the Sky naturally brings with it the question whether some – or perhaps most – of the other stars may also harbor planetary systems like our own. We live in a remarkable epoch, being the first generation that has obtained an affirmative answer to this question, that is undertaking programs to characterize the physical properties of planets outside the Solar System, and that is developing the tools to search for twins of the Earth. For the first time in human history, we are on the verge of being able to address the questions whether there are other habitable worlds, and to search for life elsewhere in the Universe with scientific methods.

The search for extrasolar planets has a long and checkered history (see e.g., Boss 1998a for an easily readable overview). Because of the enormous brightness contrast between planets and their parent stars, the direct detection of planets by taking images of the vicinity of nearby stars would be extremely difficult. Early searches for planets were therefore mostly carried out with the astrometric method, which seeks to detect the motion of the star around the center of mass of the star–planet system (see Sect. 9). First reports on the detection of massive planets ($\sim 10\,M_{\rm jup}$) were published during World War II (Strand 1943; Reuyl and Holmberg 1943), but remained controversial, both with regards to the reality of the results and to the question whether the detected bodies should be called "planets". Much painstaking work over the next few decades lead to the realization that these "detections" were spurious. Continued improvements in the astrometric accuracy finally culminated in the announcement of a planet 1.6 times as massive as Jupiter in a 24-year orbit around Barnard's Star (van de Kamp 1963). A decade earlier Otto Struve had written a remarkable paper, in which he noted the possibility that Jupiter-like planets might exist in orbits as small as 0.02 AU, proposed to search for these

Quirrenbach A (2006), Detection and characterization of extrasolar planets. In: Mayor M, Queloz D, Udry S and Benz W (eds) Extrasolar planets. Saas-Fee Adv Courses vol 31, pp 1–242
DOI 10.1007/3-540-29216-0_1

objects with high-precision radial-velocity measurements, and pointed out the feasibility of photometric searches for planets eclipsing their parent stars – all on little more than one journal page (Struve 1952).

By the mid-sixties, the search for extrasolar planets thus appeared to be a thriving field, with eight planetary companions known from astrometric observations (two of them classified as "existence not completely established"), and a number of potentially promising alternative search methods under consideration (O'Leary 1966). By the same time it had also been recognized that brown dwarfs (termed "black dwarfs" at the time) would form a class of their own, with properties intermediate between those of stars and planets. Both astrometric searches for brown-dwarf companions to low-luminosity stars, and attempts at finding them directly with high-resolution imaging techniques, seemed to be successful (Harrington et al. 1983; McCarthy et al. 1985).

Sadly, none of these early claims for detections of planets and brown dwarfs withstood the test of time. It turned out that systematic instrumental errors had been mistaken for the "planetary companion" of Barnard's Star (Gatewood and Eichhorn 1973). What appeared to be the most convincing detection of a brown dwarf, a companion to the star VB 8, could never be confirmed (Perrier and Mariotti 1987; Skrutskie et al. 1987). Other putative planets and brown dwarfs did not fare better. By the mid-nineties, all that remained was a candidate brown dwarf companion of HD 114762, detected with the radial-velocity method (Latham et al. 1989).[1]

This situation changed completely and abruptly with the discovery of 51 Peg b, a Jupiter-like planet in a 4-day orbit (Mayor and Queloz 1995), which has opened a completely new field of astronomy: the study of extrasolar planetary systems. About 150 planets outside our own Solar System are known to date, and new discoveries are announced almost every month. These developments have revolutionized our view of our own place in the Universe. We know now that other planetary systems can have a structure that is completely different from that of the Solar System, and we have set out to explore their properties and diversity.

The following chapters introduce the most important methods that have been employed (or proposed) for the detection of extrasolar planets, and for studies of their physical characteristics. Emphasis is given to observational techniques, their foundations, limitations, and their practical implementation. As far as possible, published results are mentioned in the context of the respective observing techniques, and some outstanding implications for the astrophysics of planets and planetary systems are discussed. This will hopefully elucidate the capabilities, strengths, and weaknesses of the many

---

[1] The radial-velocity technique does not allow measuring the companion mass, but only $m \sin i$, where i is the unknown inclination of the orbit (see Sect. 4). It could therefore not be excluded that HD 114762 B is a low-mass star in a nearly face-on orbit.

complementary observational approaches. It should be kept in mind, however, that the study of extrasolar planets is a rapidly expanding field, in which new and unanticipated results appear almost every month. Technical developments in fields such as adaptive optics, coronography, and interferometry are also occurring at a staggering pace. Nevertheless, the systematic introduction of the fundamental principles and methods attempted in this article will hopefully remain a useful guide for a while to come.

1.2 What is a Planet?

The Definition of “Planet”

Before we can begin to answer the question how planets outside the Solar System can be detected and characterized, we must first agree on an operational definition of the term “planet”. The Greek root of the word literally means “unsteady” or “transient”; it was historically applied to the five known “wandering stars” Mercury, Venus, Mars, Jupiter, and Saturn. The Copernican Revolution added the Earth to the list, and the discoveries of Neptune, Uranus, and Pluto completed the census of the large bodies in the Solar System as we know it. The example of Pluto clearly demonstrates the need for a clean definition of the term “planet”. With the discovery of a large number of bodies belonging to the Kuiper Belt (Jewitt and Luu 1993; Luu and Jewitt 2002) it has become clear that Pluto is but the largest member of the class of Trans-Neptunian Objects (TNOs). It has therefore be argued that Pluto should be demoted from its rank among the planets. I would side with the majority view, however, that the use of the term “planet” in the Solar System is based on historical developments and should not be changed retroactively.

The history in our own Solar System thus shows that the use of the term “planet” has been expanded from the original five members of this class, to newly discovered objects that shared the most important properties of the established examples. Two of these additions (Neptune and Uranus) were rather undramatic, one was based on the realization that the Earth shared important properties with the planets (it orbits the Sun between Venus and Mars), and one added a physically distinct and different body to the list (Pluto). Progress in our knowledge about the planets has also taught us that our list includes bodies encompassing wide ranges in mass, composition, and other physical characteristics.

When we look outside our Solar System, we should certainly expect to find a variety of objects that share many characteristics with our planets, but that may be different in one or more important ways. It is thus a matter of definition what we call a “planet” and where we draw the boundaries to other classes of objects. From a practical point of view, this definition should be based on properties that are easily verifiable observationally; this favors a definition based on mass over a definition based on the formation history. Nonetheless, we should not expect that we can easily come up with a set of criteria that will in each case allow an unambiguous classification of a newly discovered as a "planet" (or not). For example, if a maximum mass is included among the defining properties, all objects discovered with the radial-velocity technique – and thus with known $m \sin i$, see Sect. 4.1 – could strictly speaking only be called "planet candidates" before additional information on their orbital inclination is secured.

For the purposes of this article, I take a "planet" to be an object that fulfills the following criteria:

- *A planet is an object in orbit around a star or a multiple star system.* This excludes free-floating planet-mass objects. A number of such objects have been detected with direct-imaging surveys in young clusters (e.g., Zapatero Osorio et al. 2000, 2002; Béjar et al. 2001; Lucas et al. 2001).[2] Free-floating objects are not considered further here, although it is possible that some of them originally formed in a circumstellar disk, and were ejected by a collision with another planet (e.g., Bryden 2001).
- *A planet is not in orbit around another planet.* This requirement excludes moons, but one should point out that the distinction between moons and planets is also somewhat fuzzy. For example, the Pluto–Charon systems could be called a double planet rather than a planet with a moon.
- *A planet has a minimum mass of* 10^{22} kg. This distinguishes planets from planetesimals, asteroids, and comets.[3]
- *A planet has a maximum mass of* $13\,M_{\rm jup}$. This sets the boundary between planets and brown dwarfs. The value of $13\,M_{\rm jup}$ has been chosen to roughly coincide with the Deuterium burning limit (e.g., Burrows et al. 1997a). This criterion will often be applied fairly loosely, as objects with $m \sin i < 13\,M_{\rm jup}$ will also be called "planets" even if there is no additional information on i.

As with the word "planet", we will use other terms established in the Solar System and and apply them to analogous bodies and material around other stars; we can thus speak of "moons" and "rings" around extrasolar planets, about "exo-planetesimals", "exo-comets" and "exo-zodiacal dust".

The Thermal Evolution of Giant Planets

A basic understanding of the evolution of planets is an important prerequisite for a discussion of detection methods. The fundamental principle is rather

---

[2] Note that a tentative detection of free-floating planet-mass objects in the cluster M 22 (Sahu et al. 2001) has been retracted (Sahu et al. 2002).

[3] The value of 10^{22} kg is quite arbitrary, of course. I have chosen it because it separates Pluto from the "minor bodies" in the Solar System.

simple: planets are born hot and are initially self-luminous; they cool during their evolution until they reach radiative equilibrium with their parent stars. The age of a planet is therefore an important parameter that determines how difficult it is to detect its thermal emission.

The luminosity evolution of giant planets, alongside with that of brown dwarfs and low-mass stars, is shown quantitatively in Fig. 1; it can be seen from this figure that an old planet is about four orders of magnitude(!) fainter than it was at an age of 1 Myr. Another important conclusion is that luminosity alone is an extremely poor indicator of the mass of substellar objects – information about the age is crucial to distinguish between low-mass stars, brown dwarfs, and planets. (Dynamically determined masses are even better, of course.) We see, however, that at the same age the luminosity of gaseous planets increases strongly with mass. Figure 1 thus provides a very useful first orientation for general considerations about the detectability of giant planets. For more detailed calculations, one has to take into account the planet's temperature, which determines the spectral energy distribution, and modifications to the luminosity evolution due to irradiation by the host star (e.g., Burrows et al. 2003).

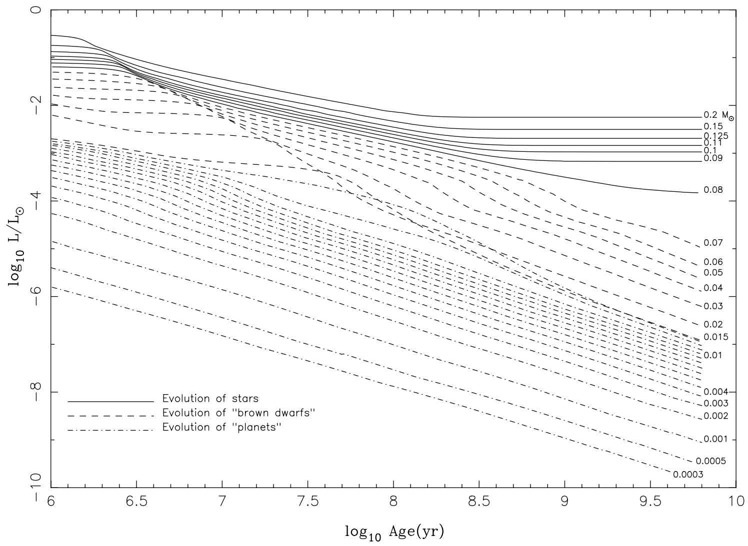

Fig. 1. Evolution of the luminosity (in $L_\odot$) of solar-metallicity M dwarfs and substellar objects vs. time (in yr) after formation. The stars, "brown dwarfs" and "planets" are shown as solid, dashed, and dot-dashed curves, respectively. In this figure, we arbitrarily designate as "brown dwarfs" those objects that burn deuterium, while we designate those that do not as "planets." The masses (in $M_\odot$) label most of the curves, with the lowest three corresponding to the mass of Saturn, half the mass of Jupiter, and the mass of Jupiter. From Burrows et al. (1997a)

1.3 Pulsar Planets

The First Extrasolar Planets

While the considerations of the preceding sections appear to give a solid framework for planet searches, the first firm discovery of objects that fulfill the above definition of a planet came totally unexpected and from a completely different line of research. The extremely stable rotation of pulsars provides a high-precision clock, which can be used for the indirect detection of planets, in a way that is quite similar to the radial-velocity method that will be discussed in detail below (Sect. 4). High-precision monitoring of the time-of-arrival (TOA) of the radio pulses can reveal subtle motions of the pulsar, such as its reflex motion due to the presence of a planetary companion. For a planet with mass m_p in a circular orbit with period P and inclination i, and a "canonical" neutron star mass of $1.35\,M_\odot$, the amplitude of the timing residuals τ is

$$\tau = 1.2\,\mathrm{ms}\left(\frac{m_p}{M_\oplus}\right)\left(\frac{P}{1\,\mathrm{yr}}\right)^{2/3}\sin i\ . \tag{1}$$

For millisecond pulsars, TOA measurements are possible with a long-term precision of a few µs (e.g., Wolszczan 1994). This implies that planets down to $\sim 0.01\,M_\oplus$ are detectable around pulsars; this limit is far lower than that of any other search method currently contemplated.

After a few false starts (e.g., Bailes et al. 1991; Lyne and Bailes 1992), two planets just a factor of ~ 3 more massive than the Earth were found orbiting the pulsar PSR B1257+12 (Wolszczan and Frail 1992). The two planets are in a 3:2 orbital resonance, which leads to accurately predictable periodic perturbations of the two orbits. The detection of this mutual gravitational attraction between the planets provided the final proof of the reality of the first pulsar planets; the same data set also revealed the presence of a third planet with even lower mass in the same system (Wolszczan 1994). The properties of the planets orbiting PSR B1257+12 are listed in Table 1.

Table 1. Parameters of the PSR B1257+12 planetary system

planet	A	B	C
semi-major axis [light-ms]	0.0035	1.3106	1.4121
eccentricity e	0.0	0.0182	0.0264
orbital period [days]	25.34	66.54	98.22
longitude of periastron	–	249°	106°
planet mass [$M_\oplus$]	$0.015/\sin i_1$	$3.4/\sin i_2$	$2.8/\sin i_3$
distance from pulsar [AU]	0.19	0.36	0.47

After Wolszczan (1999)

The Keplerian timing residuals (1) depend on $m_p \sin i$; this means that the mass of the planet and its orbital inclination cannot be determined independently. In contrast, the strength of the mutual interaction between the planets depends directly on their masses. It has thus become possible to infer the masses and inclinations of planets B and C from modeling of a long series of timing data, which now covers more than a decade (Konacki and Wolszczan 2003). The derived masses are $4.3 \pm 0.2\,M_\oplus$ and $3.9 \pm 0.2\,M_\oplus$, respectively, and the orbital inclinations $53° \pm 4°$ and $47° \pm 3°$ (or 127° and 133°), indicating that the two orbits are nearly co-planar.

Even after taking the three planets and the interaction between planets B and C into account, there remains a long-term systematic variation of the TOA residuals (see the lower two panels of Fig. 2). These residuals could be indicative of the presence of a fourth planet with longer orbital period

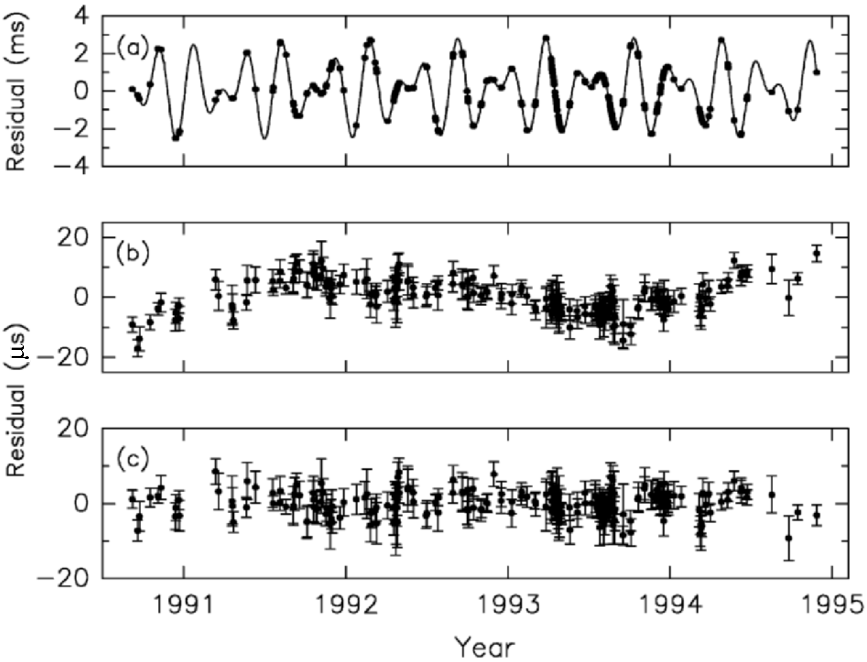

Fig. 2. Timing residuals for PSR B1257+12 at 430 MHz, for three increasingly detailed models. (**a**) Residuals after the fit of the standard timing model without planets. The time-of-arrival variations are dominated by the Keplerian orbital effects from planets B and C. (**b**) Residuals for the model including the Keplerian orbits of planets A, B, and C. Residual variations are determined by gravitational perturbations between planets B and C. (**c**) Residuals for the model including all the standard pulsar parameters, and the Keplerian and non-Keplerian orbital effects. From Konacki and Wolszczan (2003)

(Wolszczan et al. 2000). If the apparent three-year periodicity of the residual signal can be confirmed, this would point to an origin of the disturbance within the pulsar planetary system itself. It will probably be difficult to ascertain the nature of this ionized material – a "coma" ablated from a fourth body or a warped disk are among the possibilities.

Pulsar planets appear to be rare. Only one other pulsar, PSR B1620−26 near the core of the globular cluster M4, has a confirmed planet (Arzoumanian et al. 1996; Joshi and Rasio 1997). The B1620−26 system is rather interesting, too. The planet orbits an inner binary system, which consists of a millisecond pulsar and a white dwarf companion in a half-year orbit. The most likely mass of the planet is $m_p \sin i_p \approx 7\,M_{\rm jup}$, and the semi-major axis and eccentricity of its orbit are $a \approx 60\,{\rm AU}$ and $e \approx 0.45$ (Thorsett et al. 1999; Ford et al. 2000).

The Formation of Pulsar Planets

The theories for the formation of pulsar planets can be broadly divided into two classes: (a) scenarios, in which the planets were formed together with a "normal" star, and survived its evolution from the main-sequence to become a red giant and later a rapidly spinning neutron star; and (b) scenarios in which the formation of the neutron star precedes the formation or acquisition of its planets (Podsiadlowski 1993). The first category implies that the planets must be able to survive the formation of the pulsar, which involves a violent transformation in a supernova explosion, and the supernova recoil. This possibility is generally regarded as unlikely, and scenarios of type (b) are favored.

Consequently, most theories of planet formation around millisecond pulsars concern themselves with possible ways to disrupt, evaporate, ablate, or otherwise dismember the companion star, and thus to transform a fraction of the companion's mass into a gaseous disk around the neutron star (Phinney and Hansen 1993). Such a disk could be formed, for example, by an asymmetric supernova explosion in a binary system, which kicks the neutron star into its companion. In this picture, a high-velocity single neutron star with a planetary system is created from the remains of the former binary companion. One could thus speculate that the presence of planets around PSR B1257+12 is related to its unusually high proper motion.

Neutron star disks are clearly very different from those commonly found around pre-main-sequence stars. The disk is exposed to intense radiation and particle flux, close to or even above the Eddington luminosity of the neutron star ($\sim 10^{38}\,{\rm erg\,s^{-1}}$). The metallicity is very high, but initially there are no grains, and the temperature is well above the sublimation temperature of even the most refractory materials. The disk must therefore expand and cool before planets can be formed. Calculations of the evolution of such disks indicate that the formation of "terrestrial" planets such as those of the PSR B1257+12 system may indeed by possible, but the more massive and distant planet around PSR B1620−26 must have a different origin (Phinney and Hansen 1993).

The location of PSR B1620−26 near the core of the globular cluster M4 suggests that the pulsar acquired its planetary companion through an exchange interaction with a cluster star (Sigurdsson 1992, 1995). One plausible formation scenario begins with an old neutron star in a binary system, which interacts with a main-sequence star–planet system (Sigurdsson 1993). The original companion of the neutron star is ejected, while the main-sequence star and its planet are captured. The planet ends up in a wide orbit around the inner binary comprised of the neutron star and the main-sequence star. When the main-sequence star evolves to become a red giant, it transfers mass to the neutron star, spinning it up to become a millisecond pulsar. The chief difficulty of this scenario is the requirement that the age of the millisecond pulsar must be smaller than that of the triple system. However, the expected lifetime of the triple in the dense cluster core is of order $3 \cdot 10^7$ yr, while the estimated age of the binary pulsar is $\gtrsim 10^9$ yr (Ford et al. 2000). This scenario would thus require that the system, currently observed in projection near the edge of the cluster core, is in fact on an orbit that allows it to spend most of its lifetime in the far less dense cluster halo, and thus to escape disruption for a sufficiently long time.

An alternative formation scenario involves a dynamical exchange interaction between a pre-existing binary millisecond pulsar and a wide main-sequence star–planet system, in which the main-sequence star is ejected and the planet left in a wide orbit around the binary pulsar (Ford et al. 2000). Numerical simulations show that the probability of retaining the planet in the encounter is smaller than that of retaining the main-sequence star, but could still be as high as 10%. . . 30%. It is interesting to note that this scenario postulates the formation of a giant planet in a wide orbit around a “normal” star in a globular cluster environment; this is to be contrasted with the apparent absence of “hot Jupiters” in 47 Tucanae (see Sect. 6.4).

1.4 Overview of Planet Detection Methods

The Most Important Detection Techniques

Turning our attention to “normal” stars again, we will now look at the question how we might be able to find planets around them. Many different techniques have been proposed, in spite (or perhaps: because) of the difficulty of the task. The most promising strategies that are used in current detection efforts, or under development for use in the near future, are:

- Direct imaging of the star–planet system.
- Interferometric imaging of the star–planet system.
- Detection of the planetary spectrum in a composite spectrum of star and planet.

- Interferometric detection of the planetary spectrum through the wavelength dependence of the position of the photocenter of the star–planet system ("differential phase method").
- Photometry of planetary transits in front of the star.
- Spectroscopic detection of planetary transits.
- Photometric detection of the light reflected by a planet through its periodic variation with phase angle.
- Astrometric detection of the stellar motion around the star–planet center of mass.
- Radial-velocity measurement of the stellar motion around the star–planet center of mass.
- Imaging of circumstellar disks, which may show signatures of disk–planet interaction.
- Gravitational microlensing.
- Eclipse timing in binaries.

Each one of these techniques has unique strengths and weaknesses, and they vary widely in the information they can provide about the properties of the detected planets. Turning the question the other way around, we can take a list of characteristics that we would like to know about extrasolar planets, and ask which techniques can provide the requested information. Table 2 gives an overview of the most important planetary properties, and how they may be determined. More detailed discussions about the strengths and limitations of individual methods will be given in the subsequent sections. For the moment, the most important observation is that no single approach can give all the desired information; many complementary techniques will be needed to study the different aspects of extrasolar planets and planetary systems.

Typical Order-of-Magnitude Estimates

In order to understand the instrumental and observational requirements for the different planet detection techniques, we need to consider typical values for the potential observables. The large range in the properties of planets obviously implies a large difference in the difficulty to detect them. The Earth and Jupiter provide useful benchmarks (see Table 3), but one should also keep in mind that there are additional classes of planets, e.g., Uranus and Neptune in the Solar System, or the "hot Jupiters" orbiting their central stars at very small orbital radii. The properties of the host star, and the distance from the observer play important roles, too.

The chief difficulty of direct detection methods is the large contrast between the planet and its parent star at a very small angular separation. The reflex motion of the parent star due to the gravitational pull of the planet is very small, so that astrometry and the radial-velocity technique must reach extremely high precision to detect this effect. The photometric signature of transiting planets seems somewhat more easily accessible, at least for giant

Table 2. Important properties of planets, and techniques that can be used to determine them

property	technique	applicability
orbit	astrometry	++
	radial velocity	+
	direct imaging	○
mass	astrometry	++
	radial velocity	+
	microlensing	○
radius	transit photometry	++
radius, albedo	photometry of reflected light	++
radius, temperature	direct detection in mid-IR	++
surface features	photometry of reflected light	+
atmospheric composition	IR or visible spectroscopy	++
	transit spectroscopy	○
presence of moons	transit timing	+
system multiplicity	astrometry	+
	radial velocity	+

The symbols ++, +, and ○ denote how well the different methods can provide the required information

planets. However, in this case additional complications arise from the small probability that the orientation is such that transits actually occur. It is thus clear from the values listed in Table 3 that there is no "easy" technique for planet detection – this is the reason, of course, why it was not before 1995 that the first planet around a main-sequence star was discovered.

1.5 "Exotic" Concepts for Planet Detection

The subsequent chapters of this review will be devoted to introductions to some of the most promising planet detection techniques. Many more interesting approaches have been proposed, which deserve at least a brief description. It is entirely possible, of course, that one or the other of these "exotic" concepts will turn out to be more fruitful than some of the techniques that are considered "mainstream" today. The variety of physical effects that could in principle be observable should illustrate the diverse opportunities for the immediate and more distant future, and stimulate further ideas about possible ways to obtain more detailed information about planets during various phases of their life cycles, and about their interaction with the host stars.

Table 3. Typical values of observables for Jupiter-like and Earth-like planets

observable	Jupiter	Earth
angular separation	$0\farcs5$	$0\farcs1$
brightness contrast at visible $\lambda\lambda$	6×10^{-7}	1.5×10^{-10}
brightness contrast at $10\,\mu$m	1.5×10^{-7}	1.2×10^{-7}
astrometric amplitude	$500\,\mu$as	$0.3\,\mu$as
radial-velocity amplitude	$13\,\mathrm{m\,s^{-1}}$	$0.1\,\mathrm{m\,s^{-1}}$
transit probability	10^{-3}	5×10^{-3}
transit depth	1%	10^{-4}
transit duration	30 h	13 h
timing residuals	2.5 s	1.5 ms

The host star is assumed to be a Sun twin at a distance of 10 pc

Radio Emission from Extrasolar Planets

Five of the planets in the Solar System (Earth, Jupiter, Saturn, Uranus, and Neptune) produce non-thermal cyclotron radio emission, in a process that is thought to be driven by the Solar wind interacting with the planetary magnetospheres. The emission frequency is typically near the electron gyrofrequency in the magnetic field, i.e., of order 30 kHz . . . 30 MHz. The emission is very intense (at times, Jupiter is brighter than the Sun at frequencies below 20 MHz), and there exist fairly simple scaling laws that relate the observed radio power to the ram pressure of the Solar wind on the cross-sectional area of the magnetosphere (Zarka et al. 2001). There also exist scaling laws for the magnetic dipole moment of giant planets (Farrell et al. 1999); these scaling laws together can be used to predict the emitted radio power and peak frequency. In a few favorable cases the emission should be observable with current instruments, but no detections have been made so far (e.g., Bastian et al. 2000). This may either be due to the intermittent nature of the cyclotron emission, or to a smaller velocity or density of the stellar wind compared to the Solar wind, or to a smaller magnetic moment of the planet, due perhaps to tidal synchronization. In any case, future low-frequency arrays such as LOFAR or a Square Kilometer Array (Strom et al. 2001) should be able to observe the radio emission from magnetized giant planets.

Interaction-Induced Stellar Activity

Tidal or magnetic interaction between a giant planet in a short-period orbit and its host star might also increase the stellar activity, which could lead to variations in the shape of chromospheric lines in phase with the orbital period. Hints of systematic modulations of the Ca II H and K lines have been found in a few such systems, but they need further confirmation (Cuntz and Shkolnik 2002). It has also been speculated that very strong flares observed in some Solar-type stars could be due to magnetic reconnection between fields of the primary star and a close-in Jupiter-like planet (Rubenstein and Schaefer 2000). A systematic search for unusual flaring stars might thus be a new way of looking for planets, but a better understanding of the planet–star interaction would clearly be needed.

Young Planets Heated by Giant Impacts

The formation of planets proceeds through a phase of giant impacts (e.g., Wetherill 1990). The largest of these impacts may melt all or most of the surface of an Earth-size planetary embryo, and heat it to a temperature of about $1,500 \dots 2,500\,\mathrm{K}$. This would make its thermal emission detectable with a large ground-based interferometer (Stern 1994). Giant impacts on giant planets (such as the event that may have tipped the rotation axis of Uranus) will likely heat them to similar temperatures, making them even more easily detectable. The cooling times are of order a few hundred to several thousand years; this should make the number of impact-heated objects at any given time large enough to expect one detection per every few hundred pre-main-sequence stars surveyed. One would then still have to establish the planetary nature of the detected object, of course, and distinguish it from more massive companions or other possible interlopers.

Planets Swallowed by Giant Stars

As a Solar-mass star evolves off the main sequence, it expands and ascends the red giant branch of the Hertzsprung–Russell diagram. During that evolutionary phase, the star develops a large convective envelope with a radius of up to $\sim 100\,R_\odot$. Planets within that radius will be accreted by the star, and thus deposit energy, angular momentum, and elements such as Lithium in the stellar envelope. It has therefore been argued that the infrared excess (due to a substantial expansion of the star and ejection of a shell) and high Li abundance observed in $\approx 5\%$ of the G and K giants could be caused by the accretion of a giant planet or a brown dwarf (Gratton and D'Antona 1989; Siess and Livio 1999).

In an even later evolutionary stage, when the star becomes an asymptotic giant branch (AGB) star, it swells to an even larger size and develops a strong wind. Planets with even larger orbital radii can then get engulfed in the extended atmosphere, or interact with the wind flow. Episodic accretion of wind material on the planet may give rise to optical flashes and affect SiO maser emission (Struck et al. 2002). The details of these interactions are quite complex and poorly understood at the moment; this limits their potential use as a diagnostic tool and indicator for the presence of planets.

Planets Around White Dwarfs

It should be clear from the preceding paragraph that the orbits of planets can change drastically during the star's post-main-sequence evolution. Low-mass companions will spiral into the star due to the viscous and tidal forces exerted by the bloated atmosphere during the giant phase, but it might also be possible that some of them may be left in an orbit of radius $a \lesssim 1\,\mathrm{AU}$ around the ensuing white dwarf. This would be a very favorable situation for detecting the planet, because its radius would be ~ 10 times larger than that of the parent star! In the Rayleigh–Jeans portion of the combined spectrum (i.e., for observations at wavelengths much longer than that corresponding to the peak of the Planck function), the ratio of the total emission to that of just the white dwarf is given by

$$\frac{I_{\mathrm{tot}}}{I_{\mathrm{WD}}} = 1 + \frac{R_p^2 T_p}{R_{\mathrm{WD}}^2 T_{\mathrm{WD}}} \approx 1 + 100\,\frac{T_p}{T_{\mathrm{WD}}}\,. \tag{2}$$

For example, a planet with $T_p = 200\,\mathrm{K}$ orbiting a white dwarf with $T_{\mathrm{WD}} = 10,000\,\mathrm{K}$ would dominate the total emission of the system at long wavelengths.

Several groups have conducted near-infrared searches for substellar companions of white dwarfs, and some low-mass companions have been reported, but no planet has been discovered (e.g., Zuckerman and Becklin 1992). The above argument suggests, however, that searches should be conducted in the mid-infrared, where the planet can produce a strong excess over the white dwarf spectrum (Ignace 2001). The Spitzer (formerly SIRTF) infrared mission should have sufficient sensitivity to detect such planets out to a distance of $\sim 10\,\mathrm{pc}$.

Occultations by the Moon or Artificial Satellites

The planet detection schemes discussed in the previous few paragraphs intend to take advantage of special situations in which the the signature of the planet is not swamped by the nearby bright host star. In the general case, one may try to address the contrast problem by blocking the light from the star. This could either be achieved by using the dark limb of the Moon as an occulting edge (Elliot 1978), or by building a spacecraft carrying an occulting screen (Schultz et al. 1999, 2000; Copi and Starkman 2000). In either case, the observations would be carried out with a space telescope, which has to maintain alignment with the occulter to a precision of a fraction of an arcsecond (the typical angular separation of the planet from its parent star). The main obstacles for Lunar occultations are the rather large brightness of even the dark side of the Moon, and the difficulties of maneuvering the telescope. While it is possible to find orbits that give rather long ($\sim 1\,\mathrm{h}$) occultations of arbitrary stars, an enormous amount of propellant would have to be used to change targets.

Artificial occulters face similar problems. The diameter of the occulting screen clearly has to be larger than the telescope aperture, i.e., at least $\sim 10\,\mathrm{m}$.

To subtend an angle of no more than $0\rlap{.}''1$, it must therefore be placed at a separation of at least 20,000 km from the telescope. Furthermore, the intensity of the starlight in the shadow of the occulter is not zero; it must be computed with Fresnel's diffraction theory (e.g., Born and Wolf 1997).[4] Application of Babinet's Principle gives the approximation (Schultz et al. 1999)

$$\frac{I}{I_0} \approx \frac{16}{\pi^2} \cdot \frac{\lambda a}{D^2} = \frac{16}{\pi^2} \cdot \frac{\lambda}{\varphi D} \ , \tag{3}$$

where I and I_0 are the intensity in the presence and in absence of the occulting disk, λ the observing wavelength, a the distance between the occulter and the telescope, D the diameter of the occulter, and φ the angle subtended by the occulter as seen by the telescope. Equation (3) is valid for $D^2 \gg \lambda a$. For the above numbers ($D = 10\,\mathrm{m}$, $a = 20,000\,\mathrm{km}$) and $\lambda = 500\,\mathrm{nm}$, the on-axis intensity is still 16% of the value in the absence of an occulter. This shows that diffraction at the edge is a serious problem. With λ and φ fixed, and clear limits on the potential to increase D and a, the only viable way of obtaining a better starlight suppression is the use of a tapered occulter, i.e., a screen which is not completely opaque but has a transmission that continuously increases from 0 at the center to 1 at the edge (Copi and Starkman 2000). Manufacturing such a screen with precisely prescribed transmission function is a considerable technological challenge. This, together with the requirement of maneuvering the screen and telescope very precisely, has so far prevented serious consideration of this approach for a planet-detection mission.

A variation of the occultation idea is the use of a coronograph, which includes an occulting spot in the focal plane of the telescope. Compared to an external occulter, a coronograph has the disadvantage that the starlight is blocked only after passage through the telescope optics. The telescope therefore has to be built to very stringent specifications on wavefront quality and light scattering level. Nevertheless, this approach is currently regarded more promising than that of an external occulting screen.

1.6 The Search for Extraterrestrial Intelligence

The Drake Equation

Speculations about the possibility of life, of conscious beings, and of civilizations elsewhere in the Universe have a long history (Dick 1982, 1998). The search for extraterrestrial intelligence (SETI) as a scientific endeavor was born with the realization that our own technology had advanced to the point that radio signals could be transmitted and detected over interstellar distances

---

[4] One may recall Poisson's famous bright spot. Using Fresnel's theory, Poisson – who was very critical of that theory – predicted the seemingly absurd appearance of a bright spot behind a circular obstruction. This spot was almost immediately found experimentally by Arago, a great triumph of the wave theory of light.

(Cocconi and Morrison 1959). Soon the question was raised how many civilizations in the Galaxy might be engaged in attempts at communicating with each other, leading to the formulation of the famous *Drake Equation* (Drake 1962)

$$N = R_* \cdot f_p \cdot n_h \cdot f_l \cdot f_i \cdot f_c \cdot L \,. \tag{4}$$

The individual factors in this equation have the following meanings:

- N: the number of communicating civilizations in Galaxy;
- R_*: the rate of star formation in the Galaxy (expressed in stars per year);
- f_p: the fraction of stars that harbor planetary systems;
- n_h: the average number of planets (or moons) with conditions that are suitable for the genesis of life;
- f_l: the fraction of habitable planets on which life actually develops;
- f_i: the probability that evolution produces intelligent life;
- f_c: the fraction of intelligent civilizations that try to communicate over interstellar distances;
- L: the length of the communication phase (in years).

Unlike the other equations in this book, which (hopefully!) quantify our insights and knowledge, it is the main purpose of the Drake equation to organize our ignorance. We know that $R_* \approx 1\,\mathrm{yr}^{-1}$ (Trimble 1999), and we can now state fairly confidently that $f_p \geq 0.01$.[5] The determination of n_h is one of the great observational challenges for the coming ten to twenty years, as described extensively in this overview. With some luck, we might even be able to obtain an estimate for f_l from astronomical observations. This factor may also be amenable to biochemical experimentation in the tradition of the famous Miller–Urey experiments (Miller 1953) and modern attempts to generate synthetic life forms (Szostak et al. 2001). At present, in the absence of any evidence for extraterrestrial life, we have to admit that f_l could be anywhere between 10^{-9} and 1.

The next factor, f_i, is equally uncertain. Biologists are deeply divided about the question whether life necessarily evolves towards intelligence once it gets going. On the one hand, one may point out that a staggeringly improbable series of events has lead to the emergence of intelligent life on Earth (Gould 1989); on the other hand, one can argue that convergence is a ubiquitous property of life, which makes it likely that particular biological properties and features will sooner or later manifest themselves as part of the evolutionary process (Conway Morris 1998). In addition, we do not understand the biological basis of intelligence at all. What is the "quantum leap" that separates *homo sapiens* from *pan troglodytes*, the chimpanzee? Would *homo neanderthalensis* have become capable of constructing radio telescopes, if he

---

[5] This estimate is based on the number of planets detected in the Solar neighborhood (Sect. 3.1), with a "safety factor" applied for the possibility that the efficiency of planet formation may vary with the Galactic environment.

hadn't been displaced by a more advanced species? Finding an answer to these questions seems to be a key step towards a better estimate of f_i.

The factors f_c and L fall into the realm of sociology. It is tempting to speculate that $f_c \approx 1$, given the human drive for exploration, but we do not know with certainty that this extrapolation from our anthropocentric view of the world is really justified. The value of L depends on external factors such as global epidemics and giant impacts by comets and asteroids, and on internal factors that could lead to a quick end of a "semi-intelligent" civilization – wars or the exhaustion of natural resources. It appears possible that our own species and our offspring may populate the Earth at least for the remainder of the main-sequence lifetime of the Sun ($L \approx 5 \cdot 10^9\,\mathrm{yr}$), but if we are not careful, we may not live to see $L_{\mathrm{homo}} = 100\,\mathrm{yr}$.

We may thus characterize the emergent fields of exo-planetary astronomy and astrobiology as attempts to systematically explore the individual factors of the Drake Equation, from the left to the right. In contrast, SETI (which should perhaps better be called "Search for Extraterrestrial Technology" or "Search for Interstellar Communication") is an attempt to bypass this painstaking process by going directly for the grand prize. The chances of success are very uncertain, as the above arguments are consistent with estimates that range from an average distance between "neighbors" of $\approx 30\,\mathrm{pc}$, to a Galaxy that is void of life save that on a lonely, solitary Earth.

The Fermi Paradox

If civilizations are common in the Galaxy, one may ask the question why we have not found any incontrovertible evidence for their existence yet. More to the point, it has been argued that the absence of extraterrestrials from the Solar System implies that we are alone in the Galaxy, and that any searches for extraterrestrial civilizations are futile. The chain of arguments, which is known as *Fermi's Paradox*, goes as follows:

1. Lets's assume that our civilization is not the only one in our Galaxy that has developed technology.
2. Then our civilization must be "typical". This means that it is not the most advanced of all, and that other civilizations share our desire to explore.
3. Space travel is not too difficult for civilizations "slightly" more advanced than ours.
4. The time scale to colonize the whole Galaxy is $\lesssim 10^8$ yrs, i.e., small compared to the age of the Galaxy.
5. Then one must conclude that the Solar System should have been colonized a long time ago. But this is not the case.

So it appears that we have encountered a logical difficulty if we believe in the ubiquity of life. However, each step in this chain has potential loopholes, some of them more severe, others less.

The assumption in step (1.) can certainly be questioned. As explained in the previous section, we know very little about f_i, the likelihood for intelligence to emerge through evolution. If this factor is small, we may indeed be alone in the Galaxy.

Step (2.) seems to be quite plausible. The development of intelligence on Earth may have been a singular event, but if we assume that it has occurred in *one* other place, it very likely occurred in *many* other places. Then it is very unlikely that we are the most advanced civilization, given that there are many Solar-type stars (i.e., stars with comparable mass and metallicity) that are several Gyrs older. Life on a terrestrial planet around any of these stars would have a several-Gyr head start compared to the Earth. And assuming that *none* of these earlier civilizations would be interested in exploring the Galaxy (or that *all* of them would refrain from doing so for ethical reasons) seems extremely unlikely, too.

To justify step (3.), we can invoke some physical considerations. Several methods of attaining speeds necessary for interstellar travel ($v \gtrsim 0.1c$) have been suggested, including pulsed fusion and antimatter-powered rockets, light sails pushed by lasers, and interstellar ram jets (Crawford 1990). The biggest hurdle to overcome for interstellar travel are the enormous energy requirements; accelerating a spaceship to a substantial fraction of the speed of light in a reasonable time would require a few times the current global power production. This is a staggering power requirement, but it is plausible that it could be met by humanity very soon. If our power production grows at an average rate of only $\sim 1\%\,\mathrm{yr}^{-1}$, it will take less than 1,000 years, and the power requirement for interstellar travel will only be a fraction of a per cent of the global power consumption.[6] For comparison, a Saturn V rocket during lift-off consumed $\sim 0.5\%$ of the global power production. It thus seems likely that a civilization that is only slightly more advanced than ours will have the technical means to travel through interstellar distances.

The argument in step (4.) is based on the assumption that civilizations will establish colonies, and that each colony will again establish sub-colonies once it gets firmly established. With reasonable assumptions about the mean distance between these colonies, and about the time it takes for a colony to establish itself and to spawn a new settlement, one can estimate that it takes only a few Myrs to reach every habitable planet in the Galaxy (Crawford 2000).

Step (5.) also contains an assertion that can be questioned. While we have no scientifically valid evidence for the presence of other life forms in the Solar System, we do not have strong evidence for their absence, either. Observational limits on artificial probes within the Solar System are very weak, and small probes may even be hiding among us (Tough 2000).

---

[6] One should be somewhat careful with arguments based on sustained exponential growth, of course. At the same growth rate, humanity would need to generate more than $1\,L_\odot$ in less than 10,000 years.